

621.38
Щ 94

А. А. Щука

ЭЛЕКТРОНИКА

Под редакцией
профессора А. С. Сигова

Рекомендовано УМО по образованию в области радиотехники,
электроники, биомедицинской техники и автоматизации
для студентов высших учебных заведений, обучающихся
по направлению 654100 — Электроника и микроэлектроника

Санкт-Петербург
«БХВ-Петербург»

2005

Библиотека
«МАТИ» - РГТУ
им. К. Э. Циолковского

443.576

УДК 681.3(075.8)
ББК 32.85я73
Ш94

Шука А. А.

Ш94 Электроника. Учебное пособие / Под ред. проф. А. С. Сигова. — СПб.: БХВ-Петербург, 2005. — 800 с.: ил.

ISBN 5-94157-461-4

Рассмотрены разделы электроники: вакуумная и плазменная электроника, полупроводниковая электроника и микроэлектроника, оптическая и квантовая электроника, функциональная электроника. В содержание каждой части введена историческая справка об этапах становления и развития данного направления электроники. Каждый раздел содержит контрольные вопросы и задачи с решениями, некоторые из них выполнены с помощью стандартных компьютерных программ.

*Для студентов электронных, радиотехнических
и радиофизических специальностей вузов,
аспирантов и инженеров соответствующих специальностей*

УДК 681.3(075.8)
ББК 32.85я73

Группа подготовки издания:

Главный редактор	<i>Екатерина Кондукова</i>
Зам. главного редактора	<i>Людмила Еремеевская</i>
Зав. редакцией	<i>Григорий Добин</i>
Редактор	<i>Валентин Кайгородов</i>
Компьютерная верстка	<i>Ольги Сергиенко</i>
Корректор	<i>Зинаида Дмитриева</i>
Дизайн обложки	<i>Игоря Цырульникова</i>
Зав. производством	<i>Николай Тверских</i>

Рецензенты:

Заведующий кафедрой вакуумной электроники МФТИ,
доктор физико-математических наук, профессор, академик РАН,
лауреат Государственной премии А. С. Бугаев

Заведующий кафедрой электронных приборов и устройств СПбГЭТУ "ЛЭТИ",
доктор технических наук, профессор, заслуженный деятель науки и техники РФ,
лауреат Государственной премии РФ Ю. А. Быстров

Лицензия ИД № 02429 от 24.07.00. Подписано в печать 23.11.04.

Формат 70×100^{1/8}. Печать офсетная. Усл. печ. л. 64,5.

Тираж 3000 экз. Заказ № 3701.

"БХВ-Петербург", 190005, Санкт-Петербург, Измайловский пр., 29.

Гигиеническое заключение на продукцию, товар № 77.99.02.953.Д.001537.03.02
от 13.03.2002 г. выдано Департаментом ГСЭН Минздрава России.

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"

199034, Санкт-Петербург, 9 линия, 12

ISBN 5-94157-461-4

© Шука А. А., 2005
© Оформление, издательство "БХВ-Петербург", 2005
© ОМД Групп, 2005

Содержание

ПРЕДИСЛОВИЕ	15
ЧАСТЬ I. ВАКУУМНАЯ И ПЛАЗМЕННАЯ ЭЛЕКТРОНИКА	17
Введение	19
1. Краткая историческая справка	21
2. Физика и техника вакуума	27
2.1. Свойства вакуума	27
2.2. Методы создания вакуума	34
2.3. Методы измерения вакуума	38
2.4. Методы течения	49
Задачи и упражнения	52
Контрольные вопросы	56
Рекомендуемая литература	56
3. Вакуумная электроника	57
3.1. Модель прибора вакуумной электроники	57
3.2. Электронная эмиссия	58
3.2.1. Термоэлектронная эмиссия	59
3.2.2. Фотоэлектронная эмиссия	61
3.2.3. Вторичная электронная эмиссия	63
3.2.4. Кинетическая ионно-электронная эмиссия	64
3.2.5. Эмиссия горячих электронов	64
3.2.6. Экзоэлектронная эмиссия	65
3.2.7. Автоэлектронная эмиссия	65
3.2.8. Потенциальная ионно-электронная эмиссия	67
3.3. Эмиттеры свободных электронов	67
3.3.1. Электронная пушка	67
3.4. Управление потоком электронов полями	69
3.4.1. Движение электрона в однородных полях	69
3.4.2. Однородное магнитное поле	71
3.4.3. Движение электрона в скрещенных полях	75
3.5. Устройства управления электронным пучком	80
3.5.1. Электростатическая отклоняющая система	80
3.5.2. Управление с помощью электронной оптики	83
3.5.3. Управление с помощью магнитной оптики	87
3.6. Управление скоростью электронов	89
3.6.1. Резонаторные методы скоростной модуляции электронов	89
3.6.2. Нерезонансные устройства скоростной модуляции	91

3.7. Детектирование и преобразование энергии электронного потока	92
3.7.1. Наведение тока при движении электронов в вакууме	92
3.7.2. Отбор энергии от электронного потока	94
3.7.3. Процессы взаимодействия электронов с веществом детектора	95
Задачи и упражнения	98
Контрольные вопросы	108
Рекомендуемая литература	109
Приборы и устройства вакуумной электроники	110
4.1. Классификация приборов	110
4.2. Электронные лампы	110
4.3. Электровакуумные микролампы	117
4.4. СВЧ-приборы	118
4.4.1. Электронные лампы СВЧ	118
4.4.2. Клистроны	118
4.4.3. Лампы бегущей волны	121
4.4.4. Лампы обратной волны (ЛОБ)	122
4.4.5. Магнетроны	123
4.5. Электронно-лучевые приборы	126
4.5.1. Приборы типа "сигнал—свет"	126
4.5.2. Прибор типа "свет—сигнал"	128
Приборы с накоплением заряда	129
Приборы без накопления заряда	131
4.5.3. Приборы типа "сигнал—сигнал"	132
4.5.4. Приборы типа "свет—свет"	133
4.6. Фотоэлектронные приборы	134
4.6.1. Вакуумные фотоэлементы	134
4.6.2. Фотоэлектронные умножители	135
Задачи и упражнения	136
Контрольные вопросы	145
Рекомендуемая литература	145
Плазменная электроника	146
5.1. Основные понятия	146
5.2. Электрический разряд в газах	146
5.3. Процессы в плазме	149
5.4. Излучение плазмы	153
5.5. Диагностика плазмы	154
Задачи и упражнения	154
Контрольные вопросы	163
Рекомендуемая литература	163
Приборы и устройства плазменной электроники	164
6.1. Ионные приборы	164
6.2. Ионные приборы обработки и отображения информации	165
Задачи и упражнения	167
Контрольные вопросы	169
Рекомендуемая литература	169

7. Введение в зондовую микроскопию	170
7.1. Туннельная микроскопия	170
7.2. Физические явления на границе "острие — поверхность"	171
7.2.1. Туннельный ток	171
7.2.2. Силы атомных связей	174
7.3. Системный подход к созданию зондовых сканирующих микроскопов	175
7.3.1. Сенсоры	176
7.3.2. Пьезоэлектрический двигатель	177
7.3.3. Устройство управления (обратная связь)	178
7.3.4. Устройство обработки и индикации информации	179
7.4. Возможности зондовой микроскопии	179
Контрольные вопросы	181
Рекомендуемая литература	181
Заключение	182
ЧАСТЬ II. МИКРОЭЛЕКТРОНИКА	183
Введение	185
1. Микроэлектроника вчера и сегодня	186
1.1. Краткая историческая справка	186
1.1.1. Одна отгадка на все разгадки	188
1.1.2. Эпоха транзисторизации	190
1.1.3. Вторая транзисторная революция	194
1.1.4. Третья транзисторная революция — рождение микроэлектроники	196
1.1.5. Грядет ли новая транзисторная революция?	200
1.1.6. Линии развития, параллельные транзистору	201
1.2. Место микроэлектроники в сфере высоких технологий	206
2. Физика полупроводниковых структур	209
2.1. Полупроводники и их структура	209
2.2. Носители заряда в полупроводниках	211
2.3. Перенос носителей и управление переносом	217
2.4. Барьеры и контакты	219
2.4.1. Барьеры на границе кристалла	219
2.4.2. Электронно-дырочные переходы	221
2.4.3. Гетеропереходы	227
2.4.4. Контакты	229
2.4.5. Граница раздела	231
Задачи и упражнения	232
Контрольные вопросы	245
Рекомендуемая литература	246
3. Интегральные транзисторные структуры	247
3.1. Классификация транзисторных структур	247
3.2. Интегральные униполярные (полевые) транзисторы	248
3.2.1. МДП-транзисторы с индуцированным каналом	248
3.2.2. МДП-транзисторы со встроенным каналом	250

3.2.3. Комплементарные структуры.....	251
3.2.4. Транзистор с управляющим р—п-переходом	252
3.2.5. Полевой транзистор на гетероструктурах	254
3.2.6. V-MДП-транзисторы.....	257
3.3. Интегральный биполярный транзистор.....	257
3.3.1. Эпитаксиально-планарный транзистор.....	257
3.3.2. Физические основы работы биполярного транзистора	259
3.3.3. Малосигнальные параметры.....	260
3.3.4. Эпитаксиально-планарные п—р—п-транзисторы	263
3.3.5. Многоэмиттерные и многоколлекторные транзисторы	264
3.3.6. Транзисторные структуры интегрально-инжекционной логики	265
3.3.7. Транзистор с барьером Шоттки	265
3.4. Перспективные транзисторные структуры.....	266
3.4.1. Транзисторы со сверхтонким основанием.....	267
3.4.2. МОП-транзисторы с управляемой проводимостью канала.....	268
3.4.3. МОП-транзисторы с двойным затвором.....	268
3.4.4. Вертикальные МОП-транзисторы.....	269
3.4.5. Арсенид-галлиевые транзисторы	270
Задачи и упражнения.....	272
Контрольные вопросы.....	283
Рекомендуемая литература	284
Элементная база интегральных схем	285
4.1. Изоляция элементов	285
4.2. Интегральные диоды	288
4.3. Интегральные резисторы	291
4.4. Интегральные конденсаторы	293
Задачи и упражнения.....	297
Контрольные вопросы.....	300
Рекомендуемая литература	300
Технология производства интегральных схем.....	301
5.1. Технологические процессы изготовления ИС.....	301
5.2. Процессы первичной обработки материалов	301
5.3. Процессы литографии	304
5.3.1. Фотолитография	305
5.3.2. Электронолитография.....	308
5.3.3. Рентгенолитография.....	311
5.3.4. Ионная литография	314
5.3.5. Лазерная литография.....	315
5.4. Процессы локального изменения свойств полупроводников	316
5.4.1. Эпитаксия	316
5.4.2. Легирование полупроводников	319
Высокотемпературное легирование.....	320
Ионная имплантация	323
Радиационно-стимулированная диффузия	326
Лазерный отжиг.....	327
5.5. Процессы обработки поверхности	328
5.5.1. Окисные пленки кремния.....	328

5.5.2. Травление	329
Ионное травление	330
Ионно-химическое травление	331
Плазмохимическое травление	332
Лазерно-стимулированное травление	333
5.5.3. Металлизация поверхности	334
5.6. Сборка интегральных схем	338
5.7. Типовые технологические маршруты производства интегральных схем	339
5.8. Особенности производства интегральных схем	355
Задачи и упражнения	356
Контрольные вопросы	364
Рекомендуемая литература	364
6. Интегральные схемы и их классификация	365
6.1. Классификация интегральных схем	365
6.2. Условные обозначения	371
6.3. Основные параметры интегральных схем	374
6.4. Применение и эксплуатация интегральных схем	378
Рекомендуемая литература	379
7. Логические элементы интегральных схем	380
7.1. Классификация логических элементов	380
7.2. Основные характеристики логических элементов	383
7.3. Логические ИС на биполярных транзисторах	385
7.3.1. Логические элементы с передачей тока или напряжения	385
Транзисторная логика с непосредственными связями	385
Резисторно-транзисторная логика	386
Резисторно-емкостная транзисторная логика	387
Интегральная инжекционная логика	387
7.3.2. Логические элементы с логикой на входе	389
Диодно-транзисторная логика	389
Транзисторно-транзисторная логика	391
Транзисторно-транзисторная логика с диодами Шоттки	391
7.3.3. Логические схемы на переключателях тока	391
Эмиттерно-связанная логика	391
7.4. Логические элементы, реализованные на МДП-транзисторах	393
7.4.1. Схемы на транзисторах с каналами одного типа проводимости	393
7.4.2. Схемы на комплементарных транзисторах	395
7.4.3. Логические элементы на арсенид-галлиевых транзисторах	397
7.5. Логические элементы на БиКМОП-транзисторах	399
7.6. Сравнительный анализ логических элементов	400
Задачи и упражнения	401
Контрольные вопросы	404
Рекомендуемая литература	405
8. Запоминающие устройства	406
8.1. Классификация запоминающих устройств	406
8.2. Запоминающие устройства на биполярных транзисторах	408
8.3. Запоминающие устройства на МДП-транзисторах	410
8.4. Запоминающие устройства на арсенид-галлиевых структурах	414

Задачи и упражнения.....	415
Контрольные вопросы.....	420
Рекомендуемая литература.....	420
Триггеры	421
9.1. Бистабильные ячейки.....	421
9.2. Триггер Шмитта.....	422
9.3. RS-триггер.....	423
9.4. RST-триггер.....	424
9.5. D-триггер.....	424
9.6. T-триггер.....	425
9.7. JK-триггер.....	426
Задачи и упражнения.....	427
Контрольные вопросы.....	427
Рекомендуемая литература.....	428
0. Цифровые устройства на основе триггеров.....	429
10.1. Счетчики.....	429
10.2. Регистры сдвига.....	430
10.3. Сумматоры.....	430
10.4. Цифраторы и дешифраторы.....	432
10.5. Мультиплексор.....	434
Задачи и упражнения.....	435
Контрольные вопросы.....	436
Рекомендуемая литература.....	436
11. Микропроцессоры и компьютеры.....	437
11.1. Микропроцессоры.....	437
11.2. Мировой рынок микропроцессоров.....	439
11.3. Микропроцессорные системы.....	441
Задачи и упражнения.....	445
Контрольные вопросы.....	446
Рекомендуемая литература.....	446
12. Аналоговые интегральные схемы.....	447
12.1. Классификация аналоговых схем.....	447
12.2. Операционный усилитель.....	449
12.2.1. Дифференциальный усилитель.....	449
12.2.2. Структурная схема операционного усилителя.....	450
12.2.3. Аппаратурные включения операционных усилителей.....	452
Линейные включения ОУ.....	452
Нелинейные включения операционного усилителя.....	455
12.3. Преобразователи типа ЦАП — АЦП.....	457
12.3.1. Цифроаналоговый преобразователь.....	457
12.3.2. Аналого-цифровой преобразователь.....	457
Задачи и упражнения.....	458
Контрольные вопросы.....	459
Рекомендуемая литература.....	459

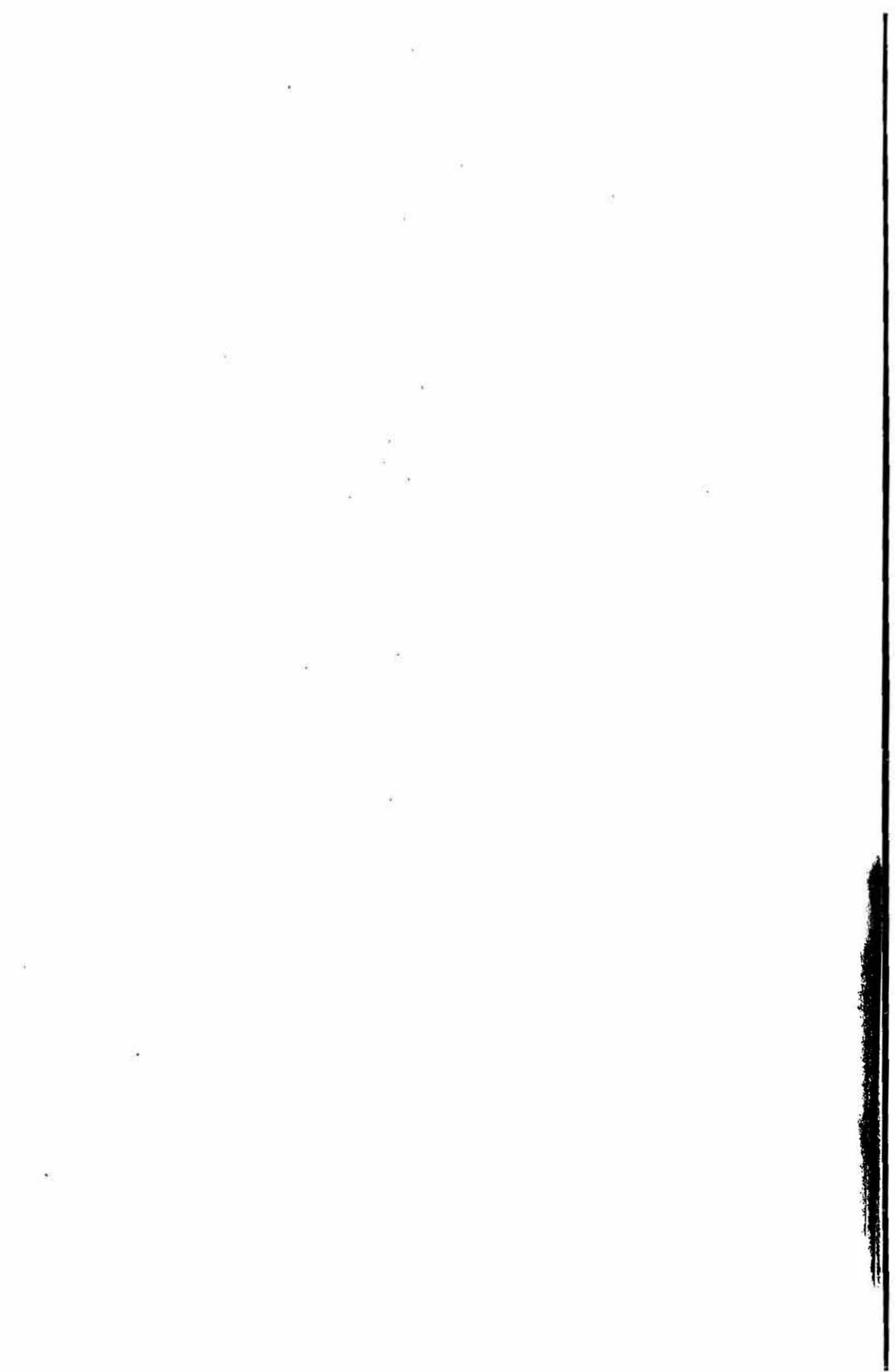
13. Базовые матричные кристаллы	460
13.1. Классификация базовых матричных кристаллов	460
13.2. Программируемые логические интегральные схемы	464
Контрольные вопросы	466
Рекомендуемая литература	466
14. Интегральные схемы СВЧ-диапазона	467
14.1. Общие положения	467
14.2. Элементная база электроники СВЧ	468
14.3. Интегральные транзисторы СВЧ-диапазона	471
14.4. Монолитные арсенид-галлиевые ИС	473
Контрольные вопросы	474
Рекомендуемая литература	474
15. От микро- к нанoeлектронике	475
15.1. Барьеры на пути перехода от микро- к нанoeлектронике	475
15.2. Начала нанoeлектроники	476
15.3. Квантовые ограничения	478
15.4. Туннелирование электронов	481
15.5. Квантовые точки	483
15.6. Квантовые эффекты в наноструктурах	484
15.6.1. Одноэлектронное туннелирование	484
15.6.2. Транспортный эффект Ааронова — Бома	486
15.7. Новые транзисторные структуры	487
15.7.1. Полевые транзисторы	488
15.7.2. Транзисторы с резонансным туннелированием	489
15.8. Квантовые приборы нанoeлектроники	490
15.8.1. Квантовый вентиль	491
15.8.2. Квантовый интерферометр	492
15.8.3. Квантовый каскадный лазер	492
15.9. Одноэлектронные приборы	493
15.10. Новые материалы нанoeлектроники	497
Контрольные вопросы	499
Рекомендуемая литература	499
Заключение	500
ЧАСТЬ III. КВАНТОВАЯ И ОПТИЧЕСКАЯ ЭЛЕКТРОНИКА	501
Краткая историческая справка	503
1. Физические основы квантовой электроники	511
1.1. Спонтанное и вынужденное излучение	511
1.2. Спектральные линии	514
1.3. Поглощение и усиление	515
Задачи и упражнения	517
Контрольные вопросы	524
Рекомендуемая литература	524

Принципы работы лазера	525
Задачи и упражнения	530
Контрольные вопросы	534
Рекомендуемая литература	535
Типы лазеров	536
3.1. Лазеры на основе конденсированных сред	536
3.1.1. Твердотельные лазеры	536
3.1.2. Полупроводниковые лазеры	538
3.1.3. Жидкостные лазеры	542
3.2. Газовые лазеры	544
3.2.1. Лазеры на нейтральных атомах	544
3.2.2. Ионные лазеры	546
3.2.3. Молекулярные лазеры	547
3.2.4. Экимерные лазеры	549
3.2.5. Газодинамические лазеры	550
3.2.6. Химические лазеры	551
3.2.7. Лазеры на парах металла	552
3.2.8. Лазеры на свободных электронах	553
Задачи и упражнения	554
Контрольные вопросы	558
Рекомендуемая литература	558
4. Оптические волноводы	559
4.1. Плоские волноводы	559
4.2. Волоконные световоды	565
Контрольные вопросы	567
Рекомендуемая литература	568
5. Устройства управления световыми потоками	569
Контрольные вопросы	573
Рекомендуемая литература	573
6. Приемники излучения	574
6.1. Фотодиоды	574
6.2. Фотоприемники с внутренним усилением	577
6.3. Гетеродинный прием оптического излучения	579
Контрольные вопросы	581
Рекомендуемая литература	582
7. Введение в интегральную оптику	583
7.1. Элементы интегральной оптики	583
7.2. Устройства и элементная база интегральной оптики	584
7.3. Интегрально-оптические схемы	587
Контрольные вопросы	589
Рекомендуемая литература	590
8. Введение в оптоэлектронику	591
8.1. Элементная база	591
8.1.1. Некогерентные излучатели	591

8.1.2. Когерентные излучатели	593
8.1.3. Приемники излучения	595
8.2. Оптоэлектронные устройства обработки информации	595
8.2.1. Оптроны	595
Контрольные вопросы	596
Рекомендуемая литература	596
9. Оптические методы обработки информации	597
9.1. Оптические сигналы	597
9.2. Голография	599
9.2.1. Принципы голографической обработки информации	599
9.2.2. Голографическая элементная база	602
9.2.3. Голографические запоминающие устройства	602
9.2.4. Интерферометрические методы	607
9.3. Когерентные оптические системы аналоговой обработки информации	609
Контрольные вопросы	613
Рекомендуемая литература	613
ЧАСТЬ IV. ФУНКЦИОНАЛЬНАЯ ЭЛЕКТРОНИКА	615
Краткая историческая справка	617
1. Начала функциональной электроники	621
Контрольные вопросы	625
Рекомендуемая литература	626
2. Функциональная акустоэлектроника	627
2.1. Физические основы	627
2.1.1. Динамические неоднородности	627
2.1.2. Континуальные среды	631
2.1.3. Генераторы динамических неоднородностей	633
2.1.4. Устройство управления динамическими неоднородностями	636
2.1.5. Детектирование динамических неоднородностей	638
2.2. Приборы функциональной акустоэлектроники	639
2.2.1. Линии задержки	640
2.2.2. Устройства частотной селекции	645
2.2.3. Генераторы на ПАВ	653
2.2.4. Усилители	654
2.3. Нелинейные устройства	656
2.3.1. Физические основы	656
2.3.2. Конвольверы	659
2.3.3. Устройства памяти	662
Акустическое запоминающее устройство	662
Экзотические устройства акустической памяти	666
2.3.4. Фурье-процессоры	667
2.4. Акустоэлектроника в системах и средствах связи	669
Контрольные вопросы	670
Рекомендуемая литература	670

Функциональная диэлектрическая электроника	671
3.1. Физические основы	671
3.1.1. Динамические неоднородности	671
3.1.2. Континуальные среды	672
3.1.3. Генераторы динамических неоднородностей	676
3.1.4. Другие элементы приборов	677
3.2. Приборы и устройства функциональной диэлектрической электроники	677
3.2.1. Слоистые структуры	677
3.2.2. Устройства памяти	678
3.2.3. Процессоры	681
Контрольные вопросы	683
Рекомендуемая литература	683
„ Функциональная полупроводниковая электроника	684
4.1. Физические основы	684
4.1.1. Динамические неоднородности в полупроводниках	684
4.1.2. Континуальные среды	688
4.1.3. Генераторы динамических неоднородностей	689
4.1.4. Устройства управления динамическими неоднородностями	695
4.1.5. Детекторы динамических неоднородностей	699
4.2. Приборы и устройства функциональной полупроводниковой электроники	700
4.2.1. Аналоговые процессоры на ПЗС-структурах	700
4.2.2. Цифровые процессоры на ПЗС-структурах	705
4.2.3. Запоминающие устройства на ПЗС-структурах	709
4.2.4. БИСПИН-приборы	712
4.2.5. Приборы на волнах пространственного заряда	717
4.2.6. Гапновские приборы	720
Контрольные вопросы	722
Рекомендуемая литература	722
5. Функциональная магнитоэлектроника	724
5.1. Физические основы	724
5.1.1. Динамические неоднородности в магнитоэлектронике	724
5.1.2. Континуальные среды	728
5.1.3. Генерация, детектирование и управление динамическими неоднородностями	728
5.2. Приборы и устройства функциональной магнитоэлектроники	731
5.2.1. Процессоры сигналов на ЦМД	731
5.2.2. Процессоры сигналов на МСВ	732
5.2.3. Запоминающие устройства на ЦМД	734
5.2.4. Запоминающие устройства на магнитных вихрях	736
Контрольные вопросы	737
Рекомендуемая литература	738
6. Функциональная оптоэлектроника	739
6.1. Физические основы	739
6.1.1. Динамические неоднородности оптической природы	739
6.1.2. Континуальные среды	741
6.1.3. Генераторы и детекторы	742
6.1.4. Устройства управления	742

6.2. Процессоры функциональной оптоэлектроники.....	743
6.3. Заломинающие устройства функциональной оптоэлектроники.....	745
Контрольные вопросы.....	747
Рекомендуемая литература	747
7. Функциональная молекулярная электроника.....	748
7.1. Физические основы	748
7.1.1. Динамические неоднородности.....	750
7.1.2. Континуальные среды	751
7.1.3. Другие элементы приборов.....	753
7.2. Молекулярные устройства.....	753
7.3. Автоволновая электроника	755
Контрольные вопросы.....	757
Рекомендуемая литература	757
8. Приборы функциональной электроники второго поколения.....	758
8.1. Приборы с акустическим переносом зарядов.....	758
8.2. Приборы акустооптики	760
Контрольные вопросы	763
Рекомендуемая литература	763
ЗАКЛЮЧЕНИЕ.....	765
ПРИЛОЖЕНИЕ	767
Вехи развития отечественной электроники.....	767
Российский период	767
Советский период	771
Новейший российский период.....	789
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ	792



ПРЕДИСЛОВИЕ

Подготовка специалистов по направлению "Электроника и микроэлектроника" затруднена, с одной стороны, обилием ранее изданных книг по этому направлению и трудностью выбора современного учебника или учебного пособия. С другой стороны, отсутствует учебник, объединивший с системных позиций вакуумную и плазменную электронику с твердотельной и микроэлектроникой, квантовую и оптическую электронику с функциональной электроникой. Все эти дисциплины взаимно связаны и являют собой физикотехнические основы создания элементной базы информационных устройств и систем.

Автор понимает, что книгу по электронике можно писать, писать несколько раз, но написать невозможно. Электроника является динамично развивающейся областью науки и техники.

С точки зрения науки в основу положены свойства носителя информационного сигнала. Первым таким носителем можно считать электрон. Он является наименьшим материальным носителем электрического заряда, обладает собственным механическим и магнитным моментом, обладает большим временем жизни, достаточно легко выводится из твердого тела, может группироваться в потоки. В последние годы используются не только свойства электрона как частицы, но и свойства электрона как волны. Использование свободных электронов как носителей информационного сигнала легло в основу вакуумной электроники и созданию класса электровакуумных приборов. Исследования процессов ионизации газов с помощью электронов привело к становлению плазменной электроники и формированию класса ионных или плазменных приборов.

Становление и развитие физики твердого тела позволило сформулировать условия использования свободных электронов и квазичастиц в твердом теле для создания твердотельных электронных приборов. На основе твердотельной электроники сформировались полупроводниковая электроника и микроэлектроника.

Исследования процессов вынужденного излучения привело к формированию и становлению новой области электроники — квантовой электроники. В ее основе лежат свойства связанных в атоме электронов, их коллективные взаимодействия с веществом и излучением.

Однако не только электроны, ионы, дырки и другие квазичастицы способны переносить информационный сигнал. Существуют так называемые динамические неоднородности, на свойствах которых можно создать класс приборов для обработки и хранения информации. К таким динамическим неоднородностям можно отнести поверхностные акустические волны, магнитно-статические волны, волны зарядовой плотности, ганновские и сегнетоэлектрические домены, флюксоны и другие квазичастицы. Это направление в электронике получило название функциональной электроники.

Свойства электрона как волны используется в нанoeлектронике, глава по которой еще не представлена в этой книге.

тесте с тем, все большее внимание привлекает к себе солитон. Он тоже обладает свойствами волны и частицы. В каждой среде у солитона свое лицо.

ожет быть, со временем это направление в электронике получит название солитонная электроника. А может быть просто — солитоника...

точки зрения техники все, что связано с электроникой, относится к высоким технологиям.

лектроника сегодня — это совокупность технологий.

начале прошлого века начала развиваться вакуумная техника. Именно успехи в вакуумной технологии позволили развиваться технологии микроэлектроники. Появилась квантовая электроника, и успехи лазерной технологии во многом определялись достижениями вакуумной технологии. Сегодня на очереди стоит развитие нанотехнологий.

этих позиций и написаны различные разделы учебника:

- 1 Часть I. Вакуумная и плазменная электроника.
- 2 Часть II. Полупроводниковая электроника и микроэлектроника.
- 3 Часть III. Квантовая и оптическая электроника.
- 4 Часть IV. Функциональная электроника.

То мысли автора учебник должен сформировать у студентов, изучающих электронику, целостное восприятие этого перспективного раздела науки и техники.

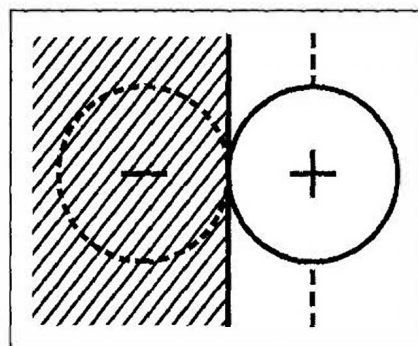
удалось ли это, судить студентам и аспирантам, для которых и предназначена эта книга.

учебник написан по материалам лекций, прочитанных автором в разные годы в Московском государственном институте радиотехники, электроники и автоматики (технический университет) и Московском физико-техническом институте (государственный университет). Автор с благодарностью вспоминает обсуждения отдельных тем и методических подходов с выдающимися лекторами профессорами Курбатовым Л. Н., Скроцким Г. В., Гуляевым Ю. В., Бондаренко Б. В., Федотовым Я. А. и др. Их лекции и семинары так или иначе органически вплелись в содержание книги.

Вместе с тем автор обязан замечательным учебникам и учебным пособиям Розанова Л. Н., Степаненко И. П., Коледова Л. А., Карлова Н. В., Пихтина А. Н. и других замечательных отечественных ученых, методические находки которых были использованы в учебном процессе и процитированы в книге. Упражнения адаптированы из популярных задачников Гольдсберга Л. М., Левитского С. М., Линча П., Терехова В. А., Успенского А. В. и других крупных методистов.

В подготовке рукописи к изданию активное участие принимали Федотов Я. А., Горбунова Г. Г., Звзрев Г. М., Долгалева Д. Г., Бура Г. Ф., Щука С. А., Поляхов Ю. Б. и др., которым приношу сердечную благодарность.

У автора нет сомнений, что в первое издание могли попасть неточности, опечатки, некорректные формулировки. С благодарностью будут приняты все замечания по адресу: 119454, Москва, Проспект Вернадского, 78, МИРЭА, кафедра интегральной электроники.



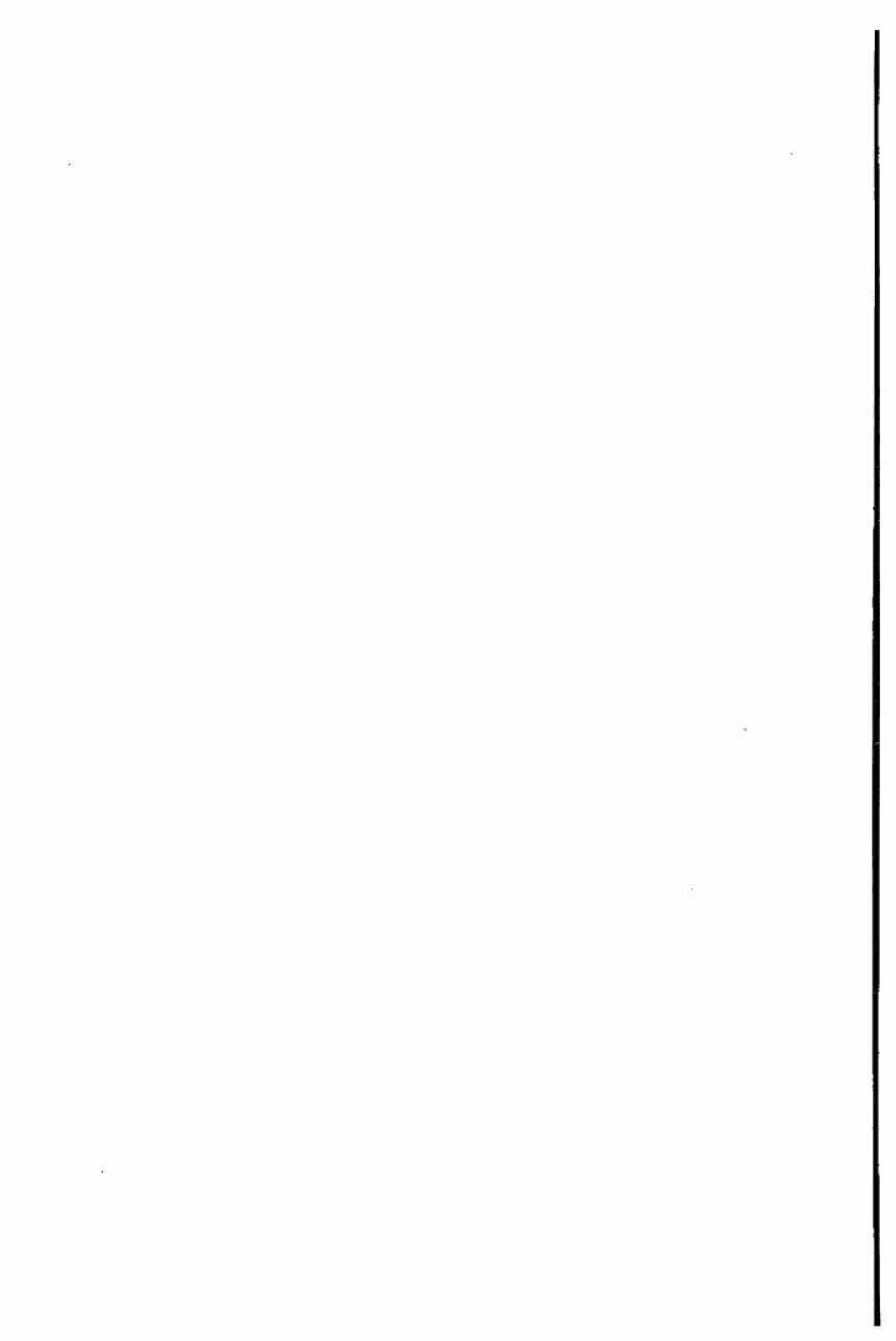
ЧАСТЬ I

Вакуумная и плазменная электроника

Введение

1. Краткая историческая справка
2. Физика и техника вакуума
3. Вакуумная электроника
4. Приборы и устройства вакуумной электроники
5. Плазменная электроника
6. Приборы и устройства плазменной электроники
7. Введение в зондовую микроскопию

Заключение



Введение

Материал части I, как и всего учебника, излагается с позиции системного анализа. В практику преподавания электроники введено понятие модели электронного прибора, определены его основные элементы на основе анализа определенного класса электронных приборов. Такой методический подход позволяет системно изложить физические основы процессов, лежащих в основе работы электронных приборов, произвести их классификацию, привести основные конструктивные решения и параметры электронных приборов. Системный подход, примененный к изложению взаимосвязанных физических явлений и процессов, относящихся к разным разделам электроники, позволяет использовать изложенный материал и для изучения физических основ технологии производства электронных приборов, играющую определяющую роль в электронике.

Автор посчитал важным в каждой части учебника привести краткий исторический раздел, в котором отражены основные вехи развития определенной области электронной техники, приведены имена ученых, включая наших отечественных, внесших решающий вклад в развитие электроники.

В каждом разделе предлагаются задачи, которые снабжены решениями. Это классические, ранее апробированные задачи, которые могут использоваться для проведения индивидуальных занятий, а также для самостоятельной работы студентов. Знакомясь с решением задач, студент может самообразовываться, самостоятельно разбираться в отдельных вопросах курса лекций. Контрольные вопросы могут использоваться при составлении экзаменационных билетов.

Первая часть учебника начинается с изложения свойств вакуума как континуальной среды распространения электронов и ионов, способов получения и измерения вакуума.

Системное изложение вакуумной электроники начинается с физической модели приборов вакуумной электроники. Она состоит из следующих элементов:

- эмиттер электронов;
- континуальная среда (вакуум), в которой распространяются электроны;
- устройства управления электронными потоками, формирующие информационный сигнал с использованием электрических и магнитных полей;
- детектор информационного сигнала.

В соответствии с такой концепцией рассмотрены физические процессы эмиссии электронов и, соответственно, основные типы катодов, электронных пушек, формирующие электронные пучки.

Для управления этими пучками предназначены устройства, в основе которых лежат электронная и магнитная оптика. Особое внимание акцентируется на процессах управления скоростью электронов в потоке, позволяющих проводить группировку электронов.

Для съема информации или вывода энергии анализируются процессы взаимодействия электронов с веществом детектора, процессы наведения токов при движении электронов в вакууме.

основе изученных физических процессов и явлений рассматриваются принципы работы и характеристики различных типов приборов вакуумной электроники: электронных, СВЧ-приборов, электронно-лучевых приборов, фотоэлектронных и рентгеновских приборов.

в книге рассмотрены некоторые перспективные направления развития вакуумной электроники, в частности, ее взаимодействие с нанoeлектроникой. Это, прежде всего, использование явления автоэлектронной эмиссии в туннельных и силовых микроскопах.

в последней главе рассмотрены основные типы ионных или плазменных приборов, принцип работы которых основан на использовании газового разряда и процессов ионизации помощью электронов.

1. Краткая историческая справка

Понятие *вакуум* (vacuum), в переводе с латинского — пустота, обозначает состояние газа при давлении меньше атмосферного. До середины XVII века этот термин использовался лишь в философии. Древнегреческий философ Демокрит одним из "начал мира" считал именно пустоту. Позднее Аристотель ввел понятие *эфира* — неосязаемой среды, способной передавать давление. В этот период знания о свойствах разреженного газа еще отсутствовали, но вакуум уже широко использовался в водоподъемных и пневматических устройствах.

В 1211 году в уставе первого в мире Парижского университета было записано, что вопросы пустоты должны заниматься богословы, но ни в коем случае не естественники. Только в XVII веке Галилео Галилей, не терпевший церковных догм, вычислил "силу боязни пустоты" из факта невозможности подъема воды всасывающим насосом на высоту более 10 м, какого бы диаметра труба ни была.

Научный этап развития вакуумной техники начинается с 1643 года, когда в Италии Эванджелиста Торричелли, ученик знаменитого Галилео Галилея, измерил атмосферное давление. Он открыл, что атмосфера создает давление, равное давлению столба ртути высотой около 760 мм. Пространство над ртутью в барометрической трубке, которое по представлению Торричелли было "абсолютной пустотой", названо в честь ученого "торричеллиевой пустотой". Теперь известно, что это пространство заполнено ртутным паром, имеющим при нормальной температуре давление $1,6 \times 10^{-1}$ Па ($1,2 \times 10^{-3}$ мм рт. ст.).

В середине XIX века немецкий химик Роберт Вильгельм Бунзен осуществил откачку газа струей быстро истекающей жидкости, увлекающей газ из изолированного сосуда.

В научную дисциплину вакуумная техника сформировалась в связи с развитием производства электровакуумных приборов.

В 1873 году русский ученый электротехник Александр Николаевич Лодыгин изобрел *первый электровакуумный прибор* — электрическую лампу накаливания с угольным стержнем.

В 1883 году американский изобретатель Томас Алва Эдисон, исследуя различные вещества для электрических ламп накаливания, открыл явление термоэлектронной эмиссии из нагретых проводников.

В 1887 году немецкий физик Генрих Рудольф Герц открыл внешний фотоэлектрический эффект. Одновременно русский физик Александр Григорьевич Столетов систематически исследовал внешний фотоэлектрический эффект и открыл первый закон фотоэффекта. Эти, а также многие другие открытия и изобретения привели к стремительному развитию вакуумной техники и вакуумной электроники.

В 1884 году итальянец А. Малиньяни впервые применил в процессе производства вакуумных ламп накаливания связывание остаточных газов парами фосфора и тем самым по-

жил начало применению различного рода газопоглотителей или так называемых *геттеров* в электровакуумных приборах. Впоследствии были разработаны геттерные насосы. В 1901 году русский физик Петр Николаевич Лебедев в своих опытах впервые использует идею удаления остаточных газов с помощью ртутного пара.

В 1904 году Джеймс Дьюар разработал способ получения вакуума путем поглощения азота активированным углем, охлаждаемым жидким азотом.

В 1905 году был изготовлен первый вращательный механический насос Геде, который позволил получить давления около 10^{-2} Па ($\sim 10^{-4}$ мм рт. ст.). В 1914—1916 гг. практически одновременно американский физик Ирвинг Ленгмюр и профессор Петроградского университета С. А. Боровик разработали ртутный диффузионный насос, способный создавать давление около 10^{-6} Па ($\sim 10^{-7}$ мм рт. ст.). За небольшой период времени в начале XX века были изобретены широко применяемые в настоящее время вакуумные насосы: вращательный (Геде, 1905), криосорбционный (Д. Дьюар, 1906), молекулярный (Геде, 1912), диффузионный (Геде, 1913).

В развитии техники измерения низких давлений следует отметить такие работы, как создание компрессионного манометра Г. Мак-Леодом (1874 г.), теплового манометра М. Пинни (1909 г.) и ионизационного манометра О. Бакли (1916 г.).

Успехам вакуумной техники способствовали многочисленные теоретические и экспериментальные работы. Здесь особенно следует отметить работы, проведенные в первые десятилетия XX века американскими учеными И. Ленгмюром и С. Дэшманом, английским ученым Н. Р. Кэмпбеллом и датским физиком М. Кнудсенем.

В СССР становление вакуумной техники связано с именем академика Сергея Аркадьевича Векшинского, организовавшего в 1928 году вакуумную лабораторию в Ленинграде, а затем возглавившего научно-исследовательский вакуумный институт в Москве.

В 1928 году появился паромасляный диффузионный насос К. Р. Бэрча, а затем большое количество других (механических, пароструйных, молекулярных) насосов, которые широко используются и сегодня в вакуумной технике и постоянно совершенствуются.

Для получения сверхвысокого вакуума были изобретены новые насосы: турбомолекулярный (Беккер, 1958), магниторазрядный (Джеспен и Холанд, 1959); были усовершенствованы паромасляные и криосорбционные насосы.

При измерении низких давлений стали использоваться давления каждой компоненты остаточных газов. Для обеспечения надежной сборки и эксплуатации сверхвысоковакуумных систем потребовалась разработка чувствительных методов определения течи в вакуумных системах: масс-спектрометрического, галоидного и других типов. Для снижения газовой нагрузки вакуумных конструкционных материалов начал применяться высокотемпературный прогрев всей вакуумной установки. Вакуумные системы стали изготавливаться целиком из сталей, интенсивно разрабатывались конструкции сверхвысоковакуумных уплотнений, вводов движения и электрических вводов в вакуум. Были усовершенствованы технологические методы получения неразъемных соединений металла со стеклом, электронно-лучевая и газовая сварка. В 60-х годах прошлого века успешное развитие вакуумной техники позволило разработать стохастические методы расчета вакуумных систем. Достижения криогенной техники в получении низких температур нашли применение в технологии получения вакуума. Криогенные вакуумные насосы начали применять в научных исследованиях, а затем и в промышленности. Разработка откачных средств, не загрязняющих откачиваемый объект, открыла новые перспективы для применения вакуумной техники.

Современная вакуумная техника способна обеспечить получение и измерение давления в 10^{-18} раз меньшего атмосферного, при котором в 1 см^3 остается всего лишь около 30 молекул газа.

Параллельно с освоением техники высокого вакуума развивались теоретические исследования термоэмиссии и электровакуумное приборостроение.

В 1896 году директор Кавендишской лаборатории Джозеф Джон Томсон, повторив опыт Эдисона, тщательно исследовал явление испускания электронов нагретыми телами. Он по праву стал первооткрывателем электрона. В 1901 году его двадцатилетний ученик Оун Вильямс Ричардсон, исследуя термоэлектронную эмиссию платины, вывел уравнение для тока термоэлектронной эмиссии для металлов. Ричардсон основывался на предположении, что эмиттированные нагретым телом частицы отрицательного электричества распределены по скоростям по закону Максвелла в соответствии с температурой испускающего их тела. В 1928 году работы Ричардсона в области термоэлектронной эмиссии были отмечены Нобелевской премией.

Впоследствии уравнение Ричардсона уточнялось с использованием методов квантовой статистики, общих принципов термодинамики. В это существенный вклад внес американский физик С. Дешман.

В 1889 году Джон Амброуз Флеминг провел серию исследований явления термоэлектронной эмиссии. В 1904 году он изобрел вакуумный диод с накаливаемым катодом, обладающий односторонней проводимостью. Свой прибор Флеминг назвал термоионной лампой, или *диодом*. Этот прибор был способен не только выпрямлять переменный ток, но также являлся чувствительным детектором малых токов (рис. 1.1).

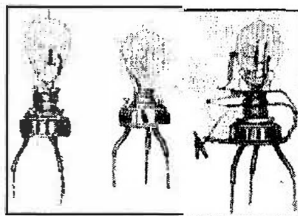


Рис. 1.1. Диоды Флеминга

В 1907 году американский инженер Ли де Форест запатентовал трехэлектродную лампу. Он обнаружил, что малые изменения напряжения на сетке приводят к заметным изменениям тока лампы, в связи с чем и появилась возможность производить электрическое усиление сигнала. Ли де Форест построил на основе своего триода усилитель низкочастотных колебаний — *аудион*. В России первые вакуумные диоды и триоды были созданы В. И. Коваленко, Н. Д. Папалекси, В. И. Волюнкиным.

В 1914 году в лаборатории завода "Русского общества беспроволочных телеграфов и телефонов" (РОБТиТ) стали выпускаться "лампы Папалекси" или *катодные реле*, являющиеся первыми в России усилительными и генераторными электровакуумными приборами. В этом же году Ли де Форест разработал генераторный ламповый триод с водяным охлаждением.

Первые электронные лампы диод и триод сразу нашли практическое применение в области обработки радиосигналов и в значительной степени способствовали становлению радиосвязи. Большой вклад в развитие радиолампостроения внесли: американский ученый Ирвинг Ленгмюр, немецкий физик Вальтер Шоттки, советские ученые Александр Михай-

евич Бонч-Бруевич, Владимир Константинович Лебединский и другие известные ученые инженеры. Лампы были незаменимы в устройствах радиосвязи, радиолокации, а также органически вписались в вычислительные машины первого поколения. С появлением телевидения возрос спрос и производство электронных ламп для телевизоров. Особое место занимают радиолампы для передающих станций. Эта одна из областей, в которой полупроводниковые приборы не вытеснили электровакуумные лампы.

Независимое направление исследований составили электронно-лучевые приборы, берущие свое начало от "катодной" трубки Брауна (1907 г.). Он впервые применил ее в качестве катодного осциллографа с электростатическим отклонением луча.

В 1907 году русский ученый, профессор Санкт-Петербургского университета Борис Львович Розинг предложил систему телевидения на основе усовершенствованной электронно-лучевой трубки Брауна с электронной разверткой изображения. Ассистировал ему молодой ученый Владимир Зворыкин, который в 1918 году эмигрировал в США и на основе изобретенного иконоскопа создал промышленное телевидение. Справедливости ради нужно отметить, что идею передающей трубки с накоплением заряда впервые высказал в 1911 году русский инженер Сергей Иванович Катаев.

Русский ученый, советский академик Дмитрий Аполлинариевич Рожанский в 1911 году создал трубку Брауна с холодными электродами и использовал ее для исследования быстроспроотекающих процессов. В 20-х годах прошлого века были опубликованы работы по электронной оптике, которые позволили поставить разработку электронно-лучевых приборов на промышленную основу.

В 1929 году американским физиком русского происхождения В. К. Зворыкиным был изобретен *кинескоп* — приемная телевизионная трубка, а в 1931 году он изобрел *иконоскоп* — передающую телевизионную трубку.

Развитие телевизионной техники привело к созданию новых передающих трубок — ортиконов, видиконов, суперортиконов, плюмбиконов и т. п. В 1933 году П. В. Шмаков и П. В. Тимофеев изобрели *супериконоскоп*, а в 1934 году Ф. Фарнсфорт (США) изобрел *диссектор* — передающую телевизионную трубку. В 1939 году Л. Роуз и Х. Ямсен (США) создали *ортикон*. Конструкция *видикона* была предложена П. Веймером, С. Формом, Р. Гудричем (США) в 1950 году. Одновременно разрабатывались новые типы приемных трубок — кинескопов.

Большой вклад в этом направлении внесли советские ученые А. А. Чернышев, А. П. Константинов, П. В. Шмаков, П. В. Тимофеева, Г. В. Брауде, Б. В. Круссер и др.

В 1905 году А. Эйнштейн впервые попытался теоретически сформулировать явление *внешнего фотоэффекта* и вывел его закон, согласно которому энергия фотоэлектронов линейно возрастает с частотой падающего света и не зависит от его интенсивности. За это открытие он был удостоен Нобелевской премии по физике.

Советским физиком Леонидом Александровичем Кубецким в 1934 году был создан первый многокаскадный фотоэлектронный умножитель, получивший название "трубка Кубецкого" (рис. 1.2).



Рис. 1.2. Трубка Кубецкого

Большой вклад в развитие отечественных фотоумножителей внесли советские ученые С. А. Вскнинский, В. Н. Авдеев, Г. С. Вильдгрубе, П. В. Тимофеев и др. Фотоэлектронные и электронно-лучевые приборы, пик расцвета которых пришелся на вторую треть прошлого века, сейчас постепенно вытесняются соответствующими полупроводниковыми и плазменными приборами, дисплеями на жидких кристаллах.

Самостоятельным направлением развития вакуумного приборостроения стали приборы СВЧ-диапазона, широко применяемые в радиолокации, системах обороны и в других областях человеческой деятельности.

В 1936—37 годах Н. Ф. Алексеевым и Д. Е. Маликовым был создан многорезонаторный магнетрон. В 1940 году академик Николай Дмитриевич Девятков и инженер Владимир Иванович Коваленко независимо друг от друга создали отражательный клистрон. Идею пролетного клистрона предложил советский академик Дмитрий Аполлинариевич Рожанский (рис. 1.3). Однако первый усилительный пролетный клистрон был создан в 1937 году братьями Р. Варнан и С. Варнан в США.

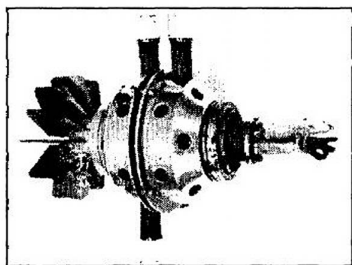


Рис. 1.3. Первый отечественный прямопролетный усилительный клистрон, разработанный в 1940 г.

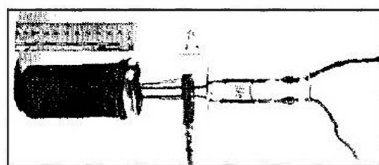


Рис. 1.4. Импульсная генераторная лампа мощностью 250 кВт, генерирующая в 4-метровом диапазоне волн. Использовалась в радиолокационной установке "Сон-2". 1942 г.

Параллельные исследования за рубежом привели к созданию ламп бегущей волны (Р. Компфнер, 1943), а также ламп обратной волны.

Все эти разработки обеспечили создание радиолокационных систем, СВЧ-систем радиорелейной связи, космической связи (рис. 1.4).

Одновременно с развитием вакуумных приборов создавались новые и совершенствовались известные конструкции газоразрядных приборов, прежде всего тиратронов, газоразрядных источников света, газоразрядных панелей.

Приборы и устройства вакуумной электроники (ЭВП) занимают значительную нишу в ряду существующих классов приборов электроники (рис. 1.5). Приборы вакуумной электроники работают в области высоких уровней мощностей 10^6 — 10^{11} Вт и частот 10^8 — 10^{12} Гц.

По своему назначению электронные приборы условно делятся на информационные, предназначенные для генерации и обработки информационных сигналов, и энергетические, применение которых связано с получением электромагнитной энергии.

В настоящем учебнике основное внимание уделено приборам и устройствам электроники, которые предназначены для обработки и хранения информации.

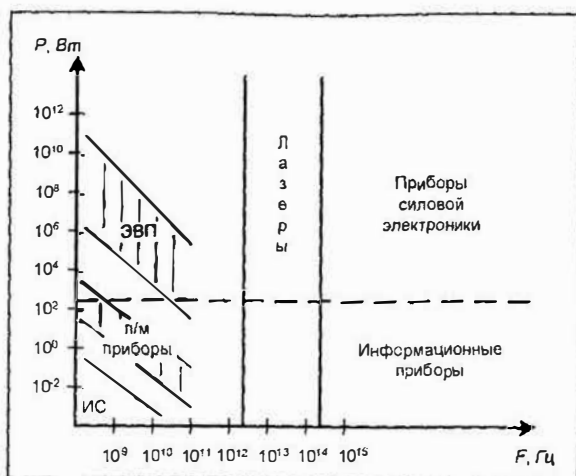


Рис. 1.5. Области применения электронных приборов (на диаграмме мощность-частота)

2. Физика и техника вакуума

2.1. Свойства вакуума

В приборах и устройствах вакуум является континуальной средой. Понятие *континуум*, в переводе с латинского — непрерывный, представляет собой совокупность всех точек пространства с одинаковыми свойствами.

Свойства газов при низких давлениях изучаются в физике вакуума, который является разделом молекулярно-кинетической теории газов. Основой физики вакуума являются следующие постулаты:

- ☐ газ состоит из отдельных движущихся молекул;
- ☐ существует постоянное распределение молекул газа по скоростям;
- ☐ при движении молекул газа нет преимущественных направлений, пространство газовых молекул изотропно;
- ☐ температура газа — величина, пропорциональная средней кинетической энергии его молекул;
- ☐ при взаимодействии с поверхностью твердого тела молекула газа адсорбируется.

Состояние газа, при котором его давление ниже атмосферного, называется *вакуумом*.

При взаимодействии молекул газа с поверхностью твердого тела нормальная составляющая изменения количества движения молекулы равна $mv \cos \Theta$; где Θ — угол между нормалью к поверхности и вектором скорости; m и v — масса и скорость молекулы (рис. 2.1).

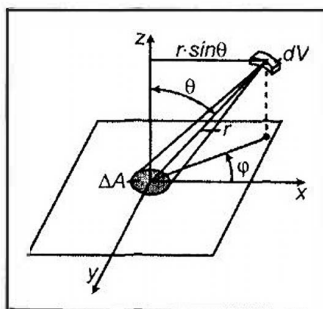


Рис. 2.1. Схема взаимодействия частицы газа с поверхностью

Когда между поверхностью и газовой средой существует энергетическое и адсорбционное равновесие, каждой адсорбированной молекуле соответствует одна десорбированная молекула с противоположным направлением вектора скорости. Таким образом, суммарное изменение количества движения ΔK адсорбированной и десорбированной молекул

$$\Delta K = 2mv \cos \Theta.$$

од давлением газа понимается средний импульс ΔK , передаваемый молекулами газа единице площади стенки сосуда ΔA в единицу времени Δt .

Согласно второму закону Ньютона, давление молекулы на поверхность твердого тела:

$$p = \frac{\Delta K}{\Delta A \cdot \Delta t} = \frac{2mv \cos \Theta}{\Delta A \cdot \Delta t}. \quad (2.1)$$

Число молекул в элементарном объеме dV , движущихся в направлении ΔA , с учетом посылки об отсутствии преимущественных направлений пропорционально телесному углу $d\omega$, под которым из центра dV видна площадка ΔA :

$$dN = n \frac{d\omega}{4\pi} dV, \quad (2.2)$$

а телесный угол определим как

$$d\omega = \cos \Theta \frac{\Delta A}{r^2}, \quad (2.3)$$

где r — расстояние между поверхностью и выделенным объемом (рис. 2.2).

В полярной системе координат элемент объема dV определяется как

$$dV = r \sin \Theta dr d\Theta d\varphi. \quad (2.4)$$

Давление газа на поверхность твердого тела найдем интегрированием по объему полушара, из которой молекулы достигают поверхности за время Δt , с радиусом $R = v\Delta t$.

С учетом выражения (2.1)

$$p = \int_V \frac{2mv \cos \Theta}{\Delta A \Delta t} dN. \quad (2.5)$$

Подставляя все значения в (2.5), можно получить значение давления:

$$p = \frac{nmv}{2\pi\Delta t} \int_0^{2\pi} d\varphi \int_0^{\pi/2} \cos^2 \Theta \sin \Theta d\Theta \int_0^R dR = \frac{nmv^2}{3}. \quad (2.6)$$

В распределение молекул по скоростям введем среднеквадратичную скорость молекул v_{KB}

$$v_{KB}^2 = \frac{1}{n} \sum_{i=1}^n v_i^2.$$

В таком случае давление газа определяется выражением

$$p = nmv_{KB}^2/3. \quad (2.7)$$

Введя значение для плотности газа $\rho = nm$, выражение (2.7) примет вид

$$p = \rho v_{KB}^2/3.$$

Условия динамического равновесия, использованные при выводе уравнения (2.7), не всегда выполняются. Примером является конденсирующая поверхность, с которой из-за очень большого времени адсорбции не происходит десорбции молекул газа.

Для точного расчета давления газа на поверхность твердого тела нужно знать соотношение потоков падающих и вылетающих молекул.

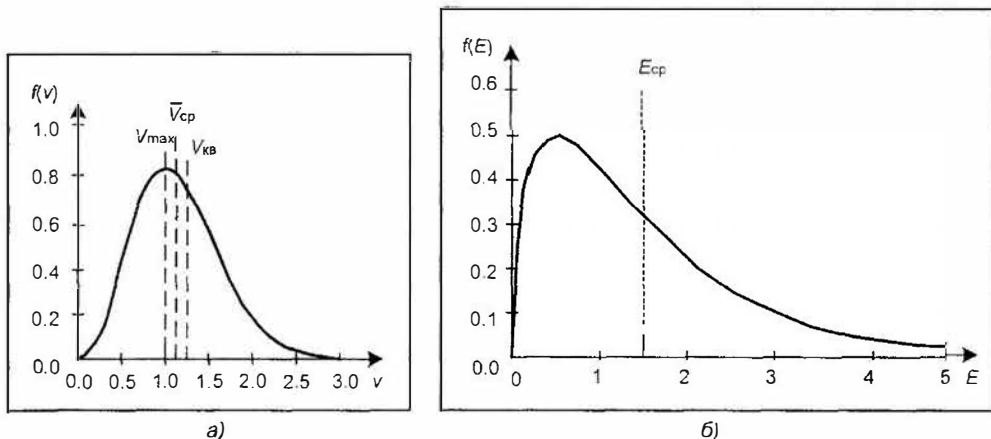


Рис. 2.2. Функции распределения Максвелла по скоростям (а) и по энергиям (б) в безразмерных единицах

Давление газа определяется величиной

$$p = nkT, \quad (2.8)$$

где k — постоянная Больцмана, $1,38 \times 10^{-16}$ эрг/К, которая связывает температуру и энергию.

Уравнение (2.8) известно как уравнение газового состояния и может быть записано в форме уравнения Менделеева — Клайперона

$$pV = \frac{N}{N_A} RT, \quad (2.9)$$

где V — объем газа, $N_A = 6,02 \times 10^{23}$ моль $^{-1}$ — число Авогадро или количество молекул в одном моле вещества, $R = kN_A = 8,31$ Дж/(К·моль $^{-1}$) — универсальная газовая постоянная.

Единицей давления в системе СИ является Па (паскаль), численно равный 1 Н/м 2 . Иногда также используют единицу гектопаскаль 1 гПа = 10^2 Па. Наиболее распространенной внесистемной единицей давления является миллиметр ртутного столба (торр). Давление газа 1 мм рт. ст. равно давлению, которое создает столбик ртути высотой 1 мм при условии, что плотность ртути равна 13595,1 кг/м 3 (при 0°C), а земное ускорение соответствует нормальному 9,80665 м/с 2 на широте 45°. Давление столба жидкости $p = \rho gh$, тогда 1 мм рт. ст. = 133,32239 Н/м 2 . В метеорологии в качестве единицы давления часто используется 1 бар = 10^5 Па. Отношения между различными единицами давления даны в табл. 2.1.

Для описания процессов, в которых давление изменяется в очень широких пределах, удобна логарифмическая единица давления. Например, она может быть задана в виде $pA = -\lg(p)$, где p — давление в физических атмосферах. Значение $pA = 0$ соответствует 1 физ. ат., $pA = 1$ соответствует 0,1 физ. ат. и т. д.

При соударениях друг с другом или со стенками вакуумной камеры молекулы газа изменяют свои скорости как по величине, так и по направлению. Пользуясь гипотезами о существовании стационарного распределения молекул по скоростям и об изотропности

содержимости) пространства газовых молекул, а также учитывая, что среднеквадратичная скорость молекул $v_{\text{кв}} = \sqrt{3kT/m}$, Максвелл получил функцию распределения молекул по скоростям, названную его именем:

$$dn_v = 4\pi n v^2 \left(\frac{m}{2\pi kT} \right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) dv, \quad (2.10)$$

где dn_v — число молекул, скорости которых заключены в пределах от v до $v+dv$.

Таблица 2.1. Соотношение между единицами давления

Единицы давления	1 Па	1 мм рт. ст.	1 дин/см ²	1 физ. ат.	1 кгс/см ²	1 кал/м ³
1 Па	1	$7,5 \cdot 10^{-3}$	10	$9,87 \cdot 10^{-6}$	$1,02 \cdot 10^{-5}$	$2,39 \cdot 10^{-1}$
1 мм рт. ст.	$1,33 \cdot 10^2$	1	$1,33 \cdot 10^3$	$1,32 \cdot 10^{-5}$	$1,36 \cdot 10^{-5}$	$3,18 \cdot 10^2$
1 дин/см ²	$1,00 \cdot 10^{-1}$	$7,50 \cdot 10^{-4}$	1	$9,87 \cdot 10^{-7}$	$1,02 \cdot 10^{-6}$	$2,39 \cdot 10^{-2}$
1 физ. ат.	$1,01 \cdot 10^5$	$7,60 \cdot 10^2$	$1,01 \cdot 10^6$	1	$1,03 \cdot 10^6$	$2,39 \cdot 10^4$
1 кгс/см ²	$9,81 \cdot 10^4$	$7,36 \cdot 10^2$	$9,81 \cdot 10^5$	$9,68 \cdot 10^{-1}$	1	$2,34 \cdot 10^4$
1 кал/м ³	$4,19 \cdot 10^0$	$3,14 \cdot 10^{-2}$	$4,19 \cdot 10^1$	$4,13 \cdot 10^{-5}$	$4,27 \cdot 10^{-5}$	1

Если ввести безразмерную скорость $y = v\sqrt{m/2kT}$ и безразмерную энергию $x = E/kT$, то распределение молекул по скоростям примет вид:

$$dn_y = f_y(y) dy,$$

а распределение по энергиям соответственно

$$dn_x = f_x(x) dx,$$

где

$$f_x(x) = \frac{2}{\sqrt{\pi}} \sqrt{x} \exp(-x)$$

и

$$f_y(y) = \frac{4}{\sqrt{\pi}} y^2 \exp(-y^2)$$

— функции, графики которых приведены на рис. 2.2.

Максимум функции $f_y(y)$ соответствует значению $y=1$, и скорость называется наиболее вероятной скоростью $v_{\text{вер}}$:

$$v_{\text{вер}} = \sqrt{2kT/m}. \quad (2.11)$$

В вакуумных расчетах часто используют:

□ среднеарифметическую скорость (средняя $v_{\text{ар}}$)

$$v_{\text{ар}} = \frac{1}{n} \int_0^{\infty} v dn_v = \sqrt{\frac{8kT}{\pi m}}, \quad (2.12)$$

□ среднеквадратичную скорость

$$v_{\text{кв}} = \sqrt{\frac{1}{n} \int_0^{\infty} v^2 dn_v} = \sqrt{\frac{3kT}{m}}. \quad (2.13)$$

Соотношения между скоростями $v_{\text{вер}}$, $v_{\text{ар}}$ и $v_{\text{кв}}$ равно 1:1,128:1,225. Так, указанные скорости для молекул азота при 0 °С составляют $v_{\text{вер}} = 402$ м/с, $v_{\text{ар}} = 453$ м/с, $v_{\text{кв}} = 492$ м/с.

Максимум дифференциальной кривой соответствует наиболее вероятной энергии $E_{\text{вер}} = 0,5kT$. Расчет среднеарифметического значения энергии молекул дает величину $E_{\text{ар}} \approx 5kT$.

Направленный молекулярный поток, содержащий в начальный момент N_0 молекул газа, за счет столкновений с хаотически движущимися молекулами с частотой K за время dt уменьшается на величину $dN = -KNdt$. Проинтегрировав это выражение, получим:

$$N = N_0 e^{-Kt} = N_0 e^{-l/L},$$

где $L = v/K$ — средняя длина свободного пробега молекул газа, определяемая отношением скорости молекул к числу столкновений в единицу времени; $l = vt$ — длина пробега молекулы за время t .

Средняя длина пробега определяется как расстояние, при прохождении которого частица в среднем сталкивается с одной молекулой среды.

Столкновение молекул произойдет, если расстояние между центрами молекул не более диаметра молекулы d_m (рис. 2.3).

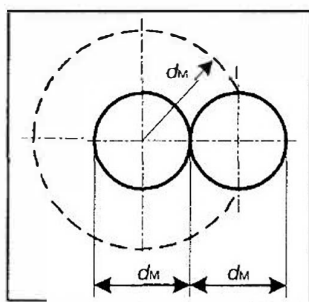


Рис. 2.3. Столкновение двух одинаковых молекул

Будем считать, что одна молекула имеет радиус d_m , а все остальные молекулы — математические точки с нулевым радиусом. При движении со скоростью v в газе с молекулярной концентрацией n за 1 с такая воображаемая молекула опишет объем $V = \pi d_m^2 v$ и испытает $K = n\pi d_m^2 v$ соударений. В этом случае средняя длина свободного пробега будет равна

$$L = \frac{v}{K} = \frac{1}{n\pi d_m^2}. \quad (2.14)$$

С учетом относительных скоростей движения молекул газа для длины свободного пробега можно получить более точное выражение:

$$L = \frac{1}{\sqrt{2}n\pi d_m^2}. \quad (2.15)$$

Из формулы (2.15) видно, что при постоянной молекулярной концентрации длина свободного пробега не должна зависеть от температуры. Эксперименты показывают, что при постоянной молекулярной концентрации с увеличением температуры длина свободного пробега увеличивается. Поэтому в (2.15) целесообразно ввести дополнительный множитель и тогда

$$L = \frac{1}{\sqrt{2}n\pi \cdot d_m^2 (1 + C/T)}, \quad (2.16)$$

где C — постоянная, равная температуре, при которой в случае постоянной молекулярной концентрации газа средняя длина свободного пробега молекул уменьшается вдвое по сравнению со значением, соответствующим бесконечно большой температуре.

Для учета взаимодействия молекул газа между собой (взаимного притяжения) вводят понятие эффективного диаметра молекулы d_i :

$$d_i^2 = d_m^2 (1 + C/T).$$

Эффективный диаметр молекулы уменьшается с увеличением температуры газа. Тогда формулу (2.16) можно представить в виде

$$L = \frac{1}{\sqrt{2}\pi \cdot d_i^2 n}.$$

Используя уравнение газового состояния, значение L можно преобразовать:

$$L = \frac{kT}{\sqrt{2}p\pi \cdot d_m^2 (1 + C/T)} = \frac{kT^2}{\sqrt{2} \pi p d_m^2 (T + C)}. \quad (2.17)$$

Для воздуха при комнатной температуре ($T = 293$ К) и давлении 1 Па из (2.17) следует, что $L_1 = 6,7 \times 10^{-3}$ мПа.

При любом другом давлении

$$L = \frac{L_1}{p} = \frac{6,7 \times 10^{-3}}{p}, \quad (2.17a)$$

где p измеряется в паскалях, а L — в метрах.

Для расчета длины свободного пробега молекул газа при различных температурах и постоянном давлении можно получить следующее выражение:

$$L_T = L_0 \frac{T^2 (T_0 + C)}{T_0^2 (T + C)}. \quad (2.18)$$

В табл. 2.2 приведены средние длины свободного пробега молекул различных газов.

В случае смеси двух газов, молекулы которых имеют массы m_1 и m_2 , среднюю длину свободного пробега L_1 частицы с массой m_1 рассчитывают по формуле:

$$L_1 = \frac{1}{\sqrt{2}n_1 d_{11}^2 + \pi n_2 d_{12}^2 \sqrt{1 + m_1/m_2}}. \quad (2.19)$$

Здесь d_{11} — эффективный диаметр молекул с массой m_1 и концентрацией n_1 ;

$d_{12} = \frac{1}{2}(d_{11} + d_{12})$, где d_{12} — эффективный диаметр молекулы с массой m_2 и концентрацией n_2 .

Таблица 2.2. Средняя длина свободного пробега молекул газов при давлении 1 Па

Газы	$L_1 10^3$, мПа при t, К				Газы	$L_1 10^3$, мПа при t, К			
	600	293	77	4,2		600	293	77	4,2
N ₂	20,8	8,67	1,26	0,0061	H ₂	28,2	12,2	0,197	0,0108
O ₂	16,9	7,02	1,00	0,0047	Хс	10,5	3,93	0,44	0,0017
Ar	16,7	6,79	0,933	0,0042	H ₂ O	13,9	4,38	0,391	0,0013
CO ₂	11,6	4,32	0,492	0,0019	Воздух	16,0	6,72	0,995	0,0048
Ne	30,7	13,9	2,50	0,0165	He	43,6	19,1	3,13	0,0174
Kr	14,1	5,52	0,691	0,0029					

Первое слагаемое в знаменателе (2.19) зависит от столкновения одинаковых частиц с массой m_1 , а второе — от столкновения частиц с массами m_1 и m_2 . Если $n_1 \ll n_2$, то получим более простое выражение:

$$L_1 = \frac{l}{\pi n_2 d_{12}^2 \sqrt{1 + m_1/m_2}} \quad (2.20)$$

Процессы в вакууме сильно зависят от соотношения между числом взаимных столкновений молекул и числом столкновений молекул со стенками вакуумной системы.

Частота столкновений между молекулами K_M обратно пропорциональна средней длине свободного пробега:

$$K_M = v_{cp}/L$$

Из общего числа соударений молекул со стенками камеры $n v_{cp} F / 4$, приходящихся на nV молекул, рассчитывается среднее число соударений со стенкой K_C , приходящихся в единицу времени на одну молекулу:

$$K_C = v_{cp} A / (4V) = v_{cp} / d_{эф},$$

где A — площадь поверхности стенок, соприкасающихся с разреженным газом; V — объем камеры; $d_{эф} = \frac{4V}{A}$ — эффективный размер вакуумной камеры.

Для молекул газа внутри сферического сосуда диаметром D эффективный размер камеры $d_{эф} = 2D/3$, для трубы бесконечной длины с диаметром D получим $d_{эф} = D$, а для двух бесконечных параллельных поверхностей, расположенных на расстоянии D друг от друга, — $d_{эф} = 2D$.

Отношение K_C / K_M называется критерием Кнудсена:

$$K_n = K_C / K_M = L / d_{эф} \quad (2.21)$$

В зависимости от значения критерия Кнудсена различают вакуум низкий, средний и высокий.

Низкий вакуум — это состояние газа, при котором взаимные столкновения между молекулами преобладают над столкновениями молекул газа со стенками вакуумной камеры.

Такое состояние газа соответствует критерию Кнудсена $K_n \ll 1$. При этом длина свободного пробега молекул газа значительно меньше размеров вакуумной камеры. Из условия изменения режима течения газа принимают $K_n \leq 5 \times 10^{-3}$.

Средний вакуум — это состояние газа, когда частоты соударений молекул друг с другом и со стенками вакуумной камеры одинаковы, при этом $L \approx d_{\text{эф}}$, а $K_n \approx 1$.

Высокий вакуум — это состояние газа, при котором столкновения молекул газа со стенками вакуумной камеры преобладают над взаимными столкновениями молекул газа.

При этом $K_n > 1$. Из условия изменения режима течения газа принимают $K_n \geq 1,5$. Тогда условие существования среднего вакуума можно записать в виде $5 \times 10^{-3} < K < 1,5$ (табл. 2.3).

Таблица 2.3

Степень вакуума	Низкий	Средний	Высокий
Критерий Кнудсена	$K_n \ll 1$, $L \ll d_{\text{эф}}$	$K_n \approx 1$, $L \approx d_{\text{эф}}$	$K_n > 1$, $L > d_{\text{эф}}$

Из этих определений следует, что степени вакуума — понятия относительные; одному и тому же давлению могут соответствовать различные степени вакуума (в зависимости от соотношения $L/d_{\text{эф}}$). Отсюда следует, что каждый разработчик или технолог должен сочетать особенности в поведении газа, зависящие от соотношения между L и $d_{\text{эф}}$, с необходимыми требованиями к абсолютному значению давления газа.

При практической работе в производстве электровакуумных приборов, расчете и конструировании вакуумных систем степени вакуума зачастую характеризуют приближенно абсолютными значениями давлений (табл. 2.4).

Однако эти области давлений соответствуют лишь степени трудности, а также различию способов их получения, но никак не связаны с различиями в свойствах и поведении газа.

Таблица 2.4

Степень вакуума		Низкий	Средний	Высокий	Сверхвысокий
Область давлений	Па	>100	$100-10^{-1}$	$10^{-1}-10^{-3}$	$<10^{-3}$
	мм рт. ст.	>1	$1-10^{-3}$	$10^{-3}-10^{-7}$	$<10^{-7}$

2.2. Методы создания вакуума

Получение, измерение и поддержание вакуума являются процессами вакуумной техники, определяющими успехи при создании приборов и устройств вакуумной электроники.

В основу процессов получения вакуума положены два принципа:

- ☐ удаление газов из откачиваемого объема;
- ☐ связывание газов за счет либо их конденсации при низких температурах, либо за счет связывания на стенках объема специальными материалами.

На первом принципе основана работа традиционных вакуумных механических насосов, относящихся к насосам объемного типа. Механические насосы с масляным уплотнением позволяют захватить порцию газа из рабочего объема и перенести ее в сторону выпуска и выбросить в атмосферу через клапан. Это насосы вращательного типа и они классифицируются на: пластинчато-роторные, пластинчато-статорные, пластинчато-статорные с пластиной в виде поршня и золотниковые или плунжерные. Во всех этих типах насосов зазоры между трущимися деталями уплотняет масло, которое предотвращает поступление в камеру атмосферного воздуха. Вместе с тем масло содержит растворенные газы, которые выделяются в откачиваемом объеме и определяют остаточное давление.

Механические вакуумные насосы с масляным уплотнением используются для получения вакуума в области давлений от 760 до 10^{-3} мм рт. ст.

В этом же диапазоне давлений работают двухроторные вакуумные насосы, которые более экономичны. Конструкция двухроторных насосов состоит из двух фигурных роторов, профиль которых напоминает правильные восьмерки. Роторы синхронно вращаются в общем корпусе навстречу друг другу, выбрасывая порцию газа из откачиваемого объема.

К механическим насосам относятся и турбомолекулярные насосы, принцип действия которых основан на сообщении молекулам газа направленной дополнительной скорости быстро движущейся твердой поверхностью (рис. 2.4). Газ увлекается лопастями насоса в направлении вращения ротора и им сообщается механический импульс.

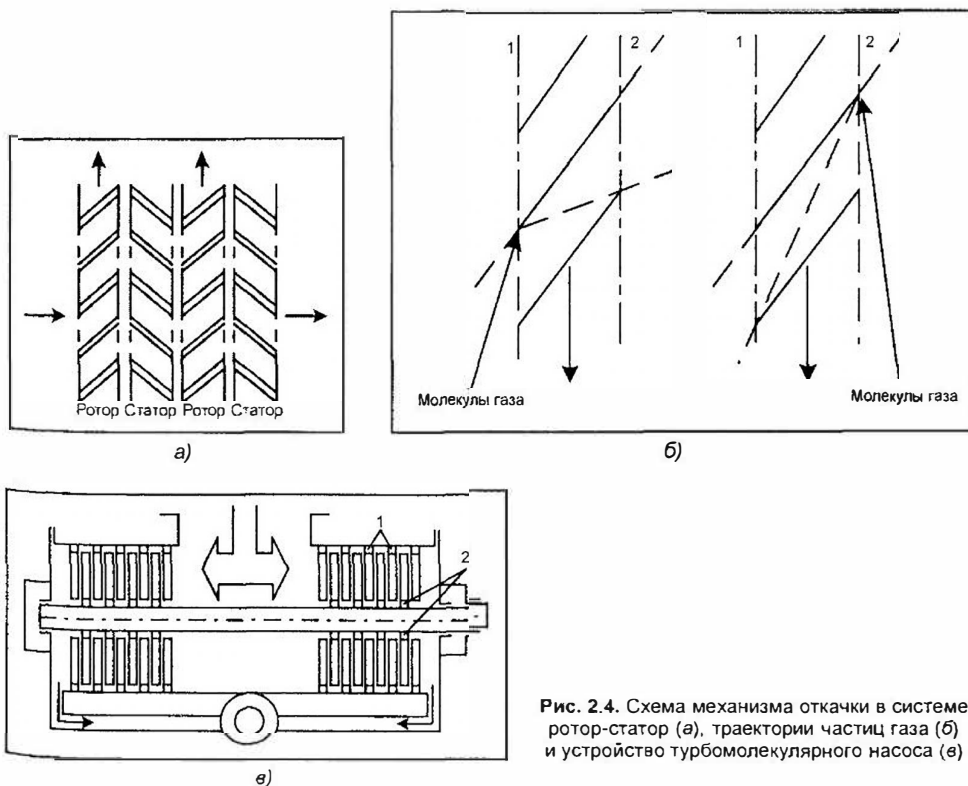


Рис. 2.4. Схема механизма откачки в системе ротор-статор (а), траектории частиц газа (б) и устройство турбомолекулярного насоса (в)

Конструкция турбомолекулярного насоса состоит из статорных 1 и роторных 2 дисков (рис. 2.4, а). В дисках имеются косые прорезы, зеркально расположенные на статорных и роторных дисках. Молекулы газа отражаются от середины ротора к краям, получая импульс от лопастей. Толщины дисков составляют несколько миллиметров, а зазоры не более миллиметра. Диски имеют большое число параллельно работающих прорезей, которые позволяют достичь высокой производительности. Откачка газа производится в поперечном направлении. Общая схема турбомолекулярного насоса представлена на рис. 2.4, б. Основное их предназначение — получение вакуума в диапазоне 10^{-3} — 10^{-9} мм рт. ст. Турбомолекулярные насосы лучше откачивают тяжелые газы, чем легкие.

К газоперемещающему типу насосов относятся струйные насосы. Они делятся на жидкоструйные насосы, газоструйные насосы и пароструйные насосы. Первые действующие вакуумные насосы были водоструйными, действие которых основывается на увлечении газа струей воды. При высокой скорости струи создается разрежение, в которое увлекаются молекулы газа.

Наибольшее распространение получили пароструйные насосы. В их конструкции предусмотрен нагрев рабочей жидкости (вакуумное масло или ртуть) до парообразного состояния. Пар поступает к соплу, из которого он с большой скоростью вытекает в рабочую камеру в виде расходящейся струи. Откачиваемый газ также поступает в камеру, захватывается струей и увлекается к охлаждаемым стенкам рабочей камеры. Рабочий пар конденсируется, и конденсат возвращается к нагревателю. Откачиваемый газ выбрасывается к насосу предварительного разрежения. В зависимости от механизма увлечения газа струей различают бустерные и диффузионные высоковакуумные насосы.

В бустерных насосах увлечение газа струей осуществляется за счет вязкостного трения между паром и газом, а также за счет диффузии газа в струю.

В диффузионных насосах увлечение газа струей осуществляется целиком за счет процесса диффузии молекул газа в струю. Современные вакуумные пароструйные насосы представляют собой многоступенчатую конструкцию с общим испарителем и общим паропроводом для питания сопел отдельных ступеней. Конструкция трехступенчатого диффузионного пароструйного насоса представлена на рис. 2.5.

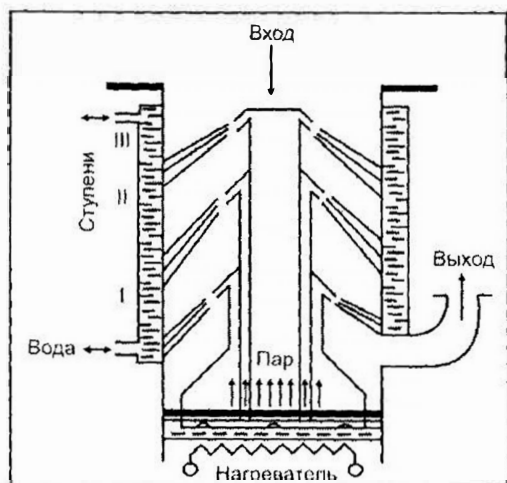


Рис. 2.5. Схема трехступенчатого диффузионного насоса

Высоковакуумные пароструйные насосы позволяют производить откачку в пределах 10^{-4} — 10^{-6} мм рт. ст.

Отметим, что использование паромасляных насосов не позволяет получить вакуум без радикалов. Применение парортутных диффузионных насосов позволяют получить "чистый" вакуум.

Для получения вакуума в пределах 10^{-6} — 10^{-10} мм рт. ст. применяют низкотемпературные ловушки.

К другому типу откачных средств относятся насосы, в основе работы которых лежат физико-химические методы получения вакуума.

Хемосорбционная откачка осуществляется путем поглощения газов поверхностью металлов: Ti, Zr, Ta, Ba, Mo, W, Hf.

В сорбционном насосе, действие которого основано на поглощении откачиваемого газа поверхностью поглотителя или сорбента газов, в качестве поглотителей используются пористые вещества с сильно развитой поверхностью (активированный уголь, цеолит, силикагель).

Такой безмаслянный способ откачки основан на способности сорбента поглощать значительные количества газа при его охлаждении до сверхнизких (азотных) температур. Насыщенные газом пористые сорбенты после прогрева практически полностью восстанавливают свои сорбционные свойства. Конструктивно цеолитовый насос состоит из цилиндрической капсулы, заполненной сорбентом, которая связана с откачиваемым объемом. При погружении капсулы в сосуд Дьюара происходит откачка. Регенерацию сорбента производят с помощью электронагревателя. Цеолитовые насосы позволяют получить вакуум в пределах 10^{-2} — 10^{-4} мм рт. ст.

В другом типе сорбционных насосов поглощающую поверхность создают напылением химически активного металла, активно реагирующего с большей частью газов. По способу получения поглощающей пленки различают испарительные геттерные ионные насосы (ГИН) и магниторазрядные.

Работа ионно-геттерных насосов (ГИН) основана на совмещении в одной конструкции двух параллельно протекающих процессов. Это процесс поглощения газов периодически или непрерывно наносимой пленкой активного вещества и процесс откачки инертных газов и углеводородов за счет ионизации и улавливания положительных ионов отрицательно заряженными ловушками насоса.

В качестве поглощающей пленки используется титан, напыляемый на внутреннюю охлаждаемую стенку насоса. Ионизация откачиваемого газа осуществляется электронами, выпускаемыми накалившимся катодом и направляющимися к анодной системе, и коллектором ионов, в качестве которого служит корпус насоса.

Отличительными качествами таких насосов являются их длительная работа без смены испарителей, наличие внутреннего нагревателя, позволяющего сокращать время на запуск насоса. Предельный вакуум в этом типе насосов может достигать значений порядка 10^{-10} мм рт. ст.

Более высокий вакуум достигается с помощью насосов орбитронного типа. Эти насосы являются продолжением усовершенствования конструкции ГИН. С целью увеличения пути движения электронов они направляются по эллиптическим спиралеобразным орбитам. Это позволяет повысить эффективность ионизации остаточных газов и увеличить быстроту откачки. В таких насосах можно получить вакуум до 10^{-13} мм рт. ст.

К этому же типу относятся магниторазрядные насосы, работа которых основана на процессах поглощения газов титаном, который распыляется высокочастотным разрядом в магнитном поле.

Конструктивной основой магниторазрядных насосов являются ячейки Пеннинга, состоящие из двух параллельных пластин-катодов и цилиндрического анода. Ось анода располагается перпендикулярно катодам и параллельно вектору индукции магнитного поля (рис. 2.6). Магнитное поле напряженностью до 1500 Э создается оксидно-бариевыми постоянными магнитами. При приложении напряжения между электродами возникает газовый разряд. Электроны движутся по спирали вдоль магнитных силовых линий, ионизируя газ на своем пути. Образующиеся положительные ионы бомбардируют катод и распыляют титан, находящийся на катодных пластинах. Активные газы в процессе химisorбции оседают на пленке титана. Инертные газы внедряются в материал катода.

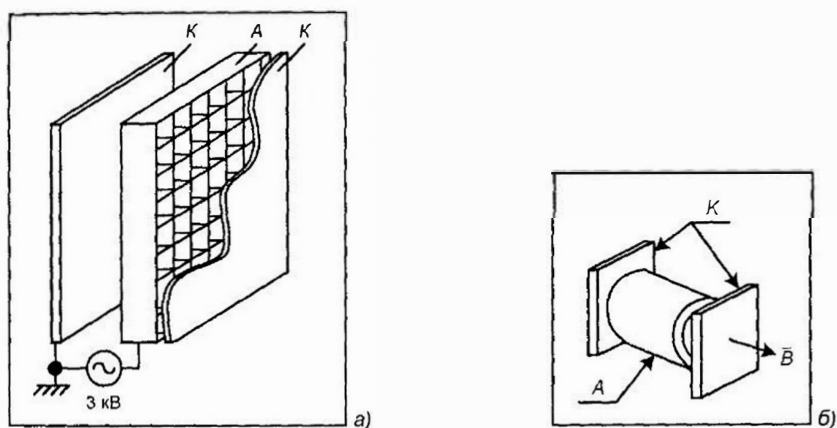


Рис. 2.6. Конструкция магниторазрядного насоса (а) и ячейки Пеннинга (б): А — анод; К — катоды; В — вектор магнитной индукции

Благодаря отсутствию в насосах этого типа нагретых и движущихся частей, а также рабочей жидкости они обладают высокой надежностью, большим сроком службы, и не выходят из строя при аварийном попадании атмосферы в вакуумную систему. Магниторазрядные насосы предназначены для работы в области высокого и сверхвысокого вакуума и позволяют получить остаточное давление порядка 10^{-10} мм рт. ст.

Криогенные насосы работают на использовании процессов вымораживания остаточных газов при температуре стенок, близких к гелиевой температуре (4,2 К). С помощью такого типа насосов давление остаточных газов в системе достигает значений 10^{-9} — 10^{-10} мм рт. ст.

2.3. Методы измерения вакуума

Способы измерения вакуума являются самостоятельным разделом вакуумной электроники, поскольку необходимо измерять давления в диапазоне ниже атмосферного от 760 до 10^{-13} мм рт. ст. (10^5 — 10^{-11} Па). Универсального метода измерений, охватывающего этот диапазон давлений, не существует. Поэтому исследуются различные физические явления и эффекты, на основе которых и разрабатываются методы измерений вакуума.

Гидростатические U -образные вакуумметры относятся к *абсолютным манометрам*, позволяющим непосредственно измерять давление. Конструкция жидкостных манометров представляет собой U -образную трубку с сообщающимися коленами, которые заполнены ртутью или вакуумным маслом с низким значением упругости пара. Процесс измерения давления сводится к измерению разности уровней жидкости в коленях, одно из которых соединено с измеряемым объемом. Диапазон измеряемых давлений лежит в пределах $760\text{—}10^{-2}$ мм рт. ст.

В деформационном вакуумметре давление измеряется по деформации упругого датчика типа мембраны или сильфона. Опорным давлением служит атмосферное давление.

К абсолютным вакуумметрам относятся компрессионный тип, например манометр Мак-Леода, использующий закон Бойля-Мариотта $pV = \text{const}$.

Все остальные типы вакуумметров относятся к *относительным манометрам*, в которых измеряются физические параметры в зависимости от давления, и в дальнейшем градуируются по абсолютным образцовым вакуумметрам.

Принцип действия *тепловых преобразователей* основан на зависимости теплопроводности разреженного газа от давления. Передача тепла происходит от нагреваемой электрическим током тонкой металлической нити к баллону, находящемуся при комнатной температуре. Уравнение теплового баланса такого прибора можно представить в следующем виде:

$$I_H^2 R = E_K + E_T + E_H + E_M, \quad (2.22)$$

где I_H — ток, проходящий через нить; R — сопротивление нити; E_K, E_T, E_H, E_M — соответствующие значения потерь тепла за счет конвекции, теплопроводности газа, излучения нити и теплопроводности материала нити.

Конвективным теплообменом в области среднего и высокого вакуума можно пренебречь, т. е. $E_K \approx 0$.

Потери тепла за счет теплопроводности через газовый промежуток

$$E_T \approx \lambda_0 (T_H - T_B) A, \quad (2.23)$$

где λ_0 — коэффициент теплопроводности газа, A — сечение нити.

В области высокого вакуума коэффициент теплопроводности пропорционален давлению газа

$$\lambda_0 = K_T p, \quad (2.24)$$

где K_T — коэффициент пропорциональности.

Потери тепла излучением и конвективным теплообменом:

$$E_H = K_H (T_H^4 + T_B^4) A,$$

здесь A — площадь поверхности нити; K_H — коэффициент излучения материала нити; T_H и T_B — соответственно температуры нити и баллона.

Тепловые потери нити за счет передачи теплоты по материалам нити и электродов, соединяющих нить с корпусом преобразователя,

$$E_M = \lambda (T_H - E_B) A_1,$$

где λ — коэффициент теплопроводности материала нити, A_1 — сечение нити.

Тогда измерительное уравнение теплового преобразователя с учетом уравнений (2.22), (2.23) и (2.24) можно записать так:

$$p = \frac{I_H^2 R - (E_H + E_M)}{K_T (T_H - T_B)}. \quad (2.25)$$

Более точное измерение давления может быть достигнуто при условии, чтобы E_T составляло значительную долю от $E_H + E_M$, т. е. чтобы сумма $E_H + E_M$ была существенно меньше мощности $I_H^2 R$, выделяющейся в нити манометра. Это условие определяет нижний предел измерений вакуума тепловым преобразователем

$$I_H^2 R - (E_H + E_M) \geq 0,01 \cdot I_H^2 R.$$

Из уравнения (2.25) видно, что давление является функцией двух переменных: тока накала I_H и температуры нити T_H .

Существует два метода работы тепловых манометров: в режимах постоянной температуры нити и постоянного тока накала. Градуировочные кривые теплового манометра показаны на рис. 2.7 для обоих методов работы. Они представляют собой в средней части параболу и гиперболу. Концы градуировочных кривых у верхнего и нижнего пределов измерения уже не описываются уравнением (2.25) и переходят в линии, параллельные оси давления.



Рис. 2.7. Градуировочные кривые теплового преобразователя:
а — при постоянном токе накала; б — при постоянной температуре нити

Для увеличения верхнего предела измерений тепловым преобразователем следует уменьшать его габариты, вследствие чего увеличивается отношение L/d и сдвигается граница низкого вакуума в сторону более высоких давлений. Зависимость коэффициента конвективного теплообмена от давления используется для измерения давлений в области низкого вакуума. Недостатком этого способа является зависимость показания прибора от его положения в пространстве.

Нижний предел измерения тепловых преобразователей можно оптимизировать путем уменьшения доли $E_H + E_M$ в сумме тепловых потерь нити. Этого можно достичь понижением температуры нити и уменьшением диаметра вводов, соединяющих нить с баллоном.

Показания тепловых преобразователей определяются соотношением (2.24) и зависят от рода газа. Преобразователь будет давать одинаковые показания при выполнении следующих условий:

$$p_1 K_{T_1} = p_2 K_{T_2} = \dots = p_i K_{T_i} = p_n K_{T_n}.$$

Выпускаемые промышленностью приборы проградуированы по сухому воздуху. Если необходимо измерить давление других газов, то нужно учитывать относительную чувствительность прибора к данному газу

$$p_i = p_B \frac{K_{TB}}{K_{Ti}} = p_B q_i, \quad (2.26)$$

где p_B и K_{TB} — давление и коэффициент теплопроводности воздуха; $q_i = K_{TB}/K_{Ti}$ — коэффициент относительной чувствительности теплового преобразователя к данному газу.

Если преобразователь измеряет давление смеси газов, то его показания будут выражены в воздушном эквиваленте p_B :

$$p_1 K_{T1} + p_2 K_{T2} + \dots + p_n K_{Tn} = p_B K_{TB}. \quad (2.27)$$

Для воздуха сумма в (2.27) равна $p_B K_{TB}$, а коэффициент относительной чувствительности теплового преобразования к данному газу примет вид $q_i = K_{TB}/K_{Ti}$.

Так как из определения относительной чувствительности следует, что $p_B = p_{см}/q_{см}$. Тогда можно записать:

$$p_{см}/q_{см} = p_1/q_1 + p_2/q_2 + \dots + p_n/q_n.$$

Разделив обе части уравнения на $p_{см}$, получим

$$\frac{1}{q_{см}} = \frac{V_1}{q_1} + \frac{V_2}{q_2} + \dots + \frac{V_n}{q_n},$$

где $V_1 \dots V_n$ — объемные концентрации соответствующих газов, причем $\sum_{i=1}^n V_i = 1$.

Таким образом, коэффициент относительной чувствительности для смеси газов определяется по формуле

$$\frac{1}{q_{см}} = \sum \frac{V_i}{q_i}. \quad (2.28)$$

Тепловые преобразователи в зависимости от способа измерения температуры делятся на *термопарные* и *преобразователи сопротивления*.

В термопарном преобразователе (рис. 2.8, а) температура нити 1 измеряется термопарой 2. Электроды расположены в стеклянном или металлическом баллоне 3, имеющем патрубок для подключения к вакуумной системе. Термо-ЭДС термопары измеряется милливольтметром, ток накала нити регулируется реостатом и измеряется миллиамперметром.

В преобразователе сопротивления для измерения температуры используется зависимость сопротивления нити от температуры. Он включается в мостовую схему (рис. 2.8, б). Ток накала нити измеряется миллиамперметром, включенным в то же самое плечо моста, что и преобразователь, а температура нити — по току гальванометра в измерительной диагонали моста. Ток накала регулируется реостатом.

Оба преобразователя могут работать как в режиме постоянного тока накала, так и в режиме постоянной температуры нити.

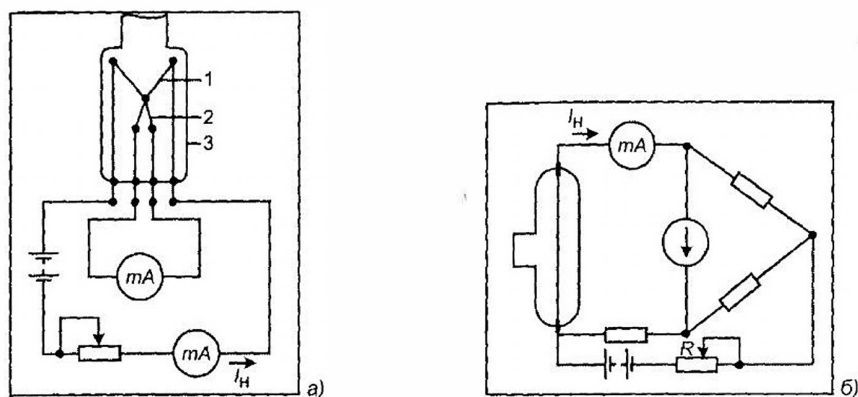


Рис. 2.8. Схемы тепловых преобразователей: а — термопарного; б — преобразователя сопротивления. 1 — нить накала; 2 — термопара; 3 — баллон

Преимуществом тепловых преобразователей является то, что они измеряют общее давление всех газов и паров, присутствующих в вакуумной системе, а также обеспечивают непрерывность измерения давления. Инерционность показаний, связанная с тепловой инерцией нити, изменяется от нескольких секунд при низких давлениях до нескольких миллисекунд при высоких давлениях.

Тепловые преобразователи как приборы для относительных измерений давления обычно градуируются по компрессионному манометру. Диапазон измеряемых рабочих давлений лежит в пределах 10^2 — 10^{-5} мм рт. ст. (5×10^3 — 10^{-1} Па).

В основе принципа действия *электронных ионизационных вакуумметров* лежит прямая зависимость между давлением остаточных газов и ионным током, образовавшимся в результате ионизации молекул газа термоэлектронами.

Существует две схемы электронного преобразователя: с внутренним и внешним коллектором, реализованных на базе триодной конструкции.

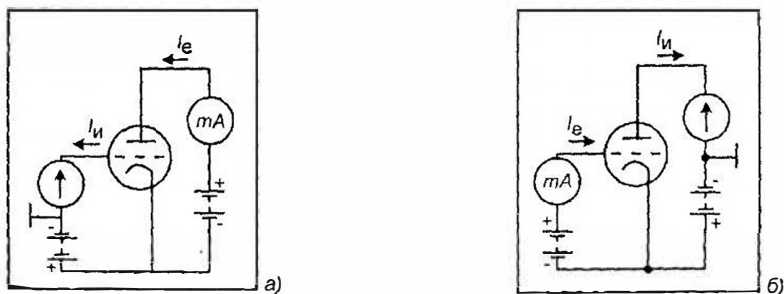


Рис. 2.9. Схемы электронных преобразователей: а — с внутренним коллектором; б — с внешним коллектором

Схема с внутренним коллектором (рис. 2.9, а) аналогична обычному триоду. Коллектором ионов является сетка, на которую относительно катода подается отрицательное напряжение величиной в несколько десятков вольт. На анод подается соответственно положительное напряжение 100—200 В. Электроны на пути от катода к аноду (ток I_e) соударяют

ся с молекулами остаточных газов и ионизируют их. Образовавшиеся положительные ионы попадают на сетку, создавая ионный ток I_H , измеряемый гальванометром.

В схеме с внешним коллектором (рис. 2.9, б) потенциалы сетки и анода меняются местами, и коллектором становится анод. Электроны, летящие от катода к сетке, совершают вокруг ее витков ряд колебаний, что увеличивает длину траектории электронов и повышает вероятность ионизации молекул остаточных газов. Это делает схему с внешним коллектором более чувствительной, несмотря на то, что часть положительных ионов, образовавшихся между сеткой и катодом, не участвует в измерении давления.

Для электронного преобразователя справедливо уравнение

$$dN = n p \varepsilon \cdot dr, \quad (2.29)$$

где dN — число положительных ионов; n — число электронов; dr — элементарная длина траектории электронов; ε — эффективность ионизации, равная количеству положительных ионов, образуемых одним электроном на единице пути при единичном давлении.

Если ввести в уравнение (2.29) электронный ток $I_e = n/t$, то получим

$$dN/t = I_e p \varepsilon \cdot dr.$$

Интегрируя это уравнение по всей длине траектории электрона с энергией, большей потенциала ионизации, получим выражение для ионного тока

$$I_H = I_e p \int_{r_1}^{r_2} \varepsilon \cdot dr,$$

которое перепишем в виде

$$I_H = K_H I_e p, \quad (2.30)$$

где $K_H = \int_{r_1}^{r_2} \varepsilon \cdot dr$ — чувствительность электронного вакуумметра, r_1 и r_2 — пределы интегрирования, определяемые геометрией манометра.

Выражение (2.30) называется уравнением электронного преобразователя. Для того чтобы измеряемый ионный ток был пропорционален давлению, необходимо во время измерения поддерживать постоянное значение электронного тока. Тогда $a = I_e K_H$ — постоянная электронного манометра, которая равна тангенсу угла наклона градуировочной кривой $I_H(p)$ к оси давления.

Конструкция электронного преобразователя с внешним коллектором представлена на рис. 2.10, а. Коллектор ионов 1 имеет форму цилиндра с электрическим вводом в верхней части баллона, сетка 2 — форму двойной спирали с двумя выводами для обезгаживания путем пропуска электрического тока. Катод 3 изготовлен из вольфрама. Постоянная a при токе эмиссии 5 мА для типового преобразователя составляет примерно 10^{-3} А/Па. Пределы давлений, которые могут быть измерены таким манометрическим преобразователем, составляют 10^{-1} — 10^{-10} мм рт. ст. (1 — 10^{-5} Па).

Верхний предел измерения типового электронного преобразователя составляет примерно 1 Па и соответствует нарушению линейности градуировочной характеристики, когда средняя длина свободного пробега электрона в объеме прибора становится меньше величины пробега электрона между электродами. Для расширения верхнего предела можно

уменьшить расстояние между электродами. Чтобы катод не сгорел при таких высоких давлениях, его изготавливают из оксидов редкоземельных металлов.

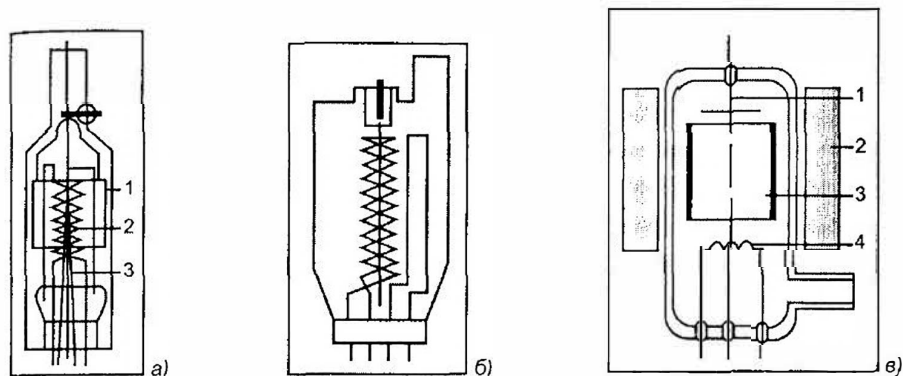


Рис. 2.10. Конструктивные схемы электронных преобразователей:
а — с внешним коллектором; б — с осевым коллектором; в — с магнитным полем

Нижний предел измерения определяется фоновыми токами в цепи коллектора, возникающими из-за эмиссии фотоэлектронов в результате мягкого рентгеновского излучения анодной сетки и ультрафиолетового излучения накаливаемого катода. Рентгеновское излучение анодной сетки является результатом ее бомбардировки электронами. Автоэлектронная эмиссия коллектора появляется под действием разности потенциалов 200—300 В между коллектором и анодной сеткой и вносит дополнительную составляющую в фоновый ток.

Фоновые электронные токи имеют одинаковое направление с ионным током и поэтому оказывают одинаковое воздействие на измерительные приборы. Максимальным фоновым током является ток рентгеновского излучения, пропорциональный эмиссионному току:

$$I_{\phi} = K_S I_e,$$

где K_S — коэффициент пропорциональности.

С учетом фоновых токов рентгеновского излучения уравнение электронного преобразователя можно записать в следующем виде:

$$I = I_H + I_{\phi} = I_e (K_N + K_S p),$$

и нижний предел измерения определить соотношением ионного и фонового токов:

$$\frac{I_H}{I_{\phi}} = \frac{K_N}{K_S p}. \quad (2.31)$$

Таким образом, для расширения нижнего предела измерения с помощью электронного вакуумметра нужно при постоянном давлении p увеличить коэффициент K_N или уменьшить K_S .

Для уменьшения фоновых токов и, следовательно, постоянной K_S был предложен преобразователь с осевым коллектором (рис. 2.10, б), в котором катод и коллектор поменялись местами. Это значительно уменьшило телесный угол, в котором рентгеновское излучение сетки попадает на коллектор. Это привело к уменьшению K_S приблизительно в 1000 раз

по сравнению с конструкцией, приведенной на рис. 2.10, а, и расширило нижний предел измерения давления до 10^{-10} мм рт. ст. (10^{-8} Па).

Чувствительность K_H можно увеличить, если поместить преобразователь в магнитное поле (рис. 2.10, в). Электроны от катода к аноду в этом случае движутся по спирали. В электронном преобразователе с магнитным полем, создаваемым катушкой 2 и направленным параллельно оси анода 3, катод 4 — термоэлектронный, а коллектор 1 расположен в верхней части баллона. За счет увеличения чувствительности такой преобразователь имеет нижний предел на 2—3 порядка ниже, чем конструкция преобразователя, показанная на рис. 2.10, а.

Электронный преобразователь имеет неодинаковую чувствительность к различным газам, т. к. эффективность ионизации зависит от рода газа.

Если преобразователь градуирован по воздуху, а применяется для измерения давления других газов, то необходимо учитывать относительную чувствительность R . Из условия равенства ионных токов запишем

$$K_1 p_1 = K_2 p_2 = \dots = K_i p_i = K_B p_B,$$

откуда

$$p_i = p_B / R_i,$$

где $R_i = K_i / K_B$ — относительная чувствительность к данному газу. При измерении давления смеси газов из условия равенства ионных токов имеем

$$K_{см} p_{см} = K_1 p_1 + K_2 p_2 + \dots + K_n p_n.$$

Поделив полученное уравнение на K_B , получим

$$\frac{K_{см} p_{см}}{K_B} = K_1 p_1 + K_2 p_2 + \dots + K_n p_n,$$

откуда

$$R_{см} = \sum_{i=1}^n K_i p_i,$$

$$V_i = p_i / p_{см}.$$

Заметные ошибки измерения возникают в случае, если скорость откачки электронных преобразователей в конструкциях преобразователей, присоединяемых через трубку с малой проводимостью, составляет 10^{-3} — 10^{-1} л/с.

Дополнительные источники погрешности измерения связаны с химическим взаимодействием газов с накаливаемым катодом и ионной десорбцией под воздействием электронной бомбардировки газов, химически поглощенных анодом.

Принцип действия магнитных преобразователей для измерения вакуума основан на зависимости тока самостоятельного газового разряда в скрещенных магнитном и электрическом полях от давления. Электродные системы, обеспечивающие поддержание самостоятельного газового разряда при высоком и сверхвысоком вакууме, бывают нескольких видов.

Ячейка Пеннинга (рис. 2.11, а) состоит из двух дисковых катодов 1 и цилиндрического анода 2, в магнетронном преобразователе (рис. 2.11, б) в отличие от ячейки Пеннинга катоды соединены между собой центральным стержнем; в инверсно-магнетронном пре-

образователе (рис. 2.11, в) центральный стержень выполняет роль анода, а наружный цилиндр становится катодом.

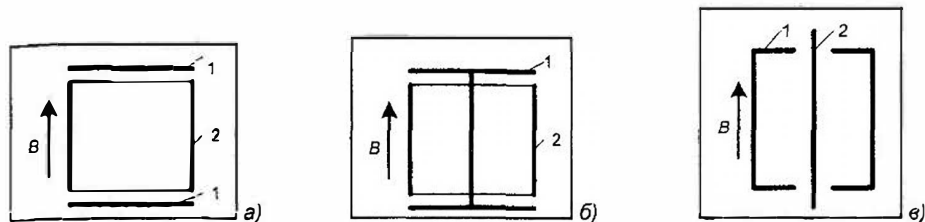


Рис. 2.11. Электродные системы магнитных преобразователей: а — ячейка Пеннинга; б — магнетронная; в — инверсно-магнетронная

Все электроды находятся в постоянном магнитном поле. На анод подается положительное относительно катода напряжение 2—6 кВ, при этом катод заземлен и соединяется с входом усилителя постоянного тока.

В ячейке Пеннинга электроны движутся по спиральным траекториям между катодными пластинами.

В конструкциях, показанных на рис. 2.11, а, б, эмитированные электроны движутся в скрещенных электрическом и магнитном полях. Траектория их движения описывается уравнением циклоиды, образованной окружностью диаметром $D = 2m|E|/(qB^2)$, катящейся по окружности радиуса r с угловой частотой вращения $\omega = qB/m$ и тангенциальной скоростью $v_T = E/B$, где E — напряженность электрического поля; B — магнитная индукция; m и q — масса и заряд электрона.

Магнитная индукция B выбирается больше критического значения, соответствующего равенству диаметра электрода и диаметра окружности, по которой движется электрон, и составляющего в современных приборах примерно 0,1 Тл.

При соударении с молекулой остаточного газа электрон теряет часть энергии на ее ионизацию и перемещается в радиальном направлении к аноду. В связи с тем, что радиальная скорость электронов значительно меньше, чем тангенциальная, при низких давлениях в разрядном промежутке образуется отрицательный объемный заряд.

Положительные ионы, образовавшиеся в результате столкновения с электронами, движутся к катоду. Так как их масса значительно больше, чем у электрона, то магнитное поле практически не влияет на траекторию движения ионов. Соударение положительных ионов с катодом приводит к появлению вторичных электронов, ток которых пропорционален ионному току.

Таким образом, разрядный ток магнитного преобразователя

$$I_p = I_\phi + I_H + I_B,$$

где I_ϕ — фоновый ток автоэлектронной эмиссии; I_H — ионный ток; I_B — ток вторичной электронной эмиссии.

Ток автоэлектронной эмиссии не зависит от давления и потому может считаться фоновым током; ионный и ток вторичной электронной эмиссии зависят от давления:

$$I_H + I_B = ap^n,$$

где $a = 10^{-2} \dots 10^{-1}$ А/Па и $n = 1 \dots 1,4$ — постоянные.

Учитывая эту зависимость и пренебрегая фоновым током, получим измерительное уравнение магнитного преобразователя

$$I_p = K_H p,$$

здесь $K_H = ap^{n-1}$ — чувствительность прибора. Разрядный ток магнитного преобразователя нелинейно зависит от давления.

Верхний предел измерения связан с ограничением максимального разрядного тока балластным сопротивлением, защищающим измерительный прибор от возникновения дугового разряда. Для расширения верхнего предела измерения следует уменьшить анодное напряжение и размеры разрядного промежутка. Обычно верхний предел измерения находится в области давлений 10—100 Па.

Нижний предел измерения определяется временем зажигания разряда и значением фонового тока. В современных приборах он составляет 10^{-11} Па. Для уменьшения фонового тока применяются специальные экраны 3 (рис. 2.12), расположенные в промежутке между катодом 2 и анодом 1, где напряженность электрического поля максимальна. Большая часть фонового тока в этом случае переходит на корпус, минуя микроамперметр, которым измеряется разрядный ток.

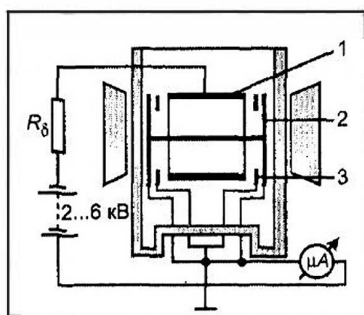


Рис. 2.12. Схема магнетронного преобразователя с уменьшенными фоновыми токами

Для обеспечения зажигания разряда при низких давлениях необходимо повышать анодное напряжение и увеличивать размеры разрядного промежутка.

Для облегчения зажигания разряда в сверхвысоком вакууме на экранных пластинах устанавливают острые иглы, увеличивающие автоэлектронную эмиссию. Наиболее надежным способом обеспечения быстрого зажигания разряда является использование нагреваемых элементов, включение которых приводит к резкому повышению давления и термоэмиссии электронов.

При применении сильных магнитных полей ($B > 0,1$ Тл) значение постоянной n в формуле, описывающей чувствительность прибора, стремится к единице. При этом диапазон работы прибора расширяется как в области низких, так и высоких давлений.

Магнитные преобразователи, так же как и электронные, имеют неодинаковую чувствительность к различным газам. Быстрота откачки колеблется для различных преобразователей в зависимости от рода газа и режимов работы в пределах от 10^{-2} до 1 л/с, что значительно больше, чем для электронных. Это приводит к увеличению погрешности измерений при наличии вакуумного сопротивления между преобразователем и вакуумной камерой.

Преимуществом магнитного преобразователя перед электронным является более высокая надежность в работе в связи с заменой накаливаемого катода холодным. К недостаткам можно отнести нестабильность, связанную с колебаниями работы выхода электронов при загрязнении катодов.

Эти нестабильности особенно заметны при работе преобразователя в вакуумных системах, где в качестве покрытий поверхности электродов используются масляные диэлектрические пленки. Продукты разложения паров масел, возникающие при ионной бомбардировке, могут в несколько раз изменить постоянную преобразователя. Во избежание этого необходимо применять самоочищающиеся магнитные преобразователи, работающие на переменном токе. В таких преобразователях катод и анод меняются местами в соответствии с полупериодами питающего напряжения, а очистка их поверхностей осуществляется ионной бомбардировкой.

На рис. 2.13 систематизированы методы измерения давления остаточных газов и диапазоны измерения. Штриховой линией обозначены предельные давления.

Современная вакуумная техника позволяет получать и измерять давления в 10^{18} раз меньше атмосферного, но даже такое состояние газа еще нельзя назвать идеальным вакуумом, поскольку в 1 м^3 такого вакуума содержатся сотни молекул газа. Идеальный вакуум как среда, в которой могут распространяться гравитационное, электромагнитное и другие поля, до сих пор является предметом тщательного исследования современной теоретической физики.

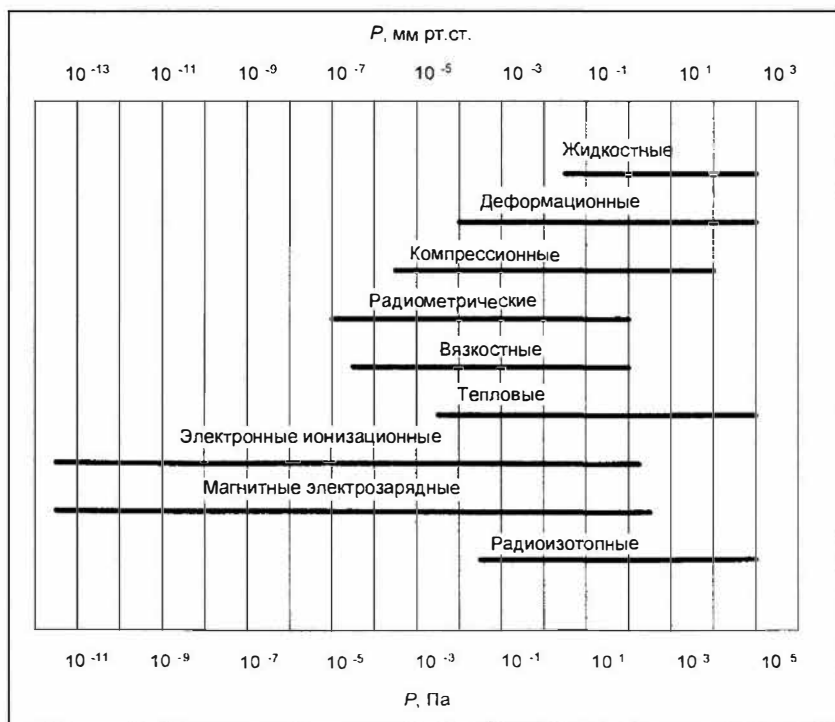


Рис. 2.13. Диапазоны измерения давления вакуумметрами различных типов

В квантовой теории поля введено такое понятие как физический вакуум. Это низшее энергетическое состояние квантовых полей, которое характеризуется отсутствием реальных частиц.

Джеймс Кларк Максвелл был одним из первых, кто исследовал вакуум. Он создал его физическую модель и на ее основе получил уравнения электромагнитного поля, известные как уравнения Максвелла. Автор рассматривал их как описание свойств вакуума. Согласно модели Максвелла вакуум представляет собой пространство, заполненное "молекулярными вихрями". Вращательное движение между вихрями передается через очень малые частицы, находящиеся между вихрями. В результате каждый вихрь заставляет вращаться соседние вихри. Сегодня такие вихри отождествляют с *солитонами*... Но это уже не относится к вакуумной электронике, поэтому тонкости этого вопроса отложим на конец книги.

2.4. Методы течеискания

Наряду с методами получения и измерения вакуума сохранение вакуума является важнейшим условием работы электронных приборов, включая электровакуумные, квантовые и другие устройства.

Одним из методов контроля приборов на герметичность являются *методы течеискания*, позволяющие устранить течи или место нарушения целостности оболочки и обеспечить герметичность приборов.

Величина течи определяется потоком натекающего через течь газа в единицу времени. Количество натекающего газа определяется произведением объема проникшего газа V на его давление P . В системе СИ поток измеряется в Вт, а на практике в качестве единицы измерения потока газа используется произведение 1 литра на давление микрон рт. ст. ($л \times мкм \text{ рт. ст.}$), что соответствует $1,33 \times 10^{-4}$ Вт.

Режим протекания газов через течи в вакуумных системах в зависимости от размеров и формы течи, природы газа, его температуры и давления может быть молекулярным, вязкостным и молекулярно-вязкостным.

Методика установления степени герметичности вакуумной системы сводится к подаче со стороны оболочки пробного, легко идентифицируемого вещества, а со стороны вакуумной среды фиксируются количественные изменения содержания пробного вещества. На сегодняшний день разработаны различные методы контроля герметичности вакуумной системы, среди которых следует отметить методы опрессовки и искрового разряда, манометрический, люминесцентный, манометрический, масс-спектрометрический метод и др.

Манометрический метод испытания на герметичность широко применяется, поскольку в каждой вакуумной установке есть хотя бы один вакуумметр, который может быть использован для выявления течей.

С помощью вакуумметра можно определить суммарный поток натекания через течи в систему и обнаружить места течей. Если при заливке жидкого азота в охлаждаемые ловушки показания вакуумметра не изменяются и составляют завышенное значение по сравнению с расчетными или полученными ранее, то это свидетельствует о наличии течей в системе. Для определения суммарного потока натекания через течи необходимо отсоединить испытываемую систему от средств откачки. По показаниям вакуумметра, снятым через определенные промежутки времени, необходимо построить зависимость давления от времени, определяемую суммарным потоком газовой выделений внутренних по-

верхностей системы и потоком натекания через течи (рис. 2.14). Суммарный поток определяется углом наклона касательной 1 к кривой 2 в ее начальной точке, а поток натекания — углом наклона прямолинейного участка кривой 2 к оси абсцисс.

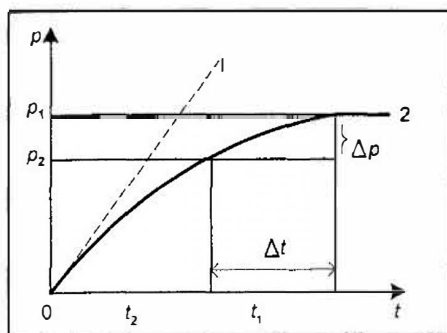


Рис. 2.14. Зависимость давления в изолированной вакуумной системе от времени

Для определения потока натекания необходимо совершить следующие действия:

1. Установить момент времени t_1 , в который начинается близкое к линейному изменение давления (при этом устанавливается равновесный поток газовой выделения и дальнейшее повышение давления в системе происходит только вследствие натекания через течи).
2. Измерить давление p_1 в вакуумной системе, соответствующее моменту времени t_1 .
3. Выждать время наблюдения Δt и измерить давление p_2 , соответствующее моменту времени t_2 .
4. Определить изменение давления Δp за время наблюдения Δt и суммарный поток натекания по формуле $Q = V \cdot \Delta p / \Delta t$, где V — объем проверяемой системы, м^3 ; Δt — время наблюдения, с.

В системах с большим газовойделением манометрический преобразователь целесообразно подсоединить через охлаждаемую ловушку, применение которой повышает эффективность поглощения продуктов газовой выделения и позволяет уменьшить время выхода на линейный участок характеристики (рис. 2.14). При измерениях уровень жидкого азота в ловушке должен быть постоянным.

При испытаниях газонаполненных систем испытуемый элемент помещают в вакуумную камеру, в которую подают воздух под избыточным давлением. О степени негерметичности судят по изменению показаний манометрических преобразователей вакуумной камеры при известном изменении избыточного давления в элементе. Место течи определяют по изменению показаний вакуумметра при подаче на отдельные участки поверхности пробного вещества. Наиболее чувствительны к течи неабсолютные манометрические преобразователи, показания которых зависят от рода газа. Для получения количественной оценки обнаруживаемых течей предварительной градуировкой манометрического преобразователя устанавливают его чувствительность по воздуху k_n и пробному газу k_p .

При работе с неабсолютным манометрическим преобразователем изменение его сигнала вследствие замены потока воздуха через течи потоком пробного вещества может быть обусловлено тремя факторами:

- ☐ различными скоростями натекания воздуха и пробного вещества через течи;
- ☐ различными значениями быстроты откачки воздуха и пробного вещества насосами испытуемой системы;

- различными чувствительностями манометрического преобразователя к воздуху и пробному веществу.

рекомендуется использовать пробное вещество, для которого все три фактора вызывают либо увеличение, либо уменьшение показаний вакуумметра, т. е. чтобы удовлетворялось неравенство

$$K_{\Pi} S_B Q_{\Pi} / K_B S_{\Pi} Q_B \gg 1 \quad (2.32)$$

или

$$K_{\Pi} S_B Q_{\Pi} / K_B S_{\Pi} Q_B \ll 1, \quad (2.33)$$

где S_B и S_{Π} — быстрота действия насоса при откачке воздуха и пробного вещества, соответственно, $\text{м}^3/\text{с}$; Q_B и Q_{Π} — поток через течь воздуха и пробного вещества, соответственно, $\text{Па} \cdot \text{м}^3/\text{с}$; K_B и K_{Π} — коэффициенты чувствительности манометрических преобразователей к воздуху и пробному веществу, соответственно, $\text{А}/\text{Па}$.

Используя манометрический преобразователь с линейной характеристикой, испытания второго этапа можно проводить с количественной оценкой обнаруживаемых течей. Для этого необходимо:

1. Снять показания p_1 манометрического преобразователя, отградуированного по воздуху до подачи пробного вещества на течь.
2. Обдуть пробным газом или смочить пробной жидкостью предполагаемые места течи и снять показания вакуумметра p_2 при вновь установившемся давлении.
3. Определить разницу давлений $\Delta p = p_2 - p_1$ по воздушному эквиваленту и рассчитать поток газа через течь, вызвавший реакцию вакуумметра для молекулярного режима:

$$B_M = v S_{\Pi} \Delta p / (\delta v - \beta). \quad (2.34)$$

Для вязкостного режима:

$$B_{\text{вз}} = \zeta S_{\Pi} \Delta p / (\delta \zeta - \beta), \quad (2.35)$$

где $v = M_{\Pi} / M_B$ — отношение молярных масс пробного вещества и воздуха; $\delta = S_{\Pi} / S_B$ — отношение эффективных скоростей реакции на пробное вещество и воздух; $\beta = k_{\Pi} / k_B$ — коэффициент относительной чувствительности манометрического преобразователя по пробному веществу; $\zeta = \eta_{\Pi} / \eta_B$ — отношение динамических вязкостей пробного вещества и воздуха.

Если в испытуемой системе применено селективное средство откачки с $S_{\Pi} = 0$, то обнаруженный поток газа через течь оценивают по формуле

$$B = V \Delta p_{\Pi} / (\Delta t \beta), \quad (2.36)$$

где Δp_{Π} — изменение давления пробного вещества за время Δt .

Жидкие пробные вещества вызывают большую реакцию вакуумметра, нежели пробный газ. Так, в частности, реакция вакуумметра на пары ацетона почти на три порядка больше, чем реакция на гелий.

Масс-спектрометрический метод основан на процессах ионизации газов и паров с последующим разделением и анализом ионов в магнитных и электрических полях. Метод отличается высокой чувствительностью и универсализмом. Основным элементом такого типа теческатора является масс-спектрометрический анализатор с магнитным отклонением пучка ионов. С помощью масс-спектрометрического теческатора контроль герме-

точности и поиск течей может осуществляться различными способами: обдув, накопление, использование гелиевых чехлов, щупа и вакуумных присосок. Масс-спектрометрические течеискатели выпускаются промышленностью (ПТИ-7А, ПТИ-9, ПТИ-10, СТИ-8) с соответствующими инструкциями по применению тех или иных способов поиска течей.

Задачи и упражнения

2.1. Средняя длина пробега электронов и молекул

Определите среднюю длину свободного пробега молекул и электронов.

Объясните смысл понятия "средняя длина свободного пробега", иллюстрируя ответ примерами перемещения как молекул, так и электронов.

Вычислите среднюю длину свободного пробега молекул неона (диаметр молекулы неона составляет 3×10^{-10} м), а также среднюю длину свободного пробега электронов в этом газе, предполагая, что газ неон — единственный газ, который присутствует в исследуемом объеме.

Решение

Средняя длина свободного пробега является важным параметром при рассмотрении проблемы электропроводности в газах. В общем виде ее можно определить как среднее расстояние, которое проходит частица в газе между последовательными соударениями.

Существуют два специфических понятия средней длины свободного пробега в газе, состоящем из одинаковых молекул:

- длина свободного пробега молекул;
- длина свободного пробега электронов.

Рассмотрим первый случай — длину свободного пробега молекул. Предположим, что в газе беспорядочно движется только одна молекула, остальные молекулы неподвижны. Эта молекула перемещается вдоль цилиндра диаметром d_m . При этом она будет соударяться с теми покоящимися молекулами, которые расположены внутри цилиндра, т. е. с теми молекулами, центры которых удалены от центра движущейся молекулы на расстояние не менее чем $2d_m$ (см. рис. 2.3).

Пусть \bar{v} — средняя скорость молекул при температуре T (К). Тогда за время dt молекула пройдет расстояние $\bar{v} \cdot dt$. Объем цилиндра, в котором находится молекула, равен $\bar{v} dt \cdot d^2$.

Следовательно, число столкновений определяется выражением

$$n = N\pi d^2 \bar{v} \cdot dt, \quad (2.1.1)$$

где N — число молекул в единице объема.

Среднее расстояние, пройденное молекулой между столкновениями, запишется в виде

$$\frac{\bar{v} dt}{n} = \frac{\bar{v} dt}{N\pi d^2 \bar{v} dt} = \frac{1}{\pi d^2 N}.$$

Действительно, если взять распределение по скоростям Максвелла — Больцмана, то средняя длина свободного пробега

$$l = \frac{1}{\sqrt{2}\pi d^2 N}. \quad (2.1.2)$$

Подставляя в это выражение величину N из уравнения состояния $p = kNT$, имеем

$$l = \frac{kT}{\sqrt{2}\pi d^2 p},$$

т. е. получаем, что

$$l \sim \frac{1}{p}. \quad (2.1.3)$$

Это соответствует ранее полученному соотношению (2.17, а).

Вычислим значение средней длины свободного пробега молекулы в газе, находящемся в условиях комнатной температуры и нормального давления:

$$l = \frac{1,38 \cdot 10^{-23} \cdot 300}{\sqrt{2} \cdot \pi \cdot 9 \cdot 10^{-20} \cdot 5 \cdot 10^{-7} \cdot 13,6 \cdot 9,81} = 155 \text{ м.}$$

Рассмотрим теперь среднюю длину свободного пробега электронов в идеальном газе. Диаметр электрона много меньше диаметра молекулы. Эффективный радиус цилиндра, в котором движется электрон, можно принять равным $d_e/2$. Более того, поскольку скорость электрона много больше скорости молекулы, множителем $\sqrt{2}$ можно пренебречь. Поэтому мы можем записать выражение для средней длины свободного пробега электрона в виде

$$l_e = \frac{4kT}{\pi d_e^2 p},$$

где p измеряется в Н/м².

Если давление измеряется в мм рт. ст., то, подставив значение постоянной Больцмана k в предыдущее выражение, получим

$$l_e = \frac{1,32 \cdot 10^{-25} T}{d_e^2 p}.$$

Используя числовые данные, находим значение l_e для объема при комнатной температуре и нормальном давлении:

$$l_e = \frac{1,32 \cdot 10^{-25} \cdot 300}{(0,3 \cdot 10^{-9})^2 \cdot 5 \cdot 10^{-7}} = 880 \text{ м.}$$

2.2. Энергия электрона и иона для ионизации

Какой минимальной кинетической энергией должны обладать электрон и ион для ионизации газа?

Решение

При ионизации газа сумма кинетических энергий сталкивающихся частиц изменяется за счет изменения внутренней энергии всех или некоторых из них.

Для получения условий ионизации атома электроном запишем уравнение сохранения энергии, считая для простоты, что атом до столкновения неподвижен, а электрон испытывает с ним лобовое столкновение.

Тогда

$$1/2 m_e v_{0e}^2 = 1/2 m_e v_{ie}^2 + 1/2 m_e v_{1e}^2 + \Delta E. \quad (2.2.1)$$

Здесь ΔE — энергия, переданная электроном атому и затраченная на увеличение его внутренней энергии. Ионизация может произойти, когда $\Delta E > qV_{ie}$. Согласно закону сохранения импульса

$$m_e v_{0e} = m_e v_{ie} + m_a v_{1a}. \quad (2.2.2)$$

Из совместного решения (2.2.1) и (2.2.2) находим, что

$$\Delta E_{\max}/E_0 = m_a/(m_e + m_a). \quad (2.2.3)$$

где $E_0 = 1/2 m_e v_{0e}^2$. Из условия (2.2.3) при $m_a \gg m_e$ видно, что на изменение внутренней энергии атома может быть израсходована почти вся кинетическая энергия электрона, т. е. $\Delta E_{\max} \approx E_0$ и ионизация имеет место при $E_0 \geq qV_{ie}$.

При столкновениях тяжелых одинаковых частиц $\Delta E_{\max} \approx E_0/2$ и, следовательно, для ионизации необходимо, чтобы $E_0 \geq 2qV_{ie}$.

2.3. Градуировка ионизационного манометра

Ионизационный манометр представляет собой плоскую триодную систему, которая схематически изображена на рис. 2.15. Расстояние между сеткой и катодом составляет 0,3 см, а между сеткой и коллектором — 1,0 см. Потенциал сетки $U_g = 250$ В, потенциал коллектора $U_a = -20$ В. Геометрическая прозрачность сетки составляет $p = 0,8$. Остаточный газ — азот.

Определить давление газа в лампе, если ионный ток на коллектор равен 5×10^{-7} А. Электронный ток I_e , эмитируемый катодом, равен 10^{-2} А.

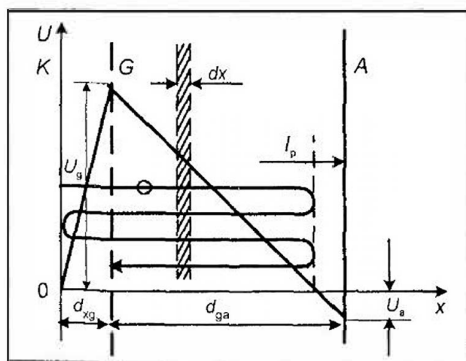


Рис. 2.15. Схема ионизационного манометра

Решение

Число актов ионизации на элементе длины пути за единицу времени составит

$$dn_i = n_e \omega_i \cdot dv,$$

где $n_e = I_e/q$ — число электронов в потоке, $dv = dx/\lambda$ — число соударений, испытываемых электроном на пути dx .

Вероятность ионизации w_i для случая $U > U_i$ аппроксимируется функцией

$$w_i = \alpha \cdot (U - U_i) \cdot \exp\left[-\frac{(U - U_i)}{\beta}\right],$$

где U_i — потенциал ионизации, α и β — константы, зависящие от рода газов.

Найдем di_p — ток от ионов, возникающих на элементе пути dx за единицу времени

$$di_p = N e d n_i = N i_e \alpha \cdot (U - U_i) \cdot e^{-\frac{U - U_i}{\beta}} \frac{dx}{\lambda},$$

где N — среднее число колебаний, совершаемых электроном вокруг сетки.

Для удобства интегрирования осуществляем замену переменной x на переменную $U - U_i$

$$E = \frac{U_g - U_a}{d_{ог}},$$

$$dx = \frac{dU}{E} = \frac{d(U - U_i)}{E};$$

$$i_p = N \frac{i_e \alpha \beta^2}{E \lambda_0} \int_{U_i}^{U_g} -\frac{(U - U_i)}{\beta} \exp[-(U - U_i)/\beta] d\left[\frac{U - U_i}{\beta}\right].$$

После подстановки пределов и значения E получим значение ионного тока i_p на коллектор

$$i_p = \frac{N i_e \alpha \beta p d_{ог}}{(U_g - U_a) \lambda_0} \left[\beta - (U_g + \beta - U_i) \exp\left(-\frac{U_g - U_i}{\beta}\right) \right];$$

$$p = \frac{i_p \lambda_0 (U_g - U_a)}{N i_e \alpha \beta p d_{ог} \left[\beta - (U_g - U_i + \beta) \exp\left(-\frac{U_g - U_i}{\beta}\right) \right]}.$$

Вероятность захвата электрона сеткой при одном пересечении им плоскости сетки составляет ориентировочно $1 - q$.

При однократном прохождении электронного потока сквозь сетку проходит 80% всех электронов, а захватывается 20%.

Таким образом, за первое колебание сквозь сетку пройдет n_0 электронов, за второе — $n_0 q$, за третье — $n_0 q^2$ и т. д.

$$n_m = n_0 q^{m-1}.$$

Суммируя эту геометрическую прогрессию для $m \rightarrow \infty$ и разделив сумму на число электронов, получим, что один электрон совершает в среднем N проходов сквозь сетку

$$N = \sum \frac{n_m}{n_0} = \frac{1}{1 - q} = 5.$$

Такая оценка была бы вполне верной, если бы геометрическую прозрачность сетки можно было отождествить с ее электрической прозрачностью. Однако поскольку сетка притягивает пролетающие электроны, в действительности N будет несколько меньше вычисленной величины.

Вычислим значение p , взяв из условия задачи и таблиц следующие значения:

$$I_p = 5 \times 10^{-7} \text{ А}; \quad \lambda_0 = 3,58 \times 10^{-4}; \quad U_g - U_a = 270 \text{ В}; \quad N = 5; \quad \alpha = 7,8 \times 10^{-3}; \quad \beta = 160.$$

Тогда

$$p = \frac{5 \cdot 10^{-7} \cdot 3,58 \cdot 10^{-4} \cdot 270}{5 \cdot 7,8 \cdot 10^{-3} \cdot 10^{-3} \cdot 160 \cdot 10^{-2} \cdot \left[160 - (250 - 16 + 160) \exp\left(-\frac{250 - 16}{160}\right) \right]} = 1,12 \cdot 10^{-6} \text{ мм рт. ст.}$$

Контрольные вопросы

1. Какие постулаты легли в основу физической теории вакуума?
2. Что такое давление газа?
3. В каких единицах измеряется давление газа? Приведите соотношения между этими единицами.
4. Что такое распределение Максвелла?
5. Что такое средняя длина пробега частицы в газе? Как ее определить?
6. Как на основе критерия Кнудсена классифицируют степень разряжения газов?
7. Какие принципы лежат в основе методов получения вакуума?
8. Что такое абсолютный манометр? Приведите примеры.
9. Каков принцип работы тепловых вакуумметров?
10. Каков принцип работы электронно-ионизационных вакуумметров?
11. На каких принципах работают магнитные вакуумметры? Что такое ячейка Пеннинга?
12. Какие методы течения вы знаете?
13. Расскажите о манометрическом методе течения.

Рекомендуемая литература

1. Вакуумная техника: Справочник. Под ред. Е. С. Фролова, В. Е. Минайчева. — М.: Машиностроение, 1992.
2. Дешман С. Научные основы вакуумной техники. — М.: Мир, 1963.
3. Левитский С. М. Сборник задач и расчетов по физической электронике. — Издательство Киевского университета, 1964.
4. Линч П., Николайдес А. Задачи по физической электронике. Под ред. проф. Г. В. Скроцкого. — М.: Мир, 1975.
5. Розанов Л. Н. Вакуумная техника. — М.: Высшая школа, 1990.
6. Швилкин Б. Н., Мискинова И. А. Физическая электроника в задачах. — М.: Наука, 1987.
7. Шейн Е. П. Основы вакуумной техники. — М.: МФТИ, 2001.

3. Вакуумная электроника

Вакуумная электроника — это раздел электроники, включающий в себя исследования взаимодействия потока свободных электронов с электрическими и магнитными полями в вакууме, а также методы создания электронных приборов и устройств, в которых это взаимодействие используется.

3.1. Модель прибора вакуумной электроники

Приборы вакуумной электроники можно классифицировать в зависимости от физического принципа действия, назначения и технологии производства. В основу классификации может быть положена и предложенная модель приборов вакуумной электроники.

Изучение приборов и устройств электроники целесообразно проводить с системных позиций. С этой целью введем модель прибора вакуумной электроники, в которой будем различать пять основных элементов:

1. Носитель информационного сигнала — ансамбль свободных электронов.
2. Генератор или эмиттер свободных электронов.
3. Континуальная среда, которой служит вакуум для электровакуумных приборов, или плазма для ионных приборов.
4. Устройство управления ансамблем электронов в континуальной среде (физические поля).
5. Детектор информационных сигналов или устройство отбора энергии от электронного потока.

На рис. 3.1 приведена схематическая модель прибора вакуумной электроники.

При таком системном подходе необходимо исследовать ряд фундаментальных проблем, среди которых:

- свойства отдельного электрона и ансамбля свободных электронов;
- свойства вакуумных континуальных сред, способы создания и измерения вакуума;
- различные способы эмиссии свободных электронов, законы электронной эмиссии из твердого тела, формирование электронных потоков;
- распространение электронных потоков в вакууме, способы управления электронными потоками с помощью электрического, магнитного и скрещенных электромагнитных полей, фокусировка электронных пучков, группирование электронов;
- детектирование информационных сигналов на основе взаимодействия потока электронов с мишенями из различных по физическим свойствам веществ, отбор энергии из электронных потоков.

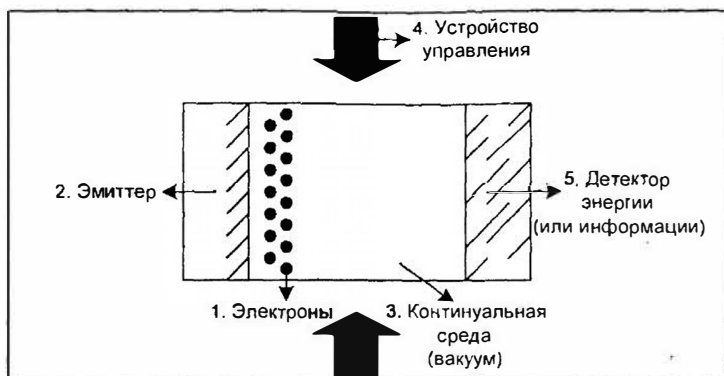


Рис. 3.1. Модель прибора вакуумной электроники

3.2. Электронная эмиссия

Электронная эмиссия — это явление испускания электронов поверхностью твердых тел в результате внешних физических воздействий.

Потенциальная энергия электронов $U(x)$ вне твердого тела и в отсутствии силовых полей принимается равной нулю (рис. 3.2).

У движущегося вне тела электрона полная энергия E_H является величиной положительной.

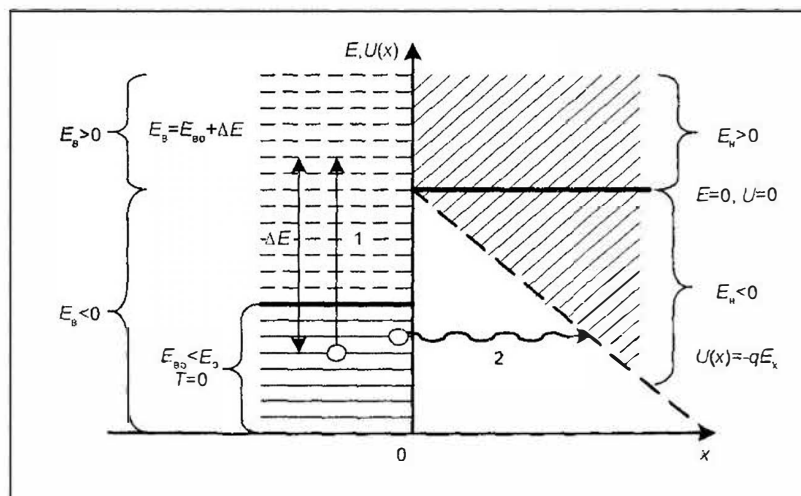


Рис. 3.2. Энергетическая диаграмма состояний электронов на границе "поверхность твердого тела — вакуум"

Внутри тела электроны в невозбужденном состоянии ($T = 0$) занимают низкие энергетические уровни $E_{b0} \leq 0$. Электроны в этих состояниях не могут покинуть тело. Для эмиссии

электронов необходимо возбудить электроны, другими словами, сообщить электронам дополнительную энергию ΔE , значение которой должно быть таковым, чтобы величина $E_{\text{в}} + \Delta E$ стала положительной. Тогда возможен переход электронов по стрелке 1 (рис. 3.2). Этот тип эмиссии называется эмиссией с предварительным возбуждением. Существуют различные методы возбуждения электронов в твердом теле и, соответственно, типы электронной эмиссии:

- ☐ термоэлектронная эмиссия;
- ☐ фотоэлектронная эмиссия;
- ☐ вторичная электронная эмиссия;
- ☐ кинетическая ионно-электронная эмиссия;
- ☐ эмиссия горячих электронов;
- ☐ экзотермическая эмиссия.

Возможна также эмиссия электронов без предварительного возбуждения электронов. Это может быть достигнуто путем снижения потенциального барьера. Внешнее однородное электрическое поле E способствует снижению этого барьера. На расстоянии x от границы тела энергия убывает по закону

$$U(x) = U_0 - eEx,$$

где U_0 — потенциальная энергия на границе тела.

При наличии силовых полей у поверхности твердого тела энергия электронов может быть отрицательной и принимать значения $E_H \approx E_{\text{в}} \leq E_0$. Эти области разделены пространственно-потенциальным барьером $U(x) \geq E_{\text{в}}$. Согласно законам квантовой механики возможен переход электронов из тела во внешнее пространство путем туннельного эффекта по стрелке 2 (рис. 3.2). Эмиссия электронов в этом случае не требует предварительного возбуждения. К этому типу эмиссии относится автоэлектронная эмиссия.

Возможен и комбинированный тип эмиссии, при котором действуют оба механизма. К этому типу эмиссии относятся:

- ☐ термоавтоэлектронная эмиссия;
- ☐ фотоавтоэмиссия;
- ☐ потенциальная ионно-электронная эмиссия.

Рассмотрим важнейшие типы электронной эмиссии.

3.2.1. Термоэлектронная эмиссия

Термоэлектронная эмиссия — это явление испускания электронов нагретыми телами в вакуум или в другую среду.

Для преодоления потенциального барьера на границе "твердое тело—вакуум" необходимо, чтобы их энергия была больше работы выхода электронов.

Значение тока термоэлектронной эмиссии поверхности в вакуум при температуре T (К) определяется формулой Ричардсона — Дешмана:

$$j(T) = AT^2(1-r)\exp[-\phi(T)/kT], \quad (3.1)$$

где A — постоянная величина, для металлов определяемая как $A = 120,4 \text{ А/см}^2 \cdot \text{К}^2$, k — постоянная Больцмана, r — коэффициент отражения электронов от потенциального барьера на границе "твердое тело—вакуум", T — температура, К.

работа выхода электронов ϕ зависит как от температуры, так и от состояния эмитирующей поверхности, наличия на поверхности адсорбированных атомов и молекул.

Для эмиттеров на основе собственных полупроводников формула Ричардсона — Дешмана принимает следующий вид:

$$j(T) = A_p T^2 (1 - r) \exp[-\phi_p(T)/kT], \quad (3.2)$$

где $A_p = A \cdot (1 - r) \cdot \exp[-\alpha/k]$, α — температурный коэффициент работы выхода при температуре эмиссии, ϕ_p — экстраполированная к $T = 0$ °K величина $\phi(T)$.

Вся комбинация сомножителей называется ричардсоновской постоянной термоэмиссии A_p .

Уравнение (3.2) можно записать в виде:

$$j = A_p T^2 \exp\left[-\frac{\phi_p}{kT}\right], \quad (3.3)$$

где A_p — ричардсоновская термоэлектронная постоянная, а ϕ_p — ричардсоновская или термоэлектронная работа выхода.

Связь между истинной, эффективной и ричардсоновской работами выходов определяется выражением:

$$j = A_0 (1 - \bar{R}) T^2 \exp\left(-\frac{\phi_{ист}}{kT}\right) = A_0 T^2 \exp\left(-\frac{\phi_{эфф}}{kT}\right) = A_p T^2 \exp\left(-\frac{\phi_p}{kT}\right). \quad (3.4)$$

Для определения величин A_p и ϕ_p из экспериментальных данных применяется так называемый "метод прямых Ричардсона". После десятичного логарифмирования формулы (3.4) она приобретает вид:

$$\lg\left(\frac{j}{T^2}\right) = \lg(A_p) - \phi_p \cdot \frac{5040}{T}.$$

Зависимость $j(T)$, построенная в координатах $y = \lg(j/T^2)$ и $x = 5040/T$, представляет собой прямую, коэффициент наклона которой равен ϕ_p , а отрезок, отсекаемый этой прямой на оси ординат, численно равен $\lg(A_p)$. Коэффициент наклона ричардсоновской прямой дает значение ϕ_p непосредственно в электронвольтах (рис. 3.3).

Термоэлектронная эмиссия неоднородных по значению работы выхода электронов ϕ поверхностей, именуемых "пятнистыми", вычисляется по уравнению Ричардсона — Дешмана при условии введения усредненных характеристик эмиссии A_p^* , ϕ_p^* .

Под действием внешнего ускоряющего электрического поля понижается потенциальный энергетический барьер, вследствие чего уменьшается работа выхода электронов на величину $\Delta\phi \approx q^{3/2} \cdot E^{1/2}$. Это так называемый эффект Шоттки, благодаря которому растет эмиссионный электронный ток. Величина E не должна превышать значения 10^5 В/см, при котором начинается автоэлектронная эмиссия.

При создании электрического поля у поверхности полупроводников наблюдается не только понижение внешнего потенциального барьера, но и проникновение электрического поля внутрь полупроводника. Глубина проникновения электрического поля зависит от концентрации свободных зарядов. Это приводит к значительно большему влиянию электрического поля на работу выхода электронов из полупроводников, чем у металлов.

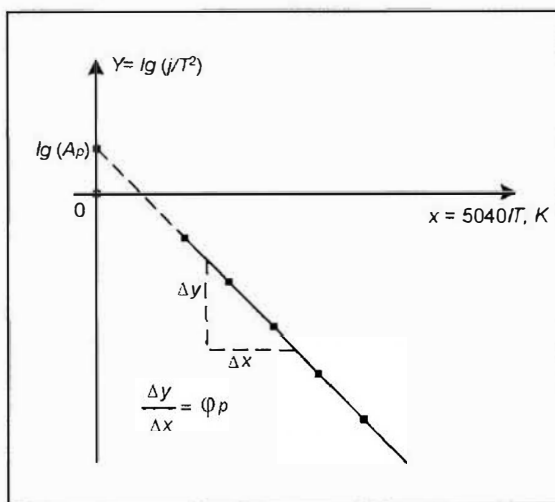


Рис. 3.3. Метод "прямых Ричардсона": теоретическая прямая и экспериментальные точки иллюстрируют зависимость $\lg(j/T^2)$ от $5040/T$

3.2.2. Фотоэлектронная эмиссия

Фотоэлектронная эмиссия или *внешний фотоэффект* — это явление испускания электронов под действием квантов электромагнитного излучения.

Фотоэлектронная эмиссия была открыта в 1887 году немецким физиком Генрихом Рудольфом Герцем. В 1888 году русский ученый Александр Григорьевич Столетов, активно исследовавший явление фотоэффекта, установил пропорциональность тока фотоэлектронной эмиссии интенсивности светового потока.

Закон Столетова гласит: при неизменном спектральном составе излучения сила тока фотоэмиссии (т. е. количество эмитируемых в единицу времени электронов) пропорциональна интенсивности падающего излучения.

В 1905 году немецкий физик Альберт Эйнштейн установил закон сохранения энергии для фотоэлектронов в виде:

$$h\nu = \varphi + E_{\text{кин. макс}}, \quad (3.5)$$

где h — постоянная Планка, φ — работа выхода электронов, $E_{\text{кин. макс}}$ — максимальная кинетическая энергия эмитированных электронов. Из этого соотношения следует, что максимальная кинетическая энергия эмитируемых электронов не зависит от интенсивности излучения при неизменном спектральном составе излучения и линейно возрастает с его частотой.

Если кинетическая энергия эмитированных электронов близка к нулевому значению, то значение частоты $\nu = \nu_0$ называется пороговой, а соответствующая ей длина волны излучения λ_0 — красной или длинноволновой границей внешнего фотоэффекта.

Энергию кванта электромагнитного излучения (фотона) удобно выражать во внесистемных единицах (электронвольтах и микроэВ):

$$h\nu(\text{эВ}) = hc/\lambda = 1,24/\lambda \text{ (мкм)} \quad (3.6)$$

или

$$1,24/\lambda(\text{мкм}) = 2,84 \times 10^{-12} \cdot \nu^2(\text{м/с}) + \varphi(\text{эВ}). \quad (3.7)$$

Из выражения (3.5) видно, что при предельном условии $\nu_{\text{гмак}} = 0$:

$$hc/\lambda_{\text{кр}} = \varphi, \quad (3.8)$$

где $\lambda_{\text{кр}}$ — длинноволновая граница фотоэффекта; φ — работа выхода фотокатода.

В металлах положение уровня Ферми, определяющее работу выхода, совпадает с наивысшим заполненным уровнем, с которого и происходит фотоэмиссия.

В 1931 году английский физик Ральф Говард Фаулсер построил полуфеноменологическую теорию фотоэмиссии металлов. В основе этой теории лежит предположение, что у электронов, поглотивших энергию фотона, энергия увеличивается, что эквивалентно понижению высоты поверхностного потенциального барьера. Проведя выкладки в том же порядке, как и при выводе уравнения Ричардсона для термоэлектронной эмиссии, можно получить выражение для плотности тока фотоэмиссии:

$$j_{\text{ф}} = aA(1-r)T^2 f(\Delta), \quad (3.9)$$

где $\Delta = \frac{h\nu - h\nu_0}{kT}$, ν_0 — пороговая частота, k — постоянная Больцмана, h — постоянная Планка, а

$$f(\Delta) = -\sum (-1)^n \exp(n\Delta)/n^2, \text{ при } \Delta \leq 0 \text{ или } \nu \leq \nu_0;$$

$$f(\Delta) = \pi/6 + \Delta/2 + \sum (-1)^n \exp(-n\Delta)/n^2, \text{ при } \Delta \geq 0 \text{ или } \nu \geq \nu_0.$$

Величины A , r и T соответствуют значениям величин в формуле Ричардсона — Дешмана. Заместим, что при $\nu = \nu_0$ значение $f(\Delta) = \pi/12$.

В полупроводниках уровень Ферми находится в запрещенной зоне, а наивысший заполненный уровень, с которого происходит фотоэмиссия электронов, соответствует верхней границе валентной зоны E_v . Порог фотоэффекта или минимальная энергия, которую нужно сообщить электрону, чтобы перевести его на уровень вакуума, соответствует энергии от потолка валентной зоны до уровня вакуума. Другими словами, порог фотоэффекта полупроводников равен сумме ширины запрещенной зоны E_G и энергии электронного средства E_a :

$$h\nu = E_G + E_a.$$

Работа выхода электронов из полупроводников равна разности энергетических уровней от вакуума до уровня Ферми $\varphi = E_0 - E_F$.

На рис. 3.4 представлены энергетические диаграммы собственного полупроводника (i), а также полупроводника с электронной (n) и дырочной (p) проводимостью.

Фотоэмиссия из металлов связана с поглощением фотонов электронами проводимости. Фотоэлектронная эмиссия полупроводников определяется процессами возбуждения электронов из валентной зоны, либо с уровней, связанных с донорной или акцепторной примесями.

Фотоэлектронную эмиссию можно интерпретировать как результат трех последовательных процессов. На первой стадии поглощается фотон и появляется возбужденный электрон с энергией выше средней. На второй стадии возбужденный электрон движется к по-

верхности, рассеивая часть энергии на других электронах, на дефектах кристаллической решетки, а также на фононах. На третьей стадии электрон преодолевает потенциальный барьер на границе раздела.

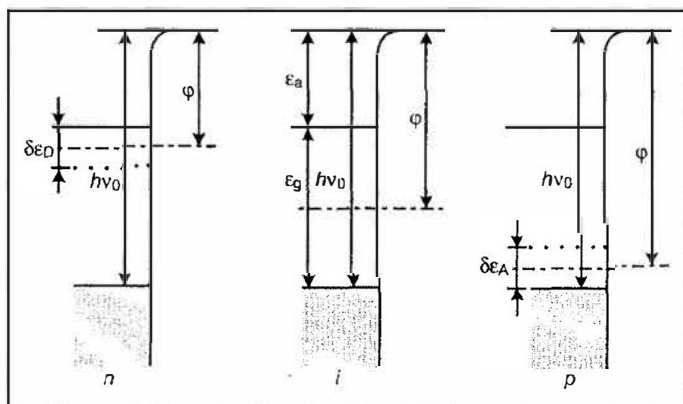


Рис. 3.4. Энергетические диаграммы полупроводника с собственной, электронной и дырочной проводимостью

Важной количественной характеристикой фотоэлектронной эмиссии является квантовый выход ξ или число эмитированных электронов, приходящихся на один фотон, падающий на поверхность. Квантовый выход электронов из металлов в видимой и ближней УФ-области спектра составляет $\xi \approx 10^{-3}$ электрон/фотон.

Для полупроводников величина ξ , главным образом, определяется процессами рассеяния энергии электронов при их движении к границе раздела, а также начальной энергией фотонов. Эффективные фотокатоды имеют квантовый выход $\xi \approx 0,1$ электрон/фотон.

При большой интенсивности падающего излучения возникает многофотонная фотоэлектронная эмиссия. Она связана с одновременным поглощением электроном двух и более фотонов.

Эмитированные в результате фотоэффекта электроны несут в себе информацию об энергии и силе электрона в твердом теле. Эта информация позволяет определить электронную структуру твердых тел, состояние их поверхности, направленность валентных связей, время жизни возбужденных электронов и дырок, квантовый выход и другие параметры твердого тела.

3.2.3. Вторичная электронная эмиссия

Вторичная электронная эмиссия — это явление испускания твердыми телами вторичных электронов при их бомбардировке первичными электронами.

Поток вторичных электронов складывается из упруго и неупруго отраженных первичных электронов и истинно вторичных электронов. Вторичные электроны имеют непрерывный энергетический спектр от нуля до энергии первичных электронов. На рис. 3.5 приведено распределение вторичных электронов по энергиям, эмитируемых вольфрамом.

Область спектра I соответствует упруго отраженным первичным электронам, область II — неупруго отраженным первичным электронам, и область III характеризует энергетический спектр вторичных электронов.

ческий спектр истинно вторичных электронов. Каждое явление характеризуется своим коэффициентом:

- r — коэффициент упруго отраженных электронов, равный отношению числа упруго отраженных электронов к числу первичных электронов;
- η — коэффициент неупругоотраженных электронов, равный отношению числа неупруго отраженных электронов к числу первичных электронов;
- δ — коэффициент истинной вторичной электронной эмиссии, равный отношению числа истинно вторичных электронов к числу первичных электронов.

Полный коэффициент вторичной электронной эмиссии σ определяется соотношением:

$$\sigma = r + \eta + \delta. \quad (3.10)$$

Все составляющие коэффициента вторичной электронной эмиссии в первую очередь зависят от величины работы выхода электронов, а также от параметров пучка первичных электронов и свойств эмиттера.

В диэлектриках и эмиттерах с широкой запрещенной зоной и малым сродством к электрону коэффициент σ имеет значение порядка 100.

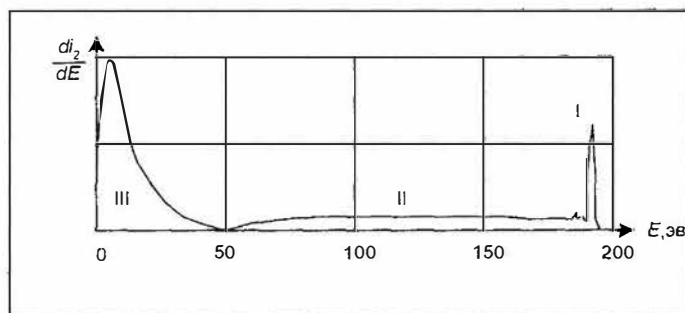


Рис. 3.5. Энергетический спектр вторичных электронов

3.2.4. Кинетическая ионно-электронная эмиссия

Кинетическая ионно-электронная эмиссия — это явление испускания электронов поверхностью твердого тела под действием поступательной энергии ионов (кинетическое выбивание).

Это явление обнаружил в 1899 году французский физик Поль Ульриш Вилард. Однако поскольку в то время отсутствовали надежные методы очистки поверхности, устойчивых закономерностей получить не удалось. Ионно-электронная эмиссия характеризуется коэффициентом $\gamma_k = n_e / n_p$, где n_e — число эмитированных электронов, n_p — число попавших на поверхность ионов за тот же промежуток времени.

3.2.5. Эмиссия горячих электронов

Эмиссия горячих электронов — это явление испускания электронов из полупроводников, электронный газ которого перегрет.

Существуют различные механизмы создания перегретых электронов, импульсы которых определяются как $p_{\text{ср}} \geq (2m^* \chi_{\text{ср}})^{1/2}$, где m^* — эффективная масса электрона в зоне проводимости полупроводника, $\chi_{\text{ср}}$ — средство электрона к эмиттеру. Одним из таких механизмов является создание в полупроводнике сильного электрического поля.

Уравнение для плотности тока эмиссии горячих электронов может быть описано уравнением типа уравнения Ричардсона:

$$j = en \left(\frac{k}{2\pi m^*} \right)^{1/2} (1-r) T_j^{1/2} \exp \left(-\frac{\chi_{\text{ср}}}{kT_j} \right), \quad (3.11)$$

где T_j — электронная температура в полупроводнике, зависящая от электрического поля в полупроводнике.

3.2.6. Экзоэлектронная эмиссия

Экзоэлектронная эмиссия (экзоэмиссия) — это явление нестационарного испускания электронов твердым телом, предварительно переведенным в термодинамически неустойчивое состояние.

Существуют различные механизмы перевода твердого тела в неустойчивое состояние. Способностью к экзоэлектронной эмиссии обладают металлы, полупроводники и диэлектрики.

Токи экзоэмиссии не превышают 10^{-13} А/см² и регистрируются с помощью электронных детекторов.

3.2.7. Автоэлектронная эмиссия

Автоэлектронная эмиссия — это явление испускания электронов твердыми телами под действием внешнего электрического поля высокой напряженности ($E \geq 10^7$ В/см).

Автоэлектронная эмиссия также носит название *полевая электронная эмиссия, туннельная эмиссия, электростатическая эмиссия, разновидность холодной эмиссии*. Автоэлектронная эмиссия не требует предварительного возбуждения электронов.

Автоэлектронная эмиссия была открыта американским физиком Робертом Уильямсом Вудом в 1887 году. В 1928—1929 годах американский физик Ральф Говард Фаулер и немецкий физик Лотар Вольфганг Нордгейм дали теоретическое объяснение автоэлектронной эмиссии на основе туннельного эффекта.

Плотность тока автоэлектронной эмиссии может быть вычислена по формуле Фаулера — Нордгейма:

$$j = A \frac{E^2}{\phi} \exp \left[-B \frac{\phi^{3/2}}{E} \theta(\gamma) \right], \quad (3.12)$$

где $A = \frac{q^2}{8\pi h}$ и $B = \frac{8\pi (2m)^{1/2}}{3 hq}$ — константы, $\theta(\gamma)$ — функция Нордгейма, затабулированная

Фаулером и Нордгеймом (рис. 3.6, а). Значения функции Нордгейма $\theta(\gamma)$ лежат в диапазоне 0,7—0,9. Теория Фаулера — Нордгейма применима только для температуры $T = 0$.

При $T > 0$, когда $kT \ll \phi$, зависимость для плотности тока остается в силе, поскольку тепловое возбуждение электронов лишь немного размывает границу Ферми (в пределах нескольких kT).

Прямые Фаулера — Нордгейма в координатах $\lg(1/V)$ и $1/V$ строятся аналогично прямым Ричардсона (см. рис. 3.3). С помощью этих прямых можно оценить площадь эмитирующей поверхности автокатода и коэффициент пропорциональности β в зависимости напряженности электрического поля E от приложенного напряжения V :

$$E = \beta \cdot V.$$

При достаточно высоких температурах вклад термически возбужденных электронов в общий эмиссионный ток становится заметным и эмиссия имеет место уже термоавтоэлектронная.

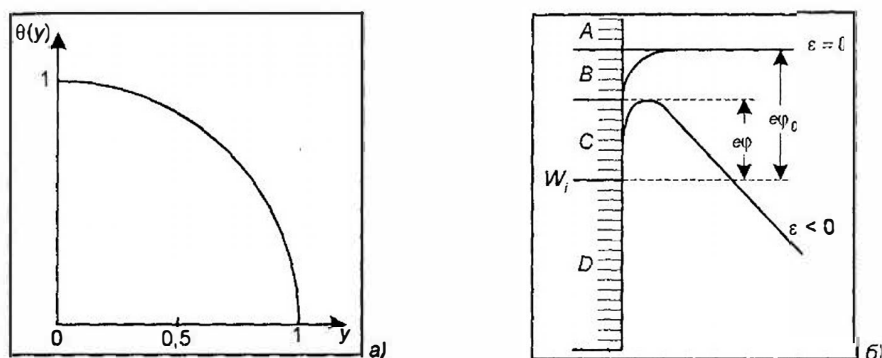


Рис. 3.6. Функция Нордгейма (а) и уровни энергии групп электронов (б)

Такая ситуация характеризуется условным разбиением электронов по энергиям на четыре группы: A , B , C и D .

Эмиссия автоэлектронов из группы D возможна при любых значениях температуры, включая $T = 0$.

Из группы C автоэлектроны могут быть эмитированы при температуре $T > 0$. Переход электронов из группы D в группу C позволяет увеличить ток автоэлектронной эмиссии.

Электроны группы B позволяют увеличивать ток термоэлектронной эмиссии за счет снижения потенциального барьера при увеличении электрического поля.

Электроны группы A обуславливают термоавтоэлектронную эмиссию даже при электрических полях $E = 0$.

Характерной особенностью автоэлектронной эмиссии являются высокие плотности токов вплоть до 10^8 А/см².

При перегрузке автокатода током может возникнуть явление взрывной электронной эмиссии. Оно обусловлено переходом материала эмиттера из конденсированной фазы в плотную плазму. С помощью взрывной электронной эмиссии удастся получить потоки электронов мощностью до 10^{13} Вт и плотностью тока до 10^9 А/см². Взрывная эмиссия электронов используется в силовых приборах силовой электроники. Основополагающие работы принадлежат советскому физiku Геннадию Андреевичу Месяцу.

Картина автоэлектронной эмиссии полупроводников усложняется за счет проникновения внешнего электрического поля в приповерхностный слой, изгиба энергетических зон, уменьшения приповерхностной концентрации электронов. В настоящее время нет полной физической теории автоэлектронной эмиссии полупроводников.

3.2.8. Потенциальная ионно-электронная эмиссия

Потенциальная ионно-электронная эмиссия — это явление испускания электронов за счет возбуждения электронов тела потенциальной энергией ионов.

Впервые это явление открыл нидерландский физик Франс Мишель Пеннинг. Потенциальная ионно-электронная эмиссия имеет место лишь для ионов таких элементов и мишеней, для которых выполняется соотношение $V_i \geq 2\phi$, где V_i — ионизационный потенциал атома, а ϕ — работа выхода мишени.

Потенциальная ионно-электронная эмиссия характеризуется величиной $\gamma_p = n_e / n_p$, где n_e — число эмитированных электронов, n_p — число попавших на поверхность ионов за тот же промежуток времени.

3.3. Эмиттеры свободных электронов

3.3.1. Электронная пушка

Для формирования электронного пучка (потока) используется *электронная пушка*, представляющая собой электронное устройство, состоящее из катода специальной формы, ускоряющего электрода — анода и фокусирующих электродов, предназначенных для формирования пучка необходимой формы.

Существует большое разнообразие конструкций электронных пушек.

Одной из характеристик электронной пушки является *первеанс* P , величина которого равна отношению тока электронного потока к ускоряющему анодному напряжению в степени трех вторых:

$$P = A/U^{3/2} [A \cdot B^{3/2}].$$

Первеанс является мерой интенсивности электронных потоков.

Потоки с $P < 10^{-8} A \cdot B^{3/2}$ являются низкоинтенсивными (их генерирует электронный проектор), а потоки с $P \geq 10^{-8} A \cdot B^{3/2}$ относятся к высокоинтенсивным или высокопервеансным, и для их генерации используется электронная пушка.

Наиболее распространенным типом электронных пушек является пушка Пирса (рис. 3.7), в которой геометрией электродов достигается распределение потенциала так же, как в диодной системе, а нормальная составляющая напряженности электрического поля на границе электронного потока равна нулю.

В электронных пушках со сходящимся электронным потоком наблюдается *кроссовер* (Kp) или сечение с минимальным радиусом пучка и нулевыми радиальными составляющими скоростей электронов. Наличие кроссовера — необходимое условие формирования пучка в магнитном поле, в системах с магнитной периодической фокусировкой и электростатической фокусировкой.

Отношение плотности тока в кроссовере к плотности тока катода называется *компресией* (\mathfrak{K}).

При равномерной плотности тока

$$\mathfrak{K} = \left(\frac{r_K}{r_{KP}} \right)^2 \sim 20 \div 50.$$

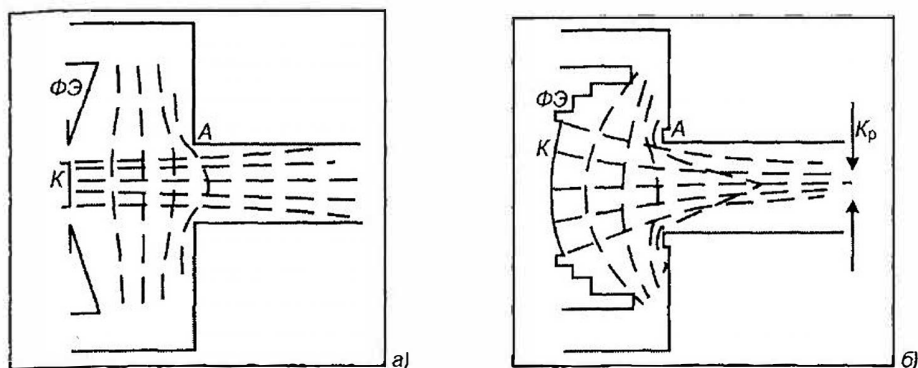


Рис. 3.7. Пушка Пирса, формирующая аксиально-симметричный электронный поток:
а — расходящийся поток; б — сходящийся поток; К — катод; А — анод;
ФЭ — фокусирующий электрод; Кр — кроссовер

С помощью электронных пушек различной конструкции можно сформировать трубчатые пушки электронов, полые пучки, ленточные пучки, многолучевые электронные пучки (рис. 3.8).

Электронный пучок — поток электронов, движущийся по близким траекториям в одном направлении, имеющий размеры в направлении движения значительно больше, чем в конечной плоскости. Электронный поток состоит из одноименно заряженных частиц, формирующий внутри потока пространственный заряд. В свою очередь этот заряд создает собственное электрическое поле.

Электроны в электронном пучке движутся по близким траекториям и формируют в пучке линейные токи, создающие собственные магнитные поля.

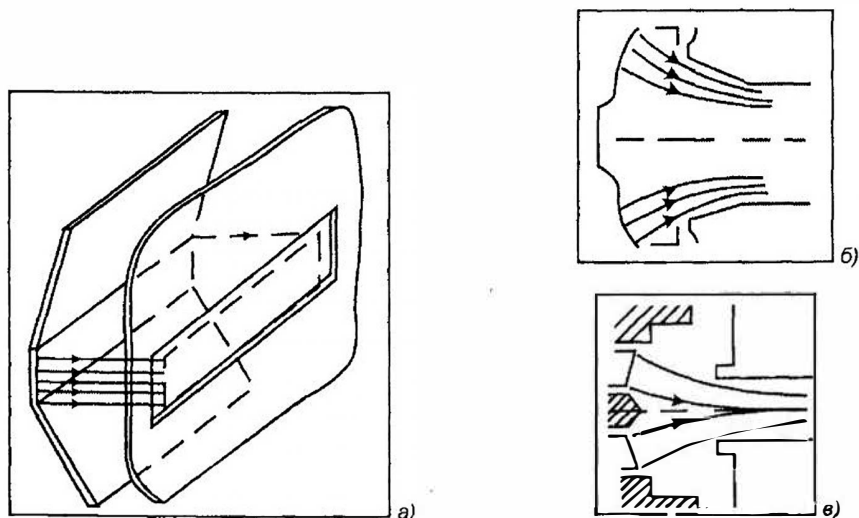


Рис. 3.8. Схемы электронных пушек: а — пушка ленточного пучка;
б — пушка полого пучка; в — многопучковая пушка

В электронном пучке наблюдаются два взаимно противоположных явления. С одной стороны, электрическое поле пространственного заряда создает силу, стремящуюся расширить пучок в соответствии с кулоновским отталкиванием одноименных зарядов. С другой стороны, магнитное поле линейных токов порождает силу Лоренца, стремящуюся сжать электронный поток. Оценки показывают, что действие пространственного заряда связывается при энергиях электронов порядка 10^3 эВ при токах 10^{-4} А. Сжимающее действие магнитного поля проявляется в релятивистских пучках, при скоростях электронов, близких к скорости света, или энергии электронов порядка 1 МэВ.

3.4. Управление потоком электронов полями

3.4.1. Движение электрона в однородных полях

Рассмотрим классическую задачу движения электрона в однородном электрическом поле, у которого величина и направление вектора напряженности во всех точках континуальной среды постоянны (рис. 3.9).

Сила, действующая на электрон в электростатическом поле, определяется только напряженностью поля в данной точке и не зависит от его скорости.

Зададим начальные условия: при $t = 0$, $x_0 = y_0 = z_0 = 0$, V_{x0} , V_{y0} , V_{z0} — заданные величины.

Запишем систему уравнений движения электрона: $F = -q\vec{E} = m\ddot{\vec{a}}$.

Тогда $\ddot{\vec{a}} = -\frac{q}{m}\vec{E} = -\eta\vec{E}$ и составляющие ускорения по осям можно записать в виде:

$$\frac{d^2x}{dt^2} = 0, \quad (3.13)$$

$$\frac{d^2y}{dt^2} = 0, \quad (3.14)$$

$$\frac{d^2z}{dt^2} = -\eta E, \quad (3.15)$$

где $\eta = q/m = 1,76 \times 10^{11}$ к/кг.

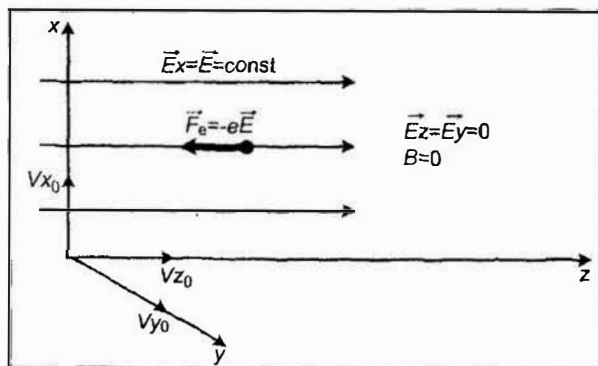


Рис. 3.9. Движение электрона в однородном электрическом поле

Проинтегрировав уравнение (3.13), получим:

$$x = Vx_0 t + C_1.$$

Постоянную интегрирования найдем из начальных условий. При $t = 0$

$$X(0) = Vx_0 \cdot 0 + C_1 = 0,$$

т. е. $C_1 = 0$ и тогда $x(t) = Vx_0 t$.

Другими словами, электрон движется равномерно и прямолинейно с начальной скоростью. В отсутствие сил нет и ускорения.

Аналогично решая уравнение (3.14), имеем:

$$y(t) = Vy_0 \cdot t. \quad (3.16)$$

Решим уравнение (3.15):

$$\frac{d^2 z}{dt^2} = -\eta E.$$

Проинтегрировав, имеем:

$$\frac{dz}{dt} = -\eta E t + C_2,$$

при $t = 0$

$$\left. \frac{dz}{dt} \right|_{t=0} = Vz_0 = -\eta E \cdot 0 + C_2,$$

или $C_2 = Vz_0$ и тогда

$$\frac{dz}{dt} = -\eta E \cdot t + Vz_0. \quad (3.17)$$

Физический смысл решения заключается в том, что с течением времени скорость убывает, если $E > 0$, и возрастает, если $E < 0$.

Проинтегрируем уравнение (3.17) еще раз:

$$z = -\frac{\eta E}{2} \cdot t^2 + Vz_0 t + C_3, \quad (3.18)$$

где C_3 имеет смысл начальной координаты.

Полученное выражение — решение в параметрической форме, где время t — параметр.

Исключим t , воспользовавшись выражением $t = \frac{x}{Vx_0}$:

$$z(x) = -\frac{\eta E}{2Vx_0^2} x^2 + \frac{Vz_0}{Vx_0} x. \quad (3.19)$$

Таким образом, в плоскости (x, z) траектория электрона представляет собой *параболу* (рис. 3.10).

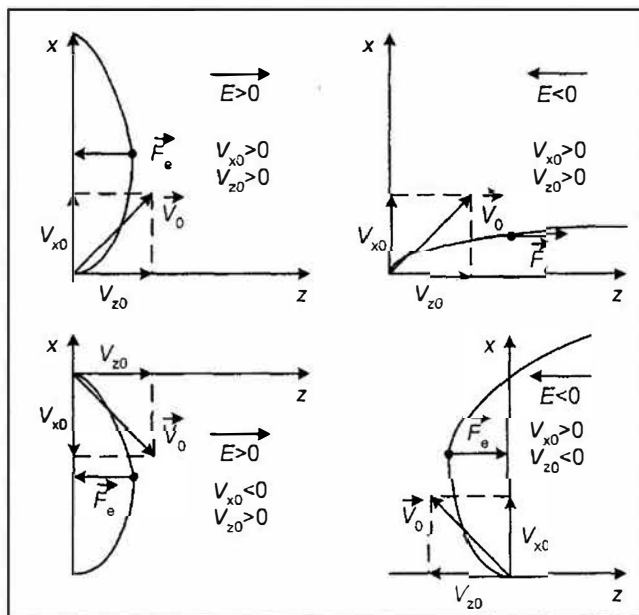


Рис. 3.10. Траектории движения электрона в однородном электрическом поле при различных начальных условиях: знаках E и начальных компонент скорости

3.4.2. Однородное магнитное поле

Под *однородным* магнитным полем будем понимать постоянное по величине и направлению значение \vec{B} во всех точках континуальной среды.

Сила воздействия магнитного поля на электрон определяется силой Лоренца

$$\vec{F}_L = -q \cdot [\vec{V}, \vec{B}].$$

С помощью квадратных скобок обозначено векторное произведение векторов \vec{V} и \vec{B} . Эта сила зависит от индукции в данной точке и от скорости электрона. Она вызывает ускорение a , которое определяется соотношением:

$$a = -\frac{q}{m} [\vec{V}, \vec{B}] = -\eta [\vec{V}, \vec{B}]. \quad (3.20)$$

Значение a легко расписать по координатам в скалярной форме:

$$\frac{d^2 x}{dt^2} = -\eta \cdot (V_y B_z - V_z B_y), \quad (3.21)$$

$$\frac{d^2 y}{dt^2} = -\eta \cdot (V_z B_x - V_x B_z), \quad (3.22)$$

$$\frac{d^2 z}{dt^2} = -\eta \cdot (V_x B_y - V_y B_x). \quad (3.23)$$

или, учитывая смысл уравнений, показанных на рис. 3.11, имеем:

$$\frac{d^2x}{dt^2} = -\eta B \frac{dy}{dt}, \quad (3.24)$$

$$\frac{d^2y}{dt^2} = \eta B \frac{dx}{dt}, \quad (3.25)$$

$$\frac{d^2z}{dt^2} = 0. \quad (3.26)$$

Эти уравнения определяют траекторию движения электрона в однородном магнитном поле.

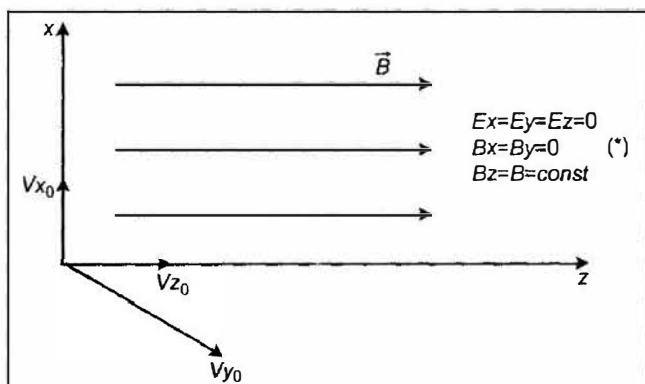


Рис. 3.11. Движение электрона в однородном магнитном поле

Решая уравнение (3.26) путем интегрирования, получим:

$$z = Vz_0 \cdot t,$$

т. е. координата возрастает линейно по времени.

Интегрируя уравнение (3.25), имеем:

$$\frac{dy}{dt} = \eta Bx + C$$

при $t = 0$,

$$\frac{dy}{dt} = Vy_0 = C.$$

Тогда

$$\frac{dy}{dt} = \eta Bx + Vy_0. \quad (3.27)$$

Подставляя (3.27) в (3.24), получим:

$$\frac{d^2x}{dt^2} = -\eta B(\eta Bx + Vy_0) = -(\eta B)^2 x + \eta B Vy_0$$

или

$$\frac{d^2x}{dt^2} + (\eta B)^2 x = -(\eta B) \cdot Vy_0. \quad (3.28)$$

Это неоднородное классическое уравнение колебаний с собственной частотой $\omega = \eta B$.

Решение этого уравнения стандартно и имеет вид:

$$x(t) = R \cos(\omega t + \varphi_0) - \frac{Vy_0}{\omega}. \quad (3.29)$$

Таким образом, электрон совершает колебательное движение по оси X с начальной фазой φ_0 и амплитудой R .

Определим значения R и φ_0 , исходя из начальных условий.

С этой целью продифференцируем уравнение (3.29) и найдем скорость электрона в направлении оси X :

$$\frac{dx}{dt} = -\omega R \sin(\omega t + \varphi_0). \quad (3.30)$$

При $t = 0$ имеем систему:

$$\begin{cases} x(0) = 0 = R \cos \varphi_0 - \frac{Vy_0}{\omega} \\ \left. \frac{dx}{dt} \right|_{t=0} = Vx_0 = -\omega R \sin \varphi_0 \end{cases}$$

Отсюда получаем следующее:

$$R \cos \varphi_0 = \frac{Vy_0}{\omega}, \quad (3.31)$$

$$R \sin \varphi_0 = -\frac{Vx_0}{\omega}. \quad (3.32)$$

Возведя в квадрат обе части этих уравнений и складывая их, получим:

$$R^2 = \left(\frac{Vy_0}{\omega} \right)^2 + \left(\frac{Vx_0}{\omega} \right)^2$$

или

$$R = \frac{1}{\omega} \sqrt{Vx_0^2 + Vy_0^2}. \quad (3.33)$$

Физический смысл радикала заключается в том, что это абсолютная величина начального вектора $|\vec{Vx}_0 + \vec{Vy}_0|$, направленного перпендикулярно силовым линиям магнитного поля.

Обозначим эту величину через V_{\perp} . В этом случае амплитуда колебаний запишется в виде:

$$R = \frac{V_{\perp}}{\eta B}.$$

Разделив (3.32) на (3.31), имеем:

$$\operatorname{tg} \varphi_0 = \frac{Vy_0}{Vx_0}$$

или

$$\varphi_0 = \operatorname{arctg} \left(\frac{Vy_0}{Vx_0} \right).$$

Подставим (3.29) в (3.27):

$$\frac{dy}{dt} = \omega R \cos(\omega t + \varphi_0) - Vy_0 + Vy_0$$

или

$$\frac{dy}{dt} = \omega R \cos(\omega t + \varphi_0),$$

откуда после интегрирования имеем:

$$y = R \sin(\omega t + \varphi_0) + C_1,$$

при $t = 0$

$$y = 0 = R \sin \varphi_0 + C_1.$$

Тогда

$$C_1 = -R \sin \varphi_0 = \frac{Vx_0}{\omega},$$

а

$$y(t) = R \sin(\omega t + \varphi_0) + \frac{Vx_0}{\omega}. \quad (3.34)$$

Другими словами, в направлении оси y электрон в однородном магнитном поле также совершает гармонические колебания, но по сравнению с колебаниями по оси x они сдвинуты по фазе на 90° .

Найдем уравнение траектории движения электрона в однородном магнитном поле. Для этого из уравнений исключим время.

Из (3.29) имеем:

$$\left(x + \frac{Vy_0}{\omega} \right)^2 = R^2 \cos^2(\omega t + \varphi_0),$$

а из (3.32) имеем:

$$\left(y - \frac{Vx_0}{\omega} \right)^2 = R^2 \sin^2(\omega t + \varphi_0).$$

Сложим левые и правые части уравнений, соответственно, и получим:

$$\left(x + \frac{Vy_0}{\omega} \right)^2 + \left(y - \frac{Vx_0}{\omega} \right)^2 = R^2. \quad (3.35)$$

Формула (3.35) является окружностью и интерпретирует собой проекцию траектории электрона на плоскость (x, y) , перпендикулярную магнитному полю.

Центр окружности смещен в точку с координатами (рис. 3.12):

$$X_{\eta} = -\frac{Vy_0}{\omega} = -\frac{Vy_0}{\eta B},$$

$$Y_{\eta} = \frac{Vx_0}{\eta B}.$$

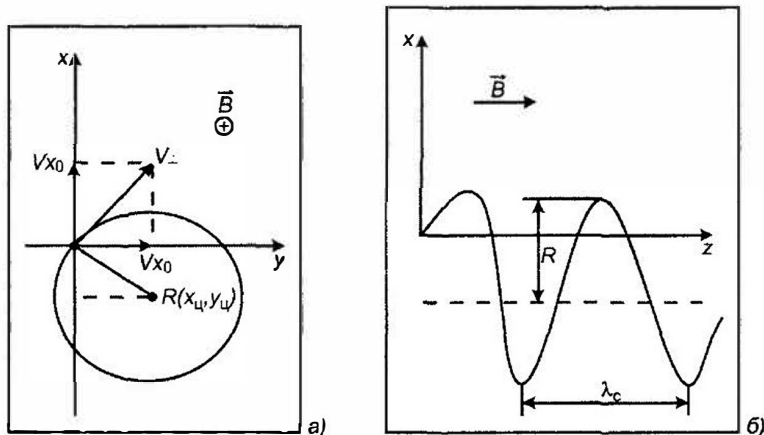


Рис. 3.12. Проекция траектории движения электронов в однородном магнитном поле:
а — на плоскость, перпендикулярную вектору магнитного поля;
б — на плоскость вдоль силовых линий магнитного поля

В плоскости (x, z) проекция траектории электрона представляет собой синусоиду с амплитудой $R = \frac{V_{\perp}}{\eta B}$ (рис. 3.12, б).

Если рассматривать пространственный образ траектории электрона в магнитном однородном поле, то она представляет собой спираль или винтовую линию с шагом или периодом колебаний T . Величина $\lambda_c = T \cdot Vz_0 = Vz_0 \frac{2\pi}{\omega}$ называется *циклотронной длиной волны*.

3.4.3. Движение электрона в скрещенных полях

Под скрещенными полями будем понимать наложенные друг на друга электрические и магнитные поля, перпендикулярные друг другу во всех точках континуального пространства.

К первому типу скрещенных полей отнесем случай, когда оба поля однородны и их векторы взаимно перпендикулярны.

Второй тип скрещенных полей состоит из однородного магнитного поля и электрического поля, обладающего осевой симметрией. Такое электрическое поле образуется в зазоре между коаксиальными цилиндрами.

На рис. 3.13 показана траектория движения электрона в скрещенных электрическом и магнитном полях.

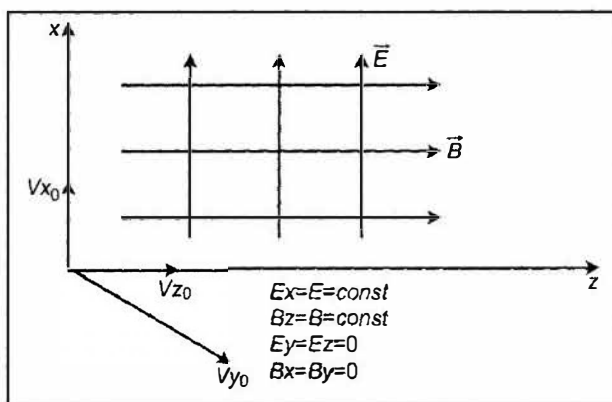


Рис. 3.13. Движение электрона в скрещенных электрическом и магнитном полях

Начальные условия запишем в виде:

$$x(0) = 0 \quad V_x(0) = V_{x0},$$

$$y(0) = 0 \quad V_y(0) = V_{y0},$$

$$z(0) = 0 \quad V_z(0) = V_{z0}.$$

В скрещенных полях на электрон действуют силы F , определяемые соотношением:

$$\vec{F} = m\vec{a} = -q\vec{E} - q[\vec{V}\vec{B}], \quad (3.36)$$

и тогда электрон движется с ускорением:

$$\vec{a} = -\eta E - \eta[\vec{V}\vec{B}]. \quad (3.37)$$

В декартовой системе координат ускорение можно записать:

$$\vec{a} = \frac{d^2x}{dt^2} \vec{i} + \frac{d^2y}{dt^2} \vec{j} + \frac{d^2z}{dt^2} \vec{k}, \quad (3.38)$$

где $\vec{i}, \vec{j}, \vec{k}$ — единичные векторы.

Аналогично:

$$\vec{V} = V_x \vec{i} + V_y \vec{j} + V_z \vec{k} = \frac{dx}{dt} \vec{i} + \frac{dy}{dt} \vec{j} + \frac{dz}{dt} \vec{k}, \quad (3.39)$$

$$\vec{E} = E_x \vec{i} + E_y \vec{j} + E_z \vec{k}, \quad (3.40)$$

$$\vec{B} = B_x \vec{i} + B_y \vec{j} + B_z \vec{k}. \quad (3.41)$$

уравнение (3.37) можно переписать в виде:

$$\begin{aligned}\frac{d^2x}{dt^2} &= -\eta Ex - \eta(VyBz - VzBy), \\ \frac{d^2y}{dt^2} &= -\eta Ey - \eta(VzBx - VxBz), \\ \frac{d^2z}{dt^2} &= -\eta Ez - \eta(VxBz - VyBx),\end{aligned}\quad (3.42)$$

где $-\eta Ez = 0$.

Тогда эта система уравнений (3.42) примет вид:

$$\frac{d^2x}{dt^2} = -\eta E - \omega \frac{dy}{dt}, \quad (3.43)$$

где $\omega = \eta B$ — циклотронная частота.

$$\frac{d^2y}{dt^2} = \omega \frac{dx}{dt}, \quad (3.44)$$

$$\frac{d^2z}{dt^2} = 0. \quad (3.45)$$

Решение уравнения (3.45) запишем в виде:

$$Z = Vz_0 t,$$

а это означает, что вдоль оси z электрон движется прямолинейно и равномерно.

Уравнение (3.44) проинтегрируем (подобно тому, как уже интегрировали до этого):

$$\frac{dy}{dt} = \omega x + Vy_0. \quad (3.46)$$

Подставим (3.46) в (3.43) и получим:

$$\frac{d^2x}{dt^2} + \omega^2 x = -\eta E - \omega Vy_0. \quad (3.47)$$

Перепишем уравнение (3.47) в виде:

$$\frac{d^2x}{dt^2} + \omega^2 x = f_0,$$

где $f_0 = -\eta E - \omega Vy_0$.

Это выражение — известное уравнение колебаний с правой частью, решение которого является функция:

$$X = R \cos(\omega t + \varphi_0) + \frac{f_0}{\omega^2}, \quad (3.48)$$

где R — амплитуда колебаний, а величина $\varphi_0 - \omega^2$ является начальной фазой.

Для рассматриваемого случая решение запишем в виде:

$$X = R \cos(\omega t + \varphi_0) - \frac{\eta E}{\omega^2} - \frac{Vy_0}{\omega}. \quad (3.49)$$

Анализ этого решения показывает, что смещение по оси x имеет постоянную составляющую, которая зависит как от электрического, так и от магнитного полей, а переменная составляющая — это колебания, частота которых зависит от магнитного поля.

Скорость по оси x периодически изменяется

$$\frac{dx}{dt} = -\omega R \sin(\omega t + \varphi_0). \quad (3.50)$$

Решая совместно уравнения (3.49) и (3.50) при $t = 0$, имеем:

$$\begin{cases} 0 = R \cos \varphi_0 - \frac{\eta E}{\omega^2} - \frac{Vy_0}{\omega} \\ Vx_0 = -\omega R \sin \varphi_0 \end{cases}$$

или

$$\begin{cases} R \cos \varphi_0 = \frac{\eta E}{\omega^2} + \frac{Vy_0}{\omega} \\ R \sin \varphi_0 = -\frac{Vx_0}{\omega} \end{cases}. \quad (3.51)$$

Возведя в квадрат и сложив оба уравнения, получим:

$$R^2 = \left(\frac{\eta E}{\omega^2} + \frac{Vy_0}{\omega} \right)^2 + \left(\frac{Vx_0}{\omega} \right)^2$$

или

$$R = \frac{1}{\eta B} \sqrt{Vx_0^2 + \left(Vy_0 + \frac{\eta E}{\omega} \right)^2}.$$

Разделив уравнения (3.51) одно на другое, имеем:

$$\operatorname{tg} \varphi_0 = -\frac{Vx_0}{\omega \left(\frac{\eta E}{\omega^2} + \frac{Vy_0}{\omega} \right)} = \frac{Vx_0}{\frac{\eta E}{\omega} + Vy_0}. \quad (3.52)$$

Таким образом, мы получили амплитуду и начальную фазу колебательного уравнения.

Теперь решим совместно уравнения (3.44) и (3.46):

$$\begin{aligned} \frac{d^2 y}{dt^2} &= \omega \frac{dx}{dt}, \\ \frac{dy}{dt} &= \omega z + Vy_0. \end{aligned}$$

Проинтегрируем уравнение (3.46) и, воспользовавшись соотношением (3.49), получим:

$$\frac{dy}{dt} = \omega R \cos(\omega t + \varphi_0) - \frac{\eta E}{\omega}.$$

Проинтегрировав это уравнение, получим выражение для траектории электрона по оси y :

$$y = R \sin(\omega t + \varphi_0) - \frac{\eta E}{\omega} t + C.$$

Константа C находится из начальных условий:

при $t = 0$

$$0 = R \sin \varphi_0 + C$$

или

$$C = -R \sin \varphi_0 = \frac{Vx_0}{\omega}.$$

Тогда

$$y(t) = R \sin(\omega t + \varphi_0) - \frac{\eta E}{\omega} t + \frac{Vx_0}{\omega}.$$

Выпишем окончательные выражения для траектории электронов по координатам в системе параметрических уравнений:

$$x(t) = R \cos(\omega t + \varphi_0) - \frac{\eta E}{\omega^2} - \frac{Vy_0}{\omega}, \quad (3.53)$$

$$y(t) = R \sin(\omega t + \varphi_0) - \frac{\eta E}{\omega} t + \frac{Vx_0}{\omega}, \quad (3.54)$$

$$z(t) = Vz_0 t. \quad (3.55)$$

Для определения траектории по координатам x и y исключим параметр t .

Итак, при $t = 0$

$$\left(x + \frac{\eta E}{\omega^2} + \frac{Vy_0}{\omega} \right)^2 + \left(y + \frac{\eta E}{\omega} t - \frac{Vx_0}{\omega} \right)^2 = R^2. \quad (3.56)$$

Это выражение — уравнение окружности с радиусом R и координатами центра, которые описываются следующим образом:

$$\begin{cases} x_y = -\frac{\eta E}{\omega^2} + \frac{Vy_0}{\omega} \\ y_y = -\frac{\eta E}{\omega} t + \frac{Vx_0}{\omega} \end{cases} \quad (3.57)$$

Анализ показывает, что траектория движения электронов в плоскости (x, y) представляет собой окружность с центром, которая равномерно смещается по оси y и одновременно перпендикулярна полям \vec{E} и \vec{B} .

Скорость смещения определяется таким образом:

$$V_{cm} = \frac{y_H}{t} = -\frac{\eta E}{\eta B} = -\frac{E}{B}, \quad (3.58)$$

Графически проекция траектории на плоскость (x, y) , которая перпендикулярна магнитному полю, изображена на рис. 3.14.

Эта кривая напоминает *циклоиду* — кривую, описываемую какой-либо точкой колеса, катящегося без скольжения. В нашем случае траектория имеет вид удлиненной циклоиды, радиус которой зависит от напряженности электрического поля и индукции магнитного поля.

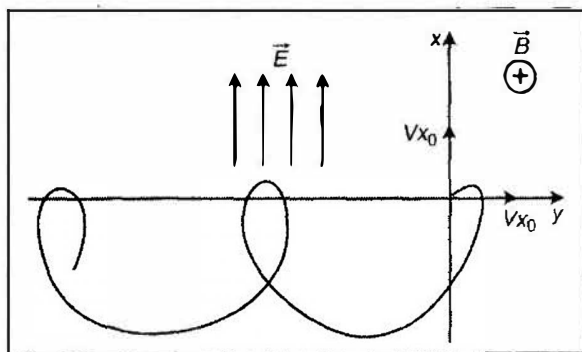


Рис. 3.14. Проекция траектории электрона, движущегося в скрещенных $E \times B$ полях

При смене знака напряженности траектория движения также меняет знак. Параметры циклоиды можно изменять путем варьирования значений V_{x0} и V_{y0} .

Циклоида может превратиться в прямую линию, если в направлении x начальная скорость отсутствует, а начальная скорость в отрицательном направлении по оси y равна скорости сноса.

Другими словами, если сила Лоренца и электростатическая силы равны $\vec{F}_e = -q\vec{E} = = F_L - q[\vec{v}\vec{B}]$, то смещение в направлении x будет отсутствовать.

3.5. Устройства управления электронным пучком

3.5.1. Электростатическая отклоняющая система

Электростатическая отклоняющая система представляет собой две параллельные пластины длиной l , которые расположены друг от друга на расстоянии d .

Если между пластинами приложено напряжение U , то возникает электрическое поле, напряженность которого равна $E = -\frac{U}{d}$ (рис. 3.15). На электрон, влетевший в систему со скоростью $V = \sqrt{2\eta U_a}$, действует сила $F = -qE$, где U_a — напряжение на ускоряющем аноде.

Уравнение движения электрона имеет вид:

$$m \frac{d^2 x}{dt^2} = -qE \quad (3.59)$$

или

$$\frac{d^2 x}{dt^2} = \eta \frac{U}{d}.$$

Проинтегрируем это уравнение:

$$\frac{dx}{dt} = V_x = \frac{\eta U}{d} \cdot t + C,$$

где константу C найдем, положив $t = 0$. Тогда $C = V_{x0} = 0$.

Учитывая, что время пролета пластины $\tau = \frac{l}{V}$, для поперечной скорости электрона имеем следующее выражение:

$$V_x = \frac{\eta U}{d} \cdot \frac{l}{V},$$

и тогда

$$\operatorname{tg} \alpha = \frac{V_x}{V} = \frac{\eta U}{d} \cdot \frac{l}{V^2} = \frac{1}{2} \cdot \frac{l}{d} \cdot \frac{U}{U_a}. \quad (3.60)$$

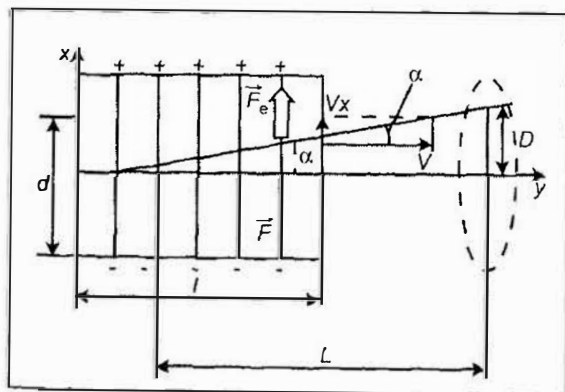


Рис. 3.15. Электростатическая отклоняющая система

Далее электрон попадает в эквипотенциальное пространство и движется по прямой линии. На экране он отклонится от нулевого положения на величину $D = l \cdot \operatorname{tg} \alpha$. Система двух пар пластин образует электростатическую отклоняющую систему конденсаторного типа, которая позволяет получить на экране линейчатый растр (рис. 3.16).

Первая пара, расположенная ближе к экрану, отклоняет луч по оси x , а вторая пара по оси y . При одновременном включении напряжения на обе пластины луч отклоняется одновременно по двум направлениям. Для получения линейчатого раstra на пластины x и y подают пилообразное напряжение. Именно линейная зависимость напряжения от времени

позволяет лучу равномерно двигаться по экрану. Однако при больших углах отклонения эта система искажает электронное пятно.

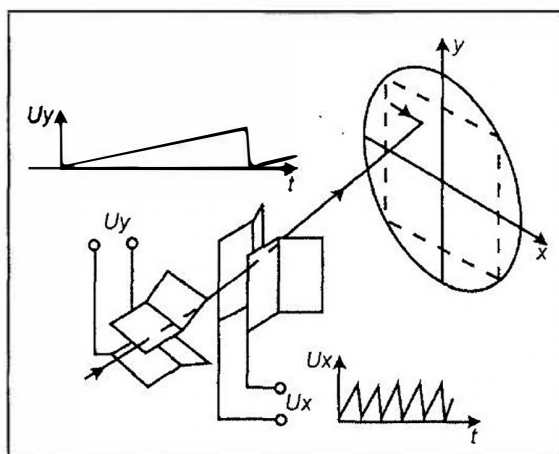


Рис. 3.16. Система отклонения пучка горизонтальными и вертикальными пластинами

Электромагнитные отклоняющие устройства представляют собой две катушки с током, плоскости которых параллельны друг другу (рис. 3.17).

В пространстве между катушками создается магнитное поле, которое можно считать однородным. На влетающий в это пространство электрон, траектория движения которого перпендикулярна магнитно-силовым линиям, действует сила Лоренца:

$$\vec{F}_{\text{Лор}} = -q[\vec{v}\vec{B}]. \quad (3.61)$$

Электромагнитные отклоняющие системы получили широкое распространение в телевизионных устройствах. Используются две пары отклоняющих катушек горизонтального и вертикального отклонения. Такие системы позволяют отклонять электронный луч на 90 и 110 градусов.

Эта сила отклоняет траекторию движения электрона. Изменением направления тока в катушке можно изменить направление вектора магнитной индукции на противоположное и, соответственно, направление вектора силы Лоренца. Меняя силу тока и его направление в магнитной катушке, можно управлять траекториями электронов.

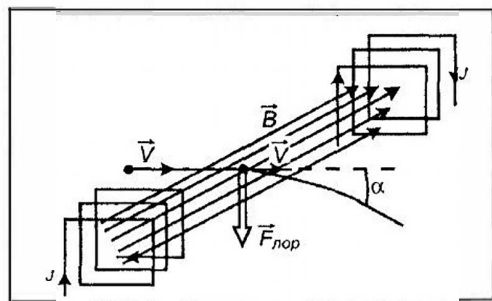


Рис. 3.17. Отклонение электрона поперечным магнитным полем

3.5.2. Управление с помощью электронной оптики

Управление электронными потоками с помощью электронной оптики началось с исследования катодных лучей. Уже тогда было понятно, что характер распространения катодных лучей подобен распространению световых лучей. В качестве *электронной линзы* используется осесимметричное магнитное поле катушки с током. Приведем некоторые аналогии из геометрической и электронной оптики.

- Законы геометрической оптики являются следствием фундаментального принципа Ферма, согласно которому световые лучи распространяются между двумя точками пространства a_1 и a_2 по такому пути S , на прохождение которого затрачивается минимальное время t . Другими словами, существует экстремум функции длины оптического пути

$$\delta \int_{a_1}^{a_2} n ds = 0, \text{ где } n — \text{показатель преломления. Аналогичный закон существует и в электронной оптике.}$$

- В световой оптике справедлив закон Снелля $\frac{\sin \varphi}{\sin \theta} = \frac{n_1}{n_2}$. Падающий и отраженный

лучи, а также нормаль к поверхности лежат в одной плоскости. Этот закон применим и к электронной оптике, причем роль показателя преломления выполняет величина \sqrt{U} . Примечательно, что в оптике справедливо неравенство $1 \leq n \leq 2$, в то время как в электронной оптике величина \sqrt{U} непрерывно меняется от точки к точке и лежит в очень широких пределах.

- Если в световой оптике линза точно очерчена своими границами, то в электронной оптике преломляющие поверхности и границы раздела всегда удовлетворяют уравнению Лапласа:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0, \text{ где } U — \text{потенциал электрического поля в области, не содержащей заряды.}$$

- При прохождении светового луча через границу раздела среды его энергия не меняется, в то время как электронный луч, проходя границу раздела, меняет свои электрические параметры.

Существует три основных типа электронных линз.

- линза, образованная диафрагмой;
- иммерсионная линза;
- одиночная линза.

Сочетание линзы с катодом, например иммерсионный объектив (фокусирующий или рассеивающий), отнесем к катодной части электронных пушек и электронных прожекторов.

Диафрагма может выполнять функцию линзы. В однородном поле эквипотенциали представляют собой параллельные плоскости. Где сильнее поле, там эквипотенциали расположены гуще.

В районе отверстия наблюдается переход от слабого поля к сильному. Функцию линзы выполняет неоднородность электрического поля. Заметим, что силовые линии направле-

ны "от плюса к минусу" или от электрода с более высоким потенциалом к электроду с более низким потенциалом (рис. 3.18).

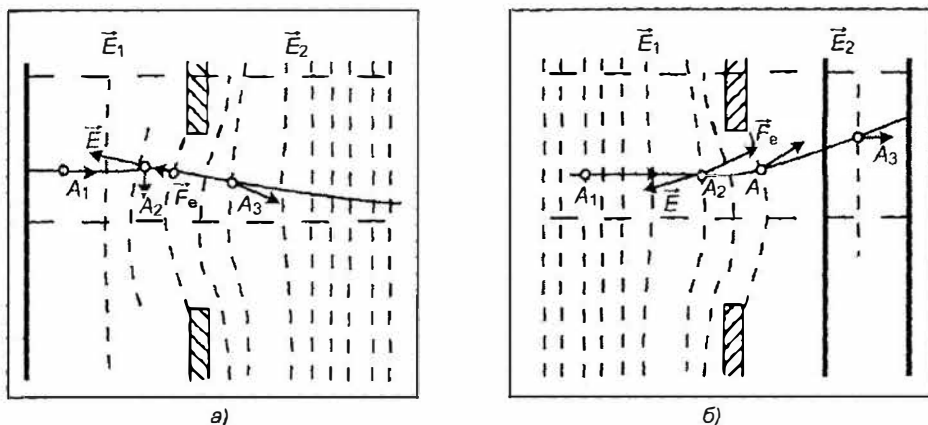


Рис. 3.18. Электронная линза, образованная диафрагмой: а — $|E_2| > |E_1|$; б — $|E_1| > |E_2|$

Вектор напряженности поля направлен по касательной к силовой линии или перпендикулярен к линии эквипотенциала. В точке A_1 на электрон действует только продольная сила потому, что поле в этой точке однородно. Эта сила ускоряет электрон вдоль оси z . В точке A_2 искажение поля уже заметно и вектор силы имеет наклон к оси. Наряду с продольной составляющей силы возникает и радиальная, направляющая электрон в сторону центральной оси. Траектория электрона искривляется и приобретает сходящность. В точке A_3 искривление мало, радиальная сила равна нулю и электрон сохраняет только горизонтальную составляющую скорости.

В этом случае говорят о фокусирующем действии линзы в виде диафрагмы.

В противоположном случае, когда $|E_1| > |E_2|$ анализ траектории электронов свидетельствует о том, что имеет место эффект рассеяния электронов (рис. 3.18, б).

Таким образом, меняя конфигурацию силовых линий поля можно сформировать либо фокусирующую, либо рассеивающую электронный поток электронную линзу.

Иммерсионная линза представляет собой два электрода, между которыми приложена разность потенциалов. На рис. 3.19 представлена линза, образованная двумя трубчатыми электродами, обращенными торцами друг к другу. Картина эквипотенциалей симметрична относительно центральной плоскости.

Положим, что $U_2 > U_1$ относительно катода и траектории движения электронов идут слева направо. Слева эквипотенциальное пространство и силы на электрон не действуют. При подлете к электроду с большим потенциалом он попадает в искривленное электрическое поле.

В точке A за счет прогиба эквипотенциали вектор напряженности электрического поля \vec{E} направлен влево от оси, а сила, действующая на электрон, направлена вправо к оси ($\vec{F} = -\vec{E}e$). Таким образом, возникает радиальная составляющая, направленная к оси и смещающая электрон к центру. Входной участок иммерсионной линзы является фокусирующим.

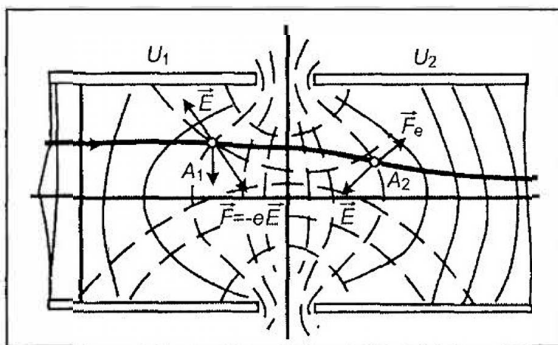


Рис. 3.19. Иммерсионная линза, образованная двумя трубчатыми электродами

На выходном участке электрическое поле направлено к зазору, а сила, действующая на электрон, имеет компоненту, направленную от оси. В этом случае выходной участок линзы является расфокусирующим.

Таким образом, иммерсионная линза скомпонована из фокусирующей и расфокусирующей линз. Так как электрон во входной части линзы находится дальше от оси и, следовательно, находится в более сильном электрическом поле, то он ускоряется этим полем.

Попав в расфокусирующий выходной участок, электрон там находится в течение меньшего времени, и поэтому его итоговая траектория приближается к осевой линии. В целом иммерсионные линзы обладают свойством фокусировки электронов.

Одиночная линза представляет собой совокупность трех электродов, центральный из которых имеет меньший потенциал. Трубчатые электроды с потенциалами $U_1 > U_2$ образуют эквипотенциали (сплошные линии) и силовые линии (штриховые линии), представленные на рис. 3.20.

Наиболее сильное поле сосредоточено в зазорах электродов, а на оси поле более слабое. Возникают замкнутые эквипотенциали, окружающие электрод. Под ним образуются "седлообразные" области.

В точке A_1 сила электрического поля направлена в соответствии с прогибом эквипотенциали наклонно к оси. Сила, действующая на электрон, отклоняет его в сторону центрального электрода.

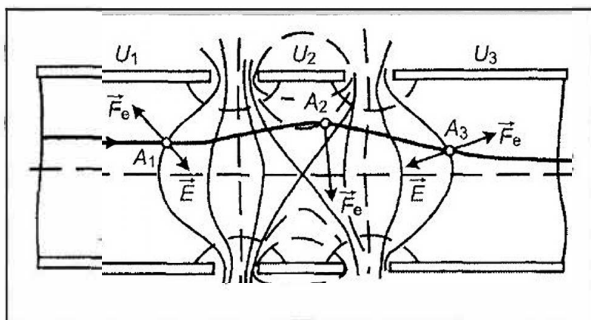


Рис. 3.20. Одиночная линза и траектория электрона на разных участках

В точке A_2 сила, действующая на электрон, направлена в сторону оси. Этот участок является фокусирующим. Скорость электрона под центральным электродом минимальна и он находится на этом участке большее время.

В точке A_3 сила вновь стремится изменить траекторию электрона в сторону от оси.

Такая конструкция линзы позволяет фокусировать электронный поток. Это объясняется тем, что в фокусирующем участке на электрон дольше действует радиальная составляющая силы. Одиночная линза имеет важную особенность: после ее прохождения электрон не меняет своей энергии.

Сформулируем основные свойства электронных линз. По аналогии со световой оптикой в электронной оптике существует понятие *толстых* и *тонких* линз. Как правило, в электронной оптике линзы относятся к типу толстых, схема построения изображения в которых представлена на рис. 3.21.

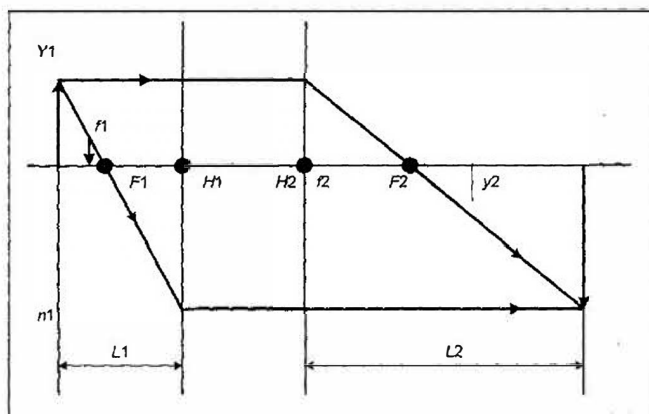


Рис. 3.21. Схема построения изображения в толстой линзе и ее кардинальные точки F_1, F_2, H_1, H_2

Введем соотношения:

$$m = \frac{\varphi_2}{\varphi_1} \text{ — угловое увеличение;}$$

$$M = \frac{y_2}{y_1} \text{ — линейное увеличение;}$$

$$f_1 = F_1 \cdot H_1 \text{ и } f_2 = F_2 \cdot H_2 \text{ — фокусные расстояния.}$$

В этом случае выполняются законы:

$$\frac{f_1}{L_1} + \frac{f_2}{L_2} = 1 \text{ — формула линзы;}$$

$$M \cdot m \frac{n_2}{n_1} = 1,$$

где n_1 и n_2 — показатели преломления.

В электронной оптике роль показателя преломления играет величина \sqrt{U} , где U — напряжение. Как и в оптике, для электронных линз характерны сферическая, хроматическая, геометрическая аберрации, искривления плоскости изображения.

Перечислим основные свойства электронных траекторий.

- Для определения траектории электрона достаточно знать распределение потенциала на оси симметрии.
- Траектории электронов и ионов будут совпадать, если у них одинаковые начальные условия (например, угол наклона).
- Если потенциал всех электродов и всех точек пространства изменить в одинаковое число раз, то траектория электронов не изменится.
- При изменении размеров электродов в n раз размеры траектории изменятся в такое же число раз. Это позволяет осуществлять масштабирование и моделирование.
- Траектории обратимы. Если направить заряженную частицу в обратном направлении, сообщив ей соответствующую скорость, то она пойдет по той же траектории, но в обратном направлении.
- Совокупность исходных точек, лежащих в одной плоскости, можно рассматривать как объект, а совокупность точек, в которых собираются электронные траектории, можно рассматривать как изображение объекта.
- При использовании параксиальных траекторий можно получить увеличенное или уменьшенное изображение без искажений.

3.5.3. Управление с помощью магнитной оптики

В качестве изображающей системы может быть использован *длинный соленоид*, представляющий собой протяженную катушку, длина которой во много раз превышает ее диаметр и по которой течет ток. В центральной части соленоида образуется практически однородное магнитное поле, направленное по оси соленоида.

В таком случае из каждой точки эмиттера будет исходить пучок спиральных линий, обладающих разными радиусами, но одинаковым шагом (рис. 3.22).

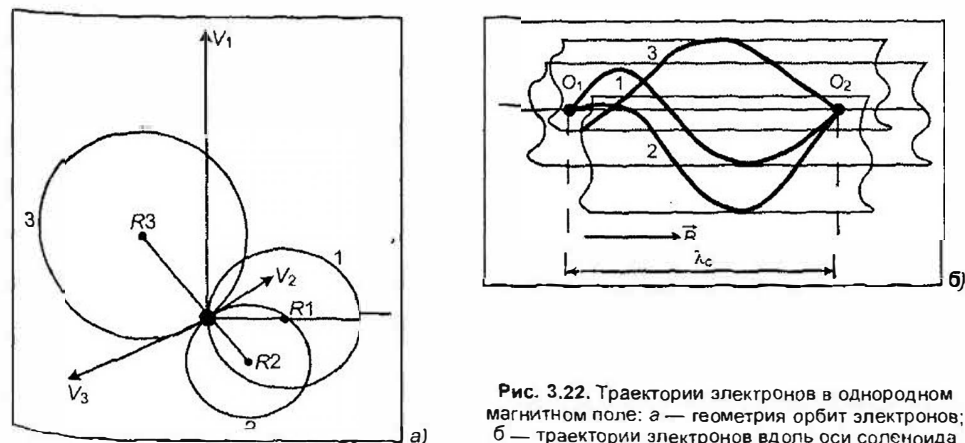


Рис. 3.22. Траектории электронов в однородном магнитном поле: а — геометрия орбит электронов; б — траектории электронов вдоль оси соленоида

Все электроны, вышедшие из точки O_1 , соберутся в точке O_2 . Точка O_2 является электронно-оптическим изображением точки O_1 .

Величина $\lambda_c = V_{\phi} \frac{2\pi}{\omega}$ представляет собой циклотронную длину волны, где $\omega = \eta B$ — циклотронная частота.

Короткая магнитная линза представляет собой катушку (виток) с током. Все силовые линии такого витка с током замкнуты, а наибольшее магнитное поле будет в центре катушки. Короткая магнитная линза даст колоколообразное распределение магнитного поля (рис. 3.23).

Условно примем, что магнитное поле сосредоточено в области, ограниченной плоскостями M_1 и M_2 .

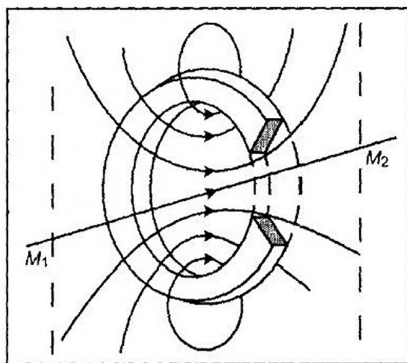


Рис. 3.23. Магнитное поле катушки с током, образующее короткую магнитную линзу

Проведем анализ траектории движения электрона, вышедшего из точки, расположенной на оси, но вне действующего магнитного поля. Из этой точки эмитируются электроны с определенной скоростью и углом наклона вектора скорости (рис. 3.24).

В точке А, вблизи магнитной линзы, на электрон действует сила Лоренца $\vec{F}_{\text{лор}} = -e[\vec{V}, \vec{B}]$,

при этом вектор магнитной индукции \vec{B} направлен по касательной к силовой линии магнитного поля. Согласно правилу левой руки векторное произведение определяет направление силы Лоренца от плоскости вверх. Эта сила $F_{\text{лор}}$ получила название азимутальной силы, которая выводит электрон из первоначальной плоскости, и вследствие этого траектория электрона получает пространственное искривление.

В точке В, расположенной в центре линзы, на электрон действует сила Лоренца. Одновременно на участке АВ под действием азимутальной силы возникает азимутальная скорость, направленная под углом 90° к силовым линиям магнитного поля.

В точке С взаимное расположение векторов \vec{V} и \vec{B} меняется, азимутальная компонента меняет свой знак и тормозится угловое движение электрона. На выходном участке линзы угловая скорость уменьшается и на выходе в точке O_2 становится равной нулю.

Точка O_2 является электронно-оптическим изображением точки O_1 . Такая линза является фокусирующей.

Выпишем без вывода уравнение траектории в меридиональной плоскости в эквипотенциальном пространстве:

$$\frac{d^2 r}{dz^2} + \frac{\eta B_z^2}{8U} r = 0,$$

где r — координата в цилиндрической системе координат.

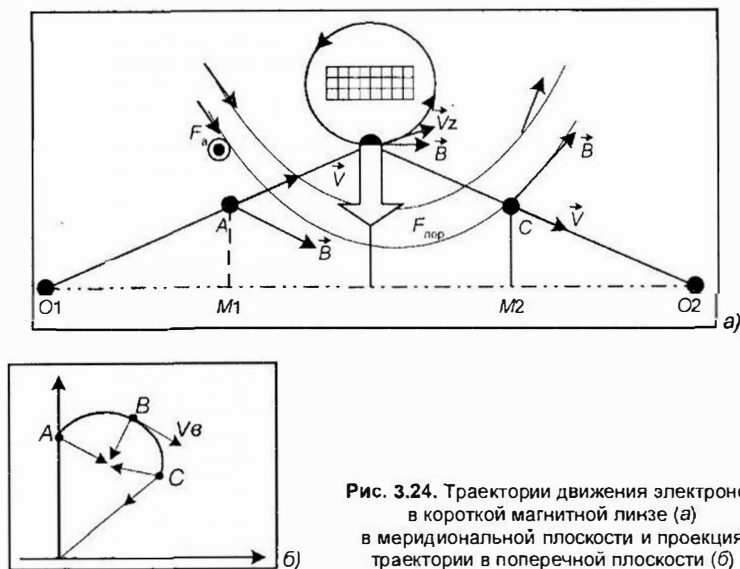


Рис. 3.24. Траектории движения электронов в короткой магнитной линзе (а) в меридиональной плоскости и проекция траектории в поперечной плоскости (б)

Это уравнение описывает траекторию электрона, лежащую в меридиональной плоскости, поворачивающейся вокруг оси с угловой скоростью:

$$\frac{d\theta}{dt} = \frac{\eta B_z}{2}. \quad (3.62)$$

Таким образом, свойства электронных траекторий в магнитном поле:

- изменение полярности поля не изменяет траекторию в меридиональной плоскости;
- траектория электрона является пространственной кривой;
- решение уравнения содержит отношение заряда к массе для данной частицы, следовательно, в одном и том же магнитном поле траектории электронов и ионов различных веществ существенно различаются;
- решения уравнений магнитной и электростатической оптики аналогичны.

Поэтому при помощи параксиальных траекторий в аксиально-симметричном магнитном поле можно получать неискаженное электронно-оптическое изображение.

3.6. Управление скоростью электронов

3.6.1. Резонаторные методы скоростной модуляции электронов

Принцип скоростной модуляции электронов в потоке является краеугольным камнем в теории работы целого класса приборов СВЧ-диапазона — *клистронов*.

В пространство в виде узкого зазора между сетками, проницаемыми электронными потоками, подается высокочастотное электромагнитное поле. Для определенности будем считать, что перед зазором начальные скорости и плотности электронного потока постоянны, а влияние угла пролета в зазоре пренебрежимо мало. Все электроны, вошедшие в зазор, имеют одинаковые скорости

$$V_0 \approx 6 \cdot 10^7 \sqrt{U_0} \text{ см/с,}$$

где U_0 — напряжение на аноде (рис. 3.25, а).

После пролета зазора скорости электронов изменяются в зависимости от фазы высокочастотного напряжения на зазоре.

Для электронов первой группы 1 поле высокочастотного напряжения направлено против постоянного поля и вычитается из него. В этом случае электроны, находящиеся в фазе 1, тормозятся полем (рис. 3.25, б).

Для электронов третьей группы 3 поле высокочастотного напряжения направлено в сторону постоянного поля, и электроны ускоряются им.

На электроны второй группы 2 поле высокочастотного напряжения не влияет (рис. 3.25, б).

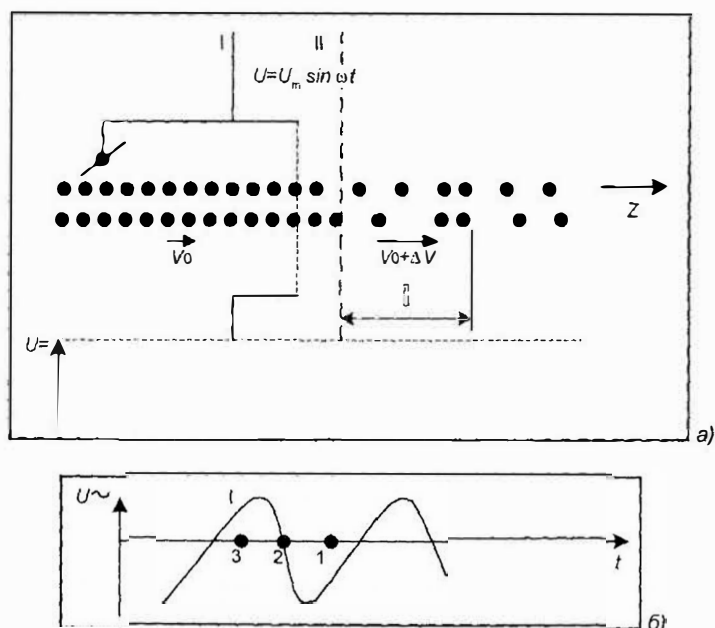


Рис. 3.25. Принцип скоростной модуляции электронного потока: а — группирование электронов; б — три типа электронов в потоке

Таким образом, электроны, влетевшие в зазор в отрицательный полупериод электрического поля, будут им тормозиться, а влетевшие в положительный полупериод, соответственно, ускоряться. Электроны, влетевшие в зазор в момент нулевого поля, будут двигаться равномерно.

Электронный поток, вышедший из зазора, разделяется на три типа электронов — ускоренные $V_0 + \Delta V$, замедленные $V_0 - \Delta V$ и двигающиеся с прежней скоростью V_0 .

Более быстрые электроны будут догонять более медленные и в результате плотность электронного потока станет неравномерной. Электроны группируются и образуют сгусток или пакет зарядов. Модуляция электронов по скоростям сводится к модуляции электронного потока по плотности. Частота сгруппированных электронов будет равна частоте переменного напряжения в зазоре.

В зависимости от фазы высокочастотного напряжения скорости электронов изменяются и, соответственно, изменяется наклон траекторий движения электронов к оси времени.

3.6.2. Нерезонансные устройства скоростной модуляции

Для скоростной модуляции в широкой полосе частот используются не резонансные системы с кратковременным взаимодействием, а системы модуляции с длительным взаимодействием электронного пучка с полем бегущей волны.

Модулирующим устройством служит замедляющая система, представляющая собой волноведущее устройство с замедленной волной, вдоль которой пропускается электронный поток. Конструктивно замедляющая система представляет собой спираль, на вход которой подается СВЧ-сигнал (рис. 3.26).

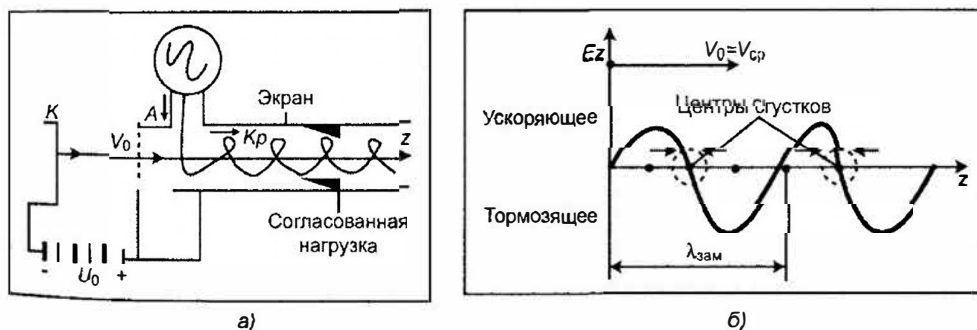


Рис. 3.26. Спиральная замедляющая система (а) и схема скоростной модуляции в процессе длительного взаимодействия электронного потока с бегущей волной (б)

На другом конце замедляющей системы располагается согласованная нагрузка. При выполнении условия синхронизации, заключающегося в равенстве скоростей электронов V_0 и фазовой скорости замедленной волны V_ϕ , электроны 1 и 2 типов остаются неподвижными (рис. 3.26, б).

Это обусловлено тем, что высокочастотное поле в этих точках равно нулю. Электроны типа 3 непрерывно ускоряются продольным полем волны, и их скорость превышает V_0 . Электроны типа 4, находящиеся в тормозящем электрическом поле, наоборот тормозятся, а их скорость становится меньше V_0 . Электронный поток, выходящий из замедляющей системы, становится промодулированным по плотности. Центром сгустков электронов становятся электроны типа 1.

3.7. Детектирование и преобразование энергии электронного потока

3.7.1. Наведение тока при движении электронов в вакууме

Вопрос о связи между движением зарядов в вакууме в анод-катодном пространстве и током, протекающим во внешней цепи диода, является принципиальным и лежит в основе физики работы электровакуумных приборов.

Рассмотрим плоский вакуумный диод, на электроды которого подано напряжение U . При отсутствии зарядов в диоде на обкладках конденсатора имеются поверхностные заряды $+Q_0$ и $-Q_0$. Величина заряда определяется по теории Гаусса: поток вектора электрической индукции \vec{D} , проходящий через замкнутую поверхность S , пропорционален полному свободному заряду Q , заключенному внутри объема V , охватываемого этой поверхностью S .

$$\oint_S \vec{D} ds = 4\pi Q = 4\pi \int_V \rho dv,$$

где ρ — объемная плотность свободного заряда.

Итак, величина поверхностного заряда определяется как:

$$Q = \epsilon_0 E_0 S, \quad (3.63)$$

где $E_0 = \frac{U}{d}$ — напряженность поля внутри диода, S — площадь электродных пластин.

Согласно закону электростатической индукции при внесении в зазор между электродами заряда q наводятся поверхностные заряды $-q_1$ и $-q_2$ (рис. 3.27, а). Величина наведенных зарядов определяется уравнением сохранения зарядов:

$$q - q_1 - q_2 = 0. \quad (3.64)$$

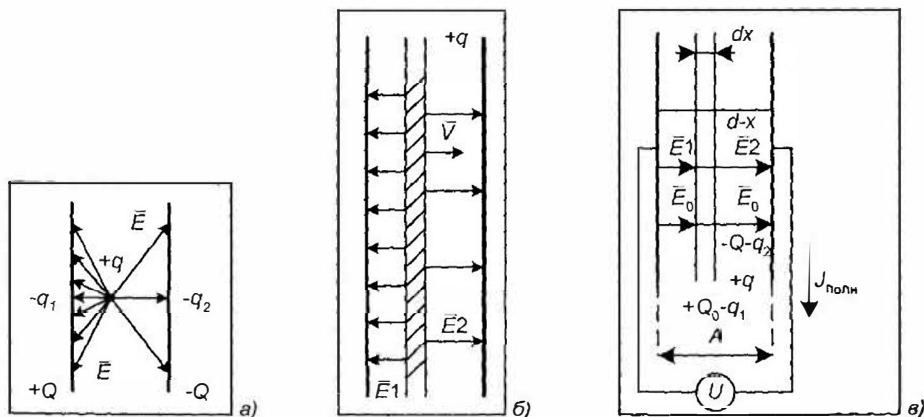


Рис. 3.27. Наведение тока во внешней цепи движущимся зарядом

Если точечный заряд распределить в заряд в виде тонкого слоя, то характер электрического поля внутри зазора изменится и будет иметь вид, изображенный на рис. 3.27, б.

Напряженности поля в зазоре слева и справа определяются соотношением:

$$E_1 = + \frac{Q_0 - q_1}{\epsilon_0 S} \quad E_2 = - \frac{Q_0 - q_2}{\epsilon_0 S} \quad (3.65)$$

Знаки в этих уравнениях обусловлены направлением электрических силовых линий по отношению к электродам.

Поскольку к диоду приложено напряжение U , то $E_1 x + E_2 (d - x) = U = E_0 d$, где x — текущая координата заряженного слоя (рис. 3.27, б).

Подставив выражения для E_1 , E_2 , E_0 через соответствующие заряды, получим:

$$q_2 (d - x) - q_1 x = 0.$$

Решая это уравнение совместно с (3.64), имеем:

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

Отсюда следует, что полные мгновенные заряды на каждой из пластин равны:

$$Q_1 = Q - q \left(1 - \frac{x}{d} \right) \quad Q_2 = -Q_0 - q \frac{x}{d}. \quad (3.66)$$

Полный мгновенный ток, регистрируемый во внешней цепи при движении заряда в зазоре, определяется величиной:

$$i_{\text{мкн}} = + \frac{dQ_1}{dt} = - \frac{dQ_2}{dt}$$

и

$$i_{\text{мкн}} = \frac{dQ_0}{dt} + \frac{q}{d} \cdot \frac{dx}{dt}.$$

Тогда

$$i_{\text{мкн}} = \frac{dQ_0}{dt} + \frac{q}{d} \cdot v,$$

где v — скорость рассматриваемого заряда.

Полный ток включает в себя две компоненты:

$i_{\text{емк}} = \frac{dQ_0}{dt}$ — компонента емкостного тока и

$i_{\text{нав}} = \frac{q}{d} \cdot \frac{dx}{dt}$ — компонента наведенного тока.

Величина наведенного тока определяется выражением:

$$i_{\text{нав}} = \frac{q \cdot v}{d}, \quad (3.67)$$

где v — скорость движения заряда в зазоре диода. Это уравнение получило название *уравнение Рамо*.

В общем случае, когда движутся n зарядов, наведенные токи от всех зарядов суммируются:

$$i_{\text{нап}} = \sum_n q_n \cdot E_m \cdot V_n. \quad (3.68)$$

При прохождении пучка с модулированным по плотности зарядом, например по гармоническому закону $\rho = \rho_0 + \rho_1 \sin \omega t$, наведенный ток будет иметь вид:

$$i_{\text{нап}} = J_0 + J_m \sin \omega t_0 \frac{\sin \frac{\omega d}{2V_0}}{\frac{\omega d}{2V_0}}, \quad (3.69)$$

где V_0 — скорость потока, d — зазор, $\frac{\omega d}{2V_0} = \theta$ — угол пролета электронов через зазор.

Введем обозначение $M = \frac{\sin \frac{\theta}{2}}{\frac{\theta}{2}}$. Тогда

$$i_{\text{нап}} = J_0 + MJ_m \sin \omega t_0, \quad (3.70)$$

где M — коэффициент взаимодействия.

3.7.2. Отбор энергии от электронного потока

Основным назначением электровакуумных СВЧ-приборов является преобразование энергии постоянного тока в высокочастотную энергию поля. Рассмотрим некоторые процессы детектирования электронного потока в приборах энергетического типа.

Наведенный ток, проходя по внешнему сопротивлению R , создает на нем падение напряжения соответствующей полярности (рис. 3.28). Электрод, по направлению к которому движется электрон, оказывается под отрицательным потенциалом.

В момент прохождения электрона создается тормозящее электрическое поле. Кинетическая энергия электрона уменьшается, в результате чего кинетическая энергия электрона при выходе из зазора меньше, чем на входе. Разность между этими значениями кинетической энергии электрона равна энергии, отданной во внешнюю цепь и рассеянной на сопротивлении. Оставшаяся кинетическая энергия электрона рассеивается на других электродах (например, на коллекторе).

Наибольшая величина наведенного тока достигается при коллинеарности векторов \vec{E} и \vec{V} . Поэтому для полного отбора энергии электроны должны двигаться вдоль силовых линий электрического поля.

Таким образом, энергия электронов передается во внешнюю цепь в процессе их движения в продольном тормозящем электрическом поле. Следовательно, возможно разделение функций электродов. Они могут выполнять роль отбора энергии и передачу ее во внешнюю цепь, а другие — роль сбора "отработанных" электронов.

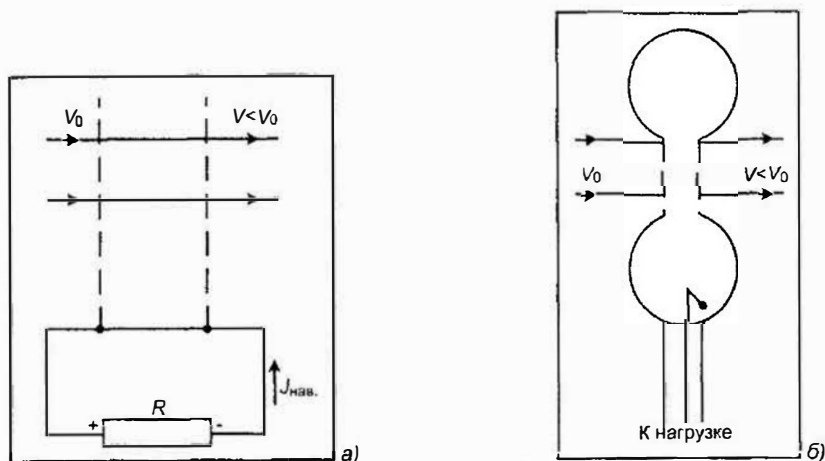


Рис. 3.28. Схема отбора энергии из электронного потока: а — в низкочастотном диапазоне на резистивной нагрузке; б — в высокочастотном диапазоне на согласованной линии

В области сверхвысоких частотных колебаний в качестве активного сопротивления используется полый резонатор, сопротивление которого равно обратной величине активной эквивалентной проводимости.

Поэтому вполне естественно использовать конструкцию, сочетающую зазор, пересекаемый электронным потоком и полый резонатор с петлей связи для вывода энергии. Максимальный отбор энергии от электронного потока достигается в случае равенства резонансной частоты и частоты следования электронов или их сгустков.

Однако отбор энергии с помощью резонаторов связан с частотной селективностью или узкополосностью резонатора. Для расширения рабочей полосы нужно увеличить связь с нагрузкой. С этой целью используются различные замедляющие системы, например, система сеток, либо спиралей, гребенок и т. п. Электронные сгустки должны проходить каждый зазор в одной и той же фазе в момент максимального тормозящего поля.

Это означает, что фазовая скорость волны $V_{ф}$, бегущей по линии, соединяющей зазоры, должна быть равна скорости электронного потока V_0 .

Передача энергии от электронов полю бегущей волны может происходить на большом промежутке замедляющей системы и носит непрерывный характер. Приборы, реализованные на замедляющих системах, называются приборами с длительным взаимодействием электронов с электромагнитным полем.

Проведя системное изучение процессов, связанных с прохождением электронного потока в вакууме, с его взаимодействием с веществом детектора или отбором энергии из потока, можно приступить к изучению приборов вакуумной электроники.

3.7.3. Процессы взаимодействия электронов с веществом детектора

Энергия электронного потока в любом из известных типов электровакуумных приборах обязательно преобразуется либо в информационный сигнал, либо в электромагнитную энергию.

В информационных приборах электронный поток несет, прежде всего, информационные сигналы, которые необходимо расшифровать.

Эти сигналы в дальнейшем можно использовать в усилительных или преобразовательных устройствах, а также можно визуализировать. В любом случае необходимо преобразовать электронный поток в последовательность информационных сигналов.

Энергия электронного потока преобразуется в энергию, удобную для конкретного практического использования.

Рассмотрим варианты преобразования энергии электронного пучка с целью получения информационного сигнала в процессе взаимодействия. Все эти процессы лежат в области взаимодействия излучения с конденсированными средами (рис. 3.29).



Рис. 3.29. Процессы взаимодействия электронов с веществом детектора

Катодолуминесценция — это люминесценция, возникающая при возбуждении вещества потоками электронов, ускоренных во внешнем электронном поле.

Существует несколько механизмов люминесценции. При возбуждении атомов они переходят из основного уровня E_1 на возбужденный уровень E_2 . В результате взаимодействия с окружающими атомами часть энергии теряется, и возбужденные атомы переходят на метастабильный уровень E_2 . Излучательный переход с уровня E_2 позволяет получить излучение, люминесцентная частота которой определяется соотношением $(E_2 - E_1) = h\nu$ (рис. 3.30).

Между зоной проводимости и валентной зоной находятся локальные уровни энергии, связанные атомами примесей или дефектами решетки. Если переходы между такими уровнями сопровождаются люминесценцией, то эти центры называются центрами люминесценции.

Более сложный механизм люминесценции имеет место в кристаллофосфорах, структура которых изображена на рис. 3.31.

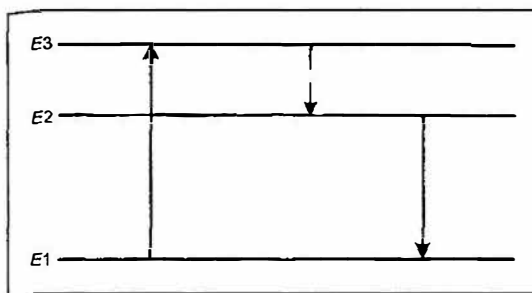


Рис. 3.30. Схема квантовых переходов при элементарных процессах люминесценции

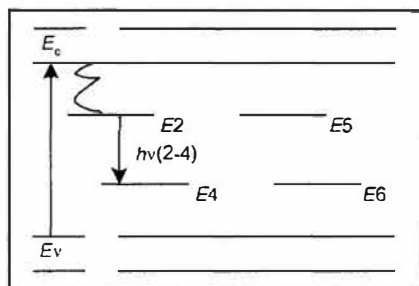


Рис. 3.31. Схема переходов при люминесценции кристаллофосфоров:
 E_v — валентная зона;
 E_c — зона проводимости;
 E_4 — уровни центра люминесценции;
 E_5 — ловушка электронов; E_6 — уровень безызлучательной рекомбинации

Электронный удар переводит электрон из валентной зоны E_v в зону проводимости E_c . Существуют уровни, характерные для каждого вещества, переход между которыми приводит к явлению люминесценции. Вещества, способные к люминесценции, называют люминофорами. Многослойный экран, состоящий из нескольких слоев люминофоров, может возбуждаться электронами разной энергии. Соответственно могут излучаться различные цвета.

Вторичная электронная эмиссия — это явление испускания электронов (вторичных) твердыми телами при их бомбардировке первичными электронами. Электроны могут эмитироваться как "на отражение", так и с тыльной стороны мишени "на прострел". Вторичные электроны имеют непрерывный энергетический спектр (см. разд. 3.2.3).

Внешний фотоэффект представляет собой перераспределение электронов по энергетическим состояниям вследствие поглощения твердым телом квантов электромагнитного излучения с последующей эмиссией электронов (см. разд. 3.2.2).

Внутренний фотоэффект связан с возникновением в твердом теле свободных носителей заряда — электронов и (или) дырок при поглощении квантов электромагнитного излучения. Внутренний фотоэффект стимулирует появление фото-ЭДС, фотопроводимость полупроводника, фотомагнитоэлектрические эффекты.

Электронно-ионная эмиссия — это явление испускания ионов поверхностью твердого тела при ее облучении потоками электронов. При электронных потоках низкой интенсивности происходит десорбция ионов с поверхности. При энергиях бомбардирующих электронов > 26 кэВ и плотности тока > 20 А/см² наблюдается истинно ионная эмиссия (см. разд. 3.2.4, 3.2.8).

Окрашивание ионных кристаллов или кристаллов с ионными (электростатическим) характером связей происходит под воздействием электронного пучка. Кристаллы могут быть галогенидами щелочных и щелочноземельных металлов, фосфатами, карбонатами.

Заряд диэлектриков происходит при облучении электронным пучком поверхности диэлектрика, расположенного на металлической подложке. В зависимости от интенсивности электронного пучка на поверхности диэлектрика формируется потенциальный рельеф глубиной до десятков вольт. Это позволяет производить многократное считывание однократно записанной информации.

Термопластическая деформация происходит на поверхности термопластического слоя под воздействием излучения. На поверхности термопластика образуется микрорельеф деформаций, соответствующий записываемой информации.

Ударная ионизация в полупроводниках происходит при бомбардировке поверхности ускоренными электронами (~ 15 кВ), вследствие чего имеет место явление умножения носителей зарядов. Для эффективного разделения зарядов различного знака полупроводниковая мишень в виде диодной структуры находится под обратным смещением (~ 300 В). Приборы, работающие на основе ударной ионизации, служат для усиления сигнала по мощности.

Механизмы бесконтактного взаимодействия пучка свободных электронов с детектором были рассмотрены ранее. Это явления наведенного тока во внешней цепи и отбор энергии от электронного потока.

Задачи и упражнения

3.1. Вывод формулы Ричардсона — Дешмана. Эффект Шоттки

Плотность тока электронов, испускаемых термоэлектронным катодом, определяется формулой Ричардсона — Дешмана:

$$J = A \cdot T^2 \exp(-\phi/kT),$$

где A — константа, а другие символы имеют свое обычное значение.

Опишите метод, с помощью которого выводится эта формула.

Пусть температура катода равна 2000 К. Анод и катод представляют собой плоскопараллельные пластины, расположенные на расстоянии 1 см друг от друга. Считайте, что пространственный заряд не ограничивает величину тока. Какую разность потенциалов надо приложить между катодом и анодом, чтобы эмиссионный ток увеличился на 10% по сравнению с величиной, даваемой формулой Ричардсона — Дешмана?

Решение

Формула Ричардсона — Дешмана выводится из рассмотрения кинетических энергий электронов, испускаемых с поверхности материала. Заметную вероятность выхода из материала будут иметь только те электроны, энергия которых больше работы выхода и которые движутся в направлении $+x$. Следовательно, плотность тока dJ_x (см. рис. 3.32) определяется выражением

$$dJ_x = qv_x \cdot dN, \quad (3.1.1)$$

где v_x — скорость вылета электронов, а dN — число электронов с энергией, достаточной для преодоления поверхностного потенциального барьера:

$$v_x = p_x/m,$$

$$dN = 2/h^3 \cdot f(E) \cdot dp_x \cdot dp_y \cdot dp_z.$$

Следовательно, подставляя в (3.1.1) выражения для dN и v_x , получим:

$$dJ_x = 2q/m \cdot h^3 \cdot f(E) \cdot dp_x \cdot dp_y \cdot dp_z, \quad (3.1.2)$$

где $f(E) = 1/[\exp(E - E_F)/kT + 1]$ — вероятностный множитель Ферми.

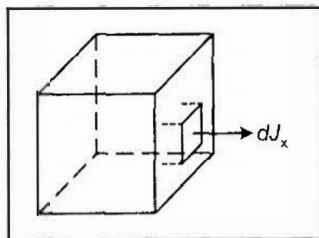


Рис. 3.32. Определение плотности тока эмиссии

Если теперь проинтегрировать выражение (3.1.2) по p_y и p_z в пределах от $-\infty$ до $+\infty$, а по компоненте p_x от p'_x до $+\infty$ (импульс, соответствующий минимальной энергии выхода ϕ), то получим формулу Ричардсона — Дешмана.

Внешнее электростатическое поле E уменьшает работу выхода на величину $\delta\phi = (qE/4\pi\epsilon_0)^{1/2}$ эВ, где ϵ_0 — диэлектрическая проницаемость вакуума.

Это так называемый эффект Шоттки, суть которого заключается в уменьшении работы выхода электронов под действием внешнего электрического поля.

Тогда формула Ричардсона — Дешмана преобразуется к виду:

$$J = AT^2 \exp[-(\phi - \delta\phi)/kT] = AT^2 \exp(-\phi/kT) \exp(\delta\phi/kT) = J_s \cdot \exp(\delta\phi/kT).$$

Пусть $J = 1,1J_s$, тогда можно написать, что $\exp(\delta\phi/kT) = 1,1$. Откуда

$$\delta\phi/kT = \ln(1,1) = 0,0953.$$

Значение величины $\delta\phi = 0,0953 \cdot 1,38 \cdot 10^{-23} \cdot 2000 = 2,63 \cdot 10^{-21}$ Дж, а

$$(2\delta\phi/q)^2 = qV/\pi\epsilon_0 d = (2 \cdot 2,63 \cdot 10^{-21} / q)^2,$$

где d — расстояние между пластинами.

Таким образом, требуемая разность потенциалов для увеличения эмиссионного тока на 10% равна

$$V = 10,8 \cdot 10^{-4} \cdot \pi \cdot 8,85 \cdot 10^{-12} \cdot 10^{-2} / 1,5 \cdot 10^{-19} = 1880 \text{ В}.$$

3.2. Понижение работы выхода вольфрама при осаждении на нем пленки цезия

Диод с вольфрамовым катодом наполнен насыщенными парами цезия. Температура баллона $T_0 = 127^\circ\text{C}$, а температура катода $T_k = 1140 \text{ К}$.

Чему будет равна работа выхода катода в этих условиях?

Решение

До тех пор пока работа выхода катода больше энергии ионизации цезия, атомы цезия будут адсорбироваться на поверхности катода в виде ионов (рис. 3.33). Благодаря этому на поверхности катода будет возникать двойной слой зарядов с плечом a и поверхностной плотностью заряда $\sigma = qN = qN_1\theta$, где $N_1\theta$ — поверхностная плотность зарядов.

В пределах этого двойного слоя создается усредненное поле:

$$E = \sigma/\epsilon_0 = qN_1\theta/\epsilon_0.$$

Благодаря этому в первом приближении работа выхода понижается на величину

$$\delta\phi = EaN_i\theta/\epsilon_0 = pN_i\theta/\epsilon_0,$$

где $p = aq$ — момент диполя. Для ионов цезия, осажденного на вольфрам, величина $a = 1,65 \times 10^{-10}$ м.

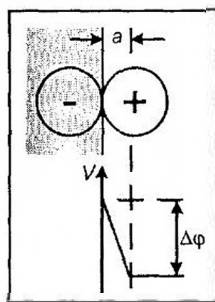


Рис. 3.33. Схема адсорбции атомов цезия

В более высоком приближении следует дополнительно учесть то, что в самом ионе из-за воздействия заряда зеркального изображения происходит смещение электронных оболочек. Иными словами, в самом ионе появляется диполь с некоторым моментом p_0 , который необходимо вычесть из основного дипольного момента p .

Величина p_0 будет пропорциональна локальному полю E_i в окрестности иона, т. е. $p_0 = aE_i$, где E_i определяется силами зеркального изображения по закону Кулона как $E_i = q / 16 \cdot \pi \epsilon_0 a^2$, а величина a — постоянная поляризации, для цезия $a = 2,74 \times 10^{-40}$.

Величина θ может быть найдена из уравнения баланса адсорбции цезия на поверхности вольфрама.

1) Давление паров цезия определяется из соотношения

$$\lg p \approx 11,05 - 1,35 \lg T - 4041 / T = 3,44 \text{ мм рт. ст.}$$

$$p = 2,75 \times 10^{-3} \text{ мм рт. ст.} = 3,66 \times 10^{-1} \text{ н/м}^2.$$

2) Определим энергию адсорбции иона цезия на поверхности вольфрама. При этом будем полагать, что энергия адсорбции будет целиком определяться только силами электростатического притяжения иона к поверхности металла.

Сила зеркального изображения равна

$$F(x) = q / 16 \pi \epsilon_0 x^2,$$

где x — расстояние заряда от поверхности металла.

Энергия адсорбции будет равна работе, которую необходимо затратить, чтобы увести адсорбированный ион на бесконечность:

$$\begin{aligned} \Phi_a &= \int_{\infty}^{\infty} F(x) dx = \frac{e^2}{16 \pi \epsilon_0} \int_{\infty}^{\infty} \frac{dx}{x^3} = \frac{e^2}{16 \pi \epsilon_0 a} = \\ &= \frac{2,56 \cdot 10^{-38}}{16 \cdot 3,14 \cdot 8,85 \cdot 10^{-12} \cdot 1,65 \cdot 10^{-10}} = 3,48 \cdot 10^{-19} \text{ Дж} = 2,18 \text{ эВ.} \end{aligned}$$

Решение

Автоэлектронная эмиссия будет достаточно интенсивной, когда ширина потенциального барьера d станет соизмеримой с длиной волны Де Бройля для электрона $\lambda_e = h / mv$.

Значение энергии электронов и ширину барьера следует взять для электронов, близких к уровню Ферми.

Для вольфрама $W_f = 5,81$ эВ.

$$\lambda_e = \frac{6,62 \cdot 10^{-34}}{9 \cdot 10^{-31} \cdot 6 \cdot 10^5 \sqrt{5,81}} = 5,1 \cdot 10^{-10} \text{ м.}$$

Соизмеримость λ_e и d выразим следующим способом: $d \approx 3\lambda_e = 1,5 \cdot 10^{-9}$ м.

С другой стороны, вследствие понижения потенциального барьера величина $qEd = E_u$.

Для вольфрама эта величина составляет 4,52 эВ.

Отсюда

$$E = E_u / qd = 4,52 / 1,5 \cdot 10^{-9} \approx 3 \cdot 10^9 \text{ В/м.}$$

3.4. Уравнение фотоэффекта Эйнштейна

Фотоэлектрическая работа выхода для калия равна 2,0 эВ. На поверхность калия падает свет с $\lambda = 350$ нм.

Определите:

1. Запирающий потенциал V_s .
2. Кинетическую энергию E_k самых быстрых электронов.
3. Скорости этих электронов.

Вычислите, на сколько изменится запирающий потенциал, если длина волны света уменьшится до 348 нм.

Решение

Энергия фотона определяется из выражения

$$E = hc / \lambda = \frac{1,24}{\lambda} \text{ эВ,}$$

где длина волны λ измеряется в мкм. Следовательно, в случае если $\lambda = 350$ нм, то

$$E = 1,24 / 0,35 = 3,54 \text{ эВ.}$$

Таким образом, энергия эмитируемых электронов E_e представляет собой разность между энергией падающего излучения $E_{\text{фот}}$ и работой выхода материала ϕ , т. е.

$$E_e = E_{\text{фот}} - \phi = 1,54 \text{ эВ.}$$

1. Поэтому запирающий потенциал равен 1,54 эВ.
2. Кинетическая энергия E_k наиболее быстрых электронов также равна 1,54 эВ, или $2,46 \cdot 10^{-19}$ Дж.
3. Скорость наиболее быстрых электронов определяется выражением $\frac{1}{2} m v_{\text{макс}}^2 =$
 $= 2,46 \cdot 10^{-19}$ Дж, откуда получаем $v_{\text{макс}} = 0,74 \times 10^6$ м/с.

уравнение фотоэффекта Эйнштейна имеет вид

$$\frac{1}{2} m v_{\text{макс}}^2 = h\nu - \varphi. \quad (3.4.1)$$

Это уравнение можно переписать следующим образом:

$$qV_s = \frac{hc}{\lambda} - \varphi,$$

где V_s — запирающий потенциал.

Предполагая, что изменение λ мало, запишем это уравнение в дифференциальной форме:

$$\partial V_s = \frac{hc}{e} \frac{\partial \lambda}{\lambda^2}.$$

Поскольку, согласно условиям задачи, $\partial \lambda = 348 - 350 = -2$ нм, а $\lambda = 350$ нм, получаем, что запирающий потенциал уменьшится на величину

$$\partial V_s = \left(\frac{6,63 \cdot 10^{-34} \cdot 3,0 \cdot 10^8}{1,6 \cdot 10^{-19}} \right) \left(\frac{2 \cdot 10^{-9}}{3,5^2 \cdot 10^{-14}} \right) = 20,4 \text{ мВ}.$$

3.5. Фокусное расстояние диафрагмы

Аксиально-симметричная электронно-оптическая система состоит из плоского катода K , вспомогательной ускорительной сетки C , находящейся под потенциалом $U_c = 200$ В, диафрагмы G и анода A . Анод и диафрагма имеют одинаковый потенциал (рис. 3.35).

Определить потенциал диафрагмы U_g (по отношению к катоду), при котором электроны фокусировались бы на аноде. Расстояния между электродами равны $d_1 = 0,5$ см, $d_2 = 0,8$ см.

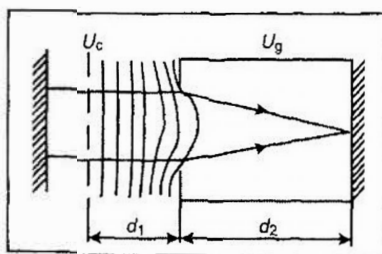


Рис. 3.35. Схема фокусирования диафрагмы

Решение

В аксиально-симметричных электрическом и магнитном поле фокусное расстояние диафрагмы с круглым отверстием имеет вид

$$F = 4U_g / (E_2 - E_1),$$

где U_g — потенциал в центре диафрагмы, E_2 и E_1 напряжения по обе стороны диафрагмы.

В указанной системе по условиям задачи

$$E_1 = U_g - U_c / d_1,$$

$$E_2 = 0, F = d_2.$$

Тогда

$$U_g = U_c d_2 / (4d_1 + d_2) = 200 \cdot 0,8 \cdot 10^{-2} / (4 \cdot 0,5 \cdot 10^{-2} + 0,8 \cdot 10^{-1}) = 57 \text{ В.}$$

3.6. Винтовое движение электрона в магнитном поле

Покажите, что электрон, пересекающий под углом θ прямую магнитную силовую линию, пересечет ее снова после того, как пролетит расстояние, равное $2mv \cos \theta / Bq$, где v — скорость электрона и B — индукция однородного магнитного потока.

Вычислите время, за которое электрон, первоначально находившийся в покое, пролетит расстояние $l = 10$ мм, если разность потенциалов, приложенная на этом расстоянии, равна:

□ постоянному напряжению $V = 100$ В;

□ переменному напряжению, изменяющемуся по синусоидальному закону, с амплитудой 25 В и частотой 3 МГц.

Предположите, что в обоих случаях градиент потенциала является постоянной величиной, а во втором случае электрон находится в покое в тот момент, когда приложенное напряжение равно нулю.

Решение

Пусть v — результирующая скорость электрона, направленная под углом θ к прямым силовым линиям магнитного поля (рис. 3.36).

Таким образом, B и v не перпендикулярны друг другу. Составляющая скорости v_y , перпендикулярная полю, при круговом движении изменяется, а составляющая скорости v_x в направлении поля остается неизменной. Сила F , вызывающая круговое движение электрона, равна

$$F = Bqv \sin \theta = \frac{mv^2 \sin^2 \theta}{r}. \quad (3.6.1)$$

Результирующее движение состоит из кругового движения, перпендикулярного полю, и поступательного движения в направлении силовых линий поля.

Поэтому, как показано на рис. 3.37, результирующее движение происходит по винтовой траектории.

Пусть шаг винтовой линии есть p :

$$p = Tv \cos \theta, \quad (3.6.2)$$

где T — время, за которое совершается один полный оборот при круговом движении.

Если ω — угловая скорость, то

$$v = \omega \cdot r = \frac{rBq}{m},$$

откуда

$$\omega = \frac{Bq}{m} = 2\pi f = \frac{2\pi}{T},$$

где f — циклотронная частота, а период T равен

$$T = \frac{2\pi m}{Bq}, \quad (3.6.3)$$

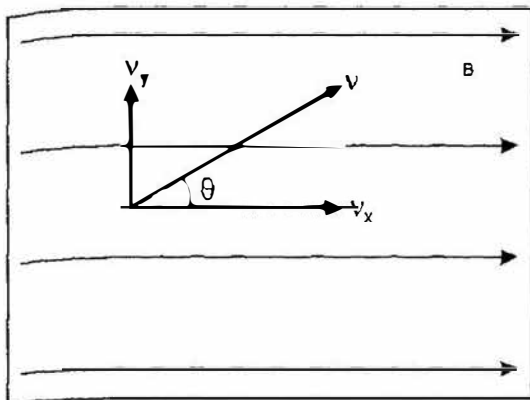


Рис. 3.36. Траектория электрона

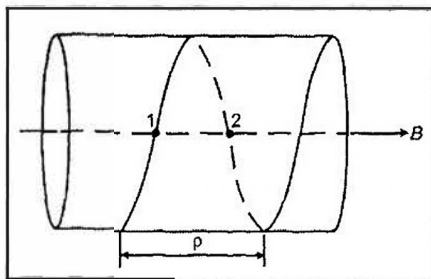


Рис. 3.37. Траектория электрона в магнитном поле

Подставив (3.6.3) в (3.6.2), для шага винтовой линии получим следующее выражение:

$$\rho = \frac{2mv\tau \cos \theta}{Bq} \quad (3.6.4)$$

1. Поскольку электрон начинает свое движение из состояния покоя, его начальная скорость $u = 0$. Поэтому расстояние, пройденное электроном, определяется по формуле

$$s = \frac{1}{2}at^2 \quad (3.6.5)$$

Следовательно, поскольку $a = qE/m$,

$$s = \frac{1}{2}q \frac{E}{m} t^2 \quad (3.6.6)$$

Из этого уравнения находим t и, подставляя в исходные данные ($E = V/l = 10^4$ В/м), получаем

$$t = \left(\frac{2sm}{qE} \right)^{1/2} = \left(\frac{2 \cdot 10 \cdot 10^{-3}}{1,76 \cdot 10^{11} \cdot 10^4} \right)^{1/2} = 3,37 \text{ нс.}$$

2. Мгновенное значение синусоидального напряжения с амплитудой 25 В и частотой $f = 3$ МГц равно $25 \sin \omega \cdot t$ В, где

$$\omega = 2\pi f = 2\pi \cdot 3 \cdot 10^6 \text{ рад/с.}$$

Напряженность электрического поля вычисляется по формуле

$$E = \frac{25 \sin \omega \cdot t}{s} \text{ В/м.} \quad (3.6.7)$$

Подставив выражение (3.6.7) в уравнение (3.6.6), получим

$$t = \left(\frac{2s^2 m}{q 25 \sin \omega \cdot t} \right)^{1/2} \quad (3.6.8)$$

Возведем в квадрат обе части уравнения (3.6.8):

$$t^2 = \frac{2s^2 m}{q25 \sin \omega \cdot t} \quad (3.6.9)$$

Обозначим ωt через θ и подставим эту величину вместо t и ωt в уравнение (3.6.9). В результате получим

$$\theta^2 = \frac{2s^2 m \omega^2}{q25 \sin \theta}.$$

Подстановка исходных данных дает для величины $\theta^2 \sin \theta$ следующее значение:

$$\theta^2 \sin \theta = \frac{2s^2 m \omega^2}{q25} = \frac{2 \cdot 10^{-4} \cdot 4\pi^2 \cdot 9 \cdot 10^{12}}{1,76 \cdot 10^{11} \cdot 25} = 0,0161. \quad (3.6.10)$$

Уравнение (3.6.10) можно решить приближенно, если предположить, что для малых значений θ , выраженных в радианах, $\sin \theta \rightarrow \theta$. Тогда

$$\theta^3 \approx 0,0161 \quad (3.6.11)$$

или

$$\theta \approx 0,252 \text{ рад.}$$

Проверим этот ответ. Вычисляя θ в градусах ($\theta = 0,252 \cdot 180 / \pi = 14,5^\circ$) и подставляя этот результат в аргумент синуса в уравнении (3.6.10), имеем

$$0,252^2 \sin 14,5^\circ = 0,016.$$

Таким образом, значение (3.6.11) найдено верно, и, следовательно, $\theta = 0,252$ рад. Но поскольку $\theta = \omega t$, отсюда находим

$$t = \frac{0,252}{2\pi \cdot 3 \cdot 10^6} = 13,4 \text{ нс.}$$

3.7. Анализ энергий электронов методом тормозящего поля

В плоскую систему электродов с тормозящим электрическим полем вводится поток электронов с начальными энергиями $E_0 \approx 10$ эВ и силой тока $I = 10^{-3}$ А.

Построить вольт-амперную характеристику тока на отрицательный электрод, считая, что:

1. Направления скорости всех электронов нормальны к поверхности электродов.
2. Направления скорости электронов распределены изотропно.
3. Направления скорости электронов распределены в пространстве по закону косинуса.

Решение

Условием попадания электрона на отрицательный электрод будет $E_k \geq qU$, где E_k — кинетическая энергия, соответствующая нормальной компоненте скорости, U — потенциал отрицательного электрода

$$E_k = mv_k^2 / 2 = mv^2 \cdot \cos^2 \theta / 2 = E_0 \cos^2 \theta, \quad (3.7.1)$$

где θ — угол между направлением скорости рассматриваемого электрона и нормалью системы (рис. 3.38).

условием попадания рассматриваемого электрона на отрицательный электрод будет

$$E_0 \cos^2 \theta \geq qU. \quad (3.7.2)$$

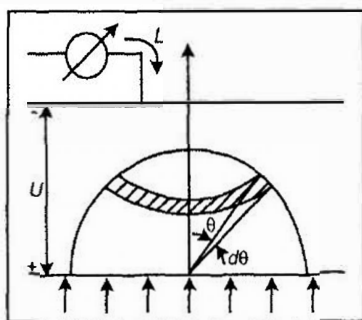


Рис. 3.38. Схема движения электронов в плоской системе электродов

1. Если для всех электронов угол $\theta = 0$, то условием попадания станет

$$E_0 \geq |qU|. \quad (3.7.3)$$

Поэтому

$$i = I \text{ при } |U| < E_0 / q$$

$$i = 0 \text{ при } |U| > E_0 / q$$

2. При изотропном распределении направлений скоростей электронов количество электронов в потоке $dN(\theta, \varphi)$, движущемся в пределах элемента телесного угла $d\Omega(\theta, \varphi)$, пропорционально величине этого угла

$$dN(\theta, \varphi) \cdot a \cdot d\Omega = a \cdot \sin \theta \cdot d\theta \cdot d\varphi,$$

где a — некоторая константа пропорциональности, которая может быть найдена путем нормирования функции $N(\theta, \varphi)$.

Условию попадания электрона на отрицательный электрод будут удовлетворять электроны, у которых $\cos^2 \theta \geq |qU| / E_0$. Поэтому на отрицательный электрод должны попадать только те электроны, у которых угол θ меньше некоторого предельного угла θ_{mp}

$$\cos \theta_{mp} = \sqrt{|qU| / E_0}. \quad (3.7.4)$$

Поток электронов, направления скоростей которых лежат в пределах углов от 0 до некоторого θ , равен

$$N(\theta) = a \int_0^{2\pi} \int_0^\theta \sin \theta d\theta d\varphi = 2\pi \cdot a(1 - \cos \theta).$$

В частном случае, если $\theta = \pi / 2$, то, очевидно, $N(\theta) = N_0$ или всему потоку электронов. Так же можно найти $a = N_0 / 2$.

Величина тока на коллектор в зависимости от потенциала коллектора равна

$$i(U) = qN(U) = qN_0(1 - \cos\theta_{mp}) = I(1 - \sqrt{|qU|/E_0}) \text{ для } |U| < E_0/q$$

$$\text{и } i = 0 \text{ для } |U| \geq E_0/q.$$

При распределении направлений скоростей электронов по закону косинуса

$$dN(\theta, \varphi) = a_1 \cdot \cos\theta \cdot d\Omega,$$

$$N(0) = a_1 \int_0^{2\pi} \int_0^\theta \sin\theta \cos\theta d\theta d\varphi = \pi \cdot a_1 (1 - \cos^2\theta).$$

Нормированием находим величину $a_1 = N_0 / \pi$, тогда

$$N(\theta) = N_0(1 - \cos^2\theta).$$

Подставляя величину θ_{mp} определяем вольт-амперную характеристику (рис. 3.39)

$$i = I(1 - |qU|/E_0) \text{ для } |U| < E_0/q.$$

$$\text{и } i = 0 \text{ для } |U| \geq E_0/q.$$

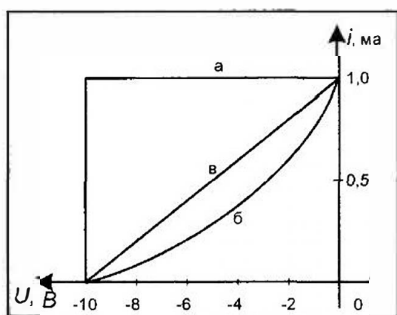


Рис. 3.39. Вольт-амперные характеристики: а — направления скорости всех электронов перпендикулярны к поверхности электродов; б — изотропное распределение скоростей электронов; в — направления скоростей распределены по закону косинуса

Контрольные вопросы

1. Что такое вакуумная электроника?
2. Какие элементы входят в модель прибора вакуумной электроники?
3. Что представляет собой явление электронной эмиссии?
4. Что такое явление термоэлектронной эмиссии? Выпишите уравнение Ричардсона — Дешмана.
5. В чем заключается эффект Шоттки?
6. В чем заключается явление фотоэлектронной эмиссии?
7. Выпишите закон Эйнштейна для фотоэффекта. Что такое красная граница фотоэффекта?
8. Выпишите закон Фаулера для фотоэффекта.
9. Что такое автоэлектронная эмиссия? Выпишите закон Фаулера — Нордгейма.
0. Что такое электронная пушка? Как она устроена?
1. Опишите поведение электрона в скрещенном электрическом и магнитном полях.

12. Какие устройства управления электронным пучком вы знаете?
13. Изложите основные положения управления пучком с помощью электронной оптики.
14. Какие элементы магнитной оптики вы знаете?
15. Каковы физические принципы резонансного метода скоростной модуляции?
16. Каковы физические принципы нерезонансного метода скоростной модуляции?
17. Выпишите закон Рамо для наведенного в цепи тока.
18. Каков принцип отбора энергии из электронного пучка?
19. Перечислите основные физические явления при воздействии электронного пучка на мишень.

Рекомендуемая литература

1. Делямовский И. В. Электронные пучки и электронные пушки. — М.: Советское радио, 1966.
2. Добрецов Л. Н., Гомоюнова М. В. Эмиссионная электроника. — М.: Наука, 1966.
3. Киселев А. Б. Металлооксидные катоды электронных приборов. — М.: МФТИ, 2001.
4. Лебедев И. В. Техника и приборы СВЧ. — М.: Высшая школа, 1972.
5. Левитский С. М. Сборник задач и расчетов по физической электронике. — Изд. Киевского университета, 1964.
6. Линч П., Николаидес А. Задачи по физической электронике. Под ред. проф. Г. В. Скроцкого. — М.: Мир, 1975.
7. Ненакаливаемые катоды. Под ред. М. И. Елинсона. — М.: 1974.
8. Щука А. А. Вакуумная электроника. Учебное пособие. — М.: МИРЭА. 2002.

4. Приборы и устройства вакуумной электроники

4.1. Классификация приборов

Первым прибором вакуумной электроники была лампа А. Н. Лодыгина. Затем был создан электровакуумный диод Дж. А. Флеминга.

Электронные лампы стали основой разработанных систем радиосвязи, радиовещания. Разработка и производство электронно-лучевых приборов способствовало возникновению и развитию телевидения. В 20—50-е годы прошлого века приборостроение устройств вакуумной электроники стало самостоятельной областью знаний, что сопровождалось быстрым ростом промышленности и появлением соответствующих научных учреждений.

Качественно новый этап развития вакуумной электроники наступил при освоении области частот свыше 500 МГц. В 40—50-х годах были разработаны принципы динамического управления электронным потоком. Появились приборы новых классов — клистроны, магнетроны, лампы бегущей волны и другие, работа которых основывалась на взаимодействии электронов с электромагнитными полями. Сформировалась такое направление, как СВЧ-электроника — самостоятельное направление в вакуумной электронике.

В 40—50-е годы на основе приборов вакуумной электроники было создано первое поколение вычислительных машин.

В настоящее время приборы вакуумной электроники применяются в тех областях техники и технологии, где необходимы мощные источники электромагнитного поля, работа в условиях интенсивного радиационного или мощного импульсного электромагнитного облучения. Это, прежде всего, передающее радиовещание, телевидение, дальняя связь, радиолокация и атомная техника.

В зависимости от принципа действия и назначения приборы и устройства вакуумной электроники делятся на электронные лампы, СВЧ-, электронно-лучевые, фотоэлектронные и рентгеновские приборы.

Классификация приборов вакуумной электроники приведена на рис. 4.1.

4.2. Электронные лампы

Электронные лампы — это электровакуумные приборы с термоэлектронным катодом и электростатическим управлением электронным потоком, которые служат для детектирования, генерации и преобразования электрических сигналов.

Диод — двухэлектродный электровакуумный прибор, имеющий *анод* и *катод* (рис. 4.2, а). Электровакуумный диод проводит ток в одном направлении — от катода к аноду и используется для преобразования переменного тока в постоянный (кенотрон).

Эмитированные катодом электроны создают пространственный заряд между катодом и анодом. При положительном потенциале на аноде отрицательный потенциальный барьер

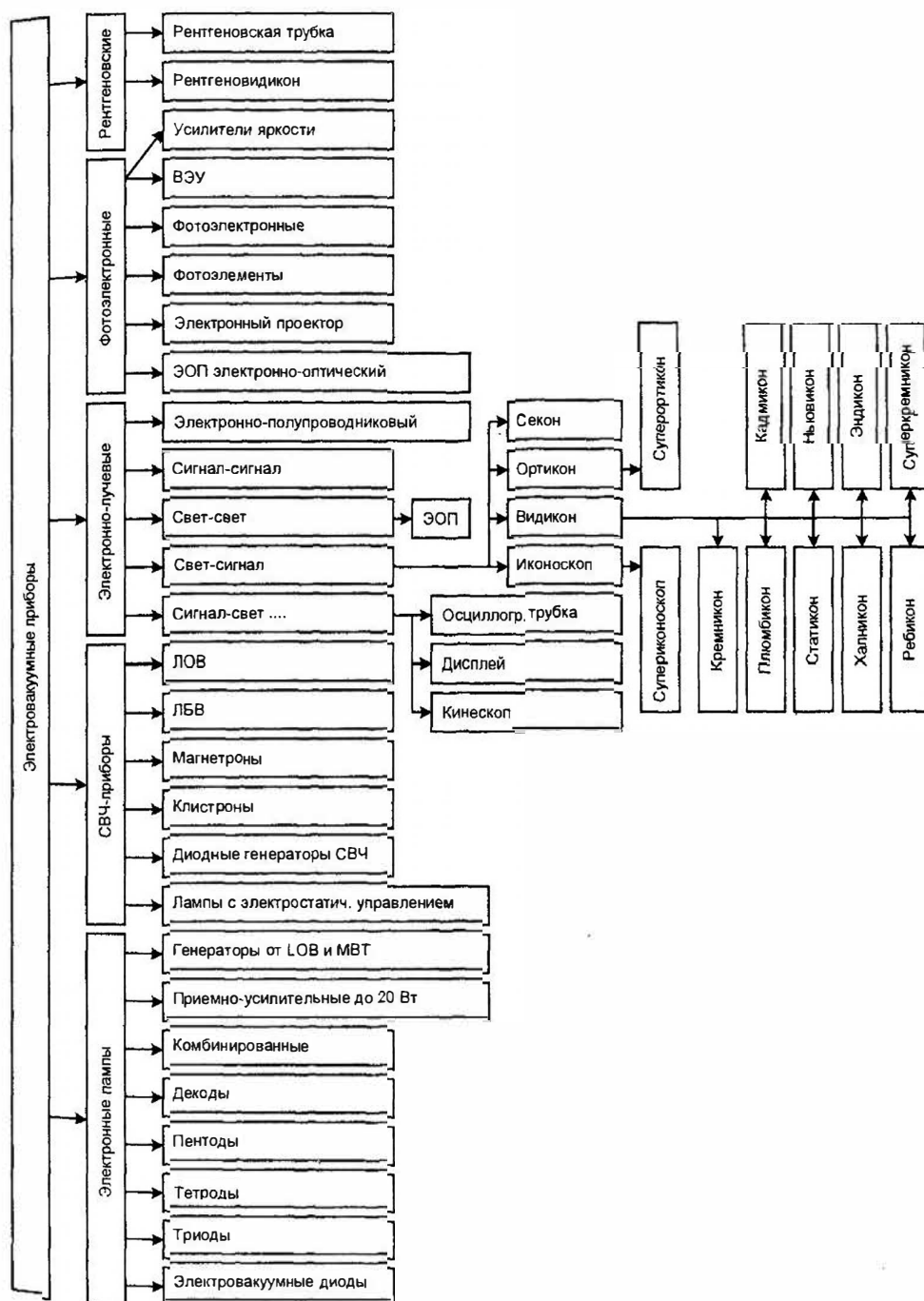


Рис. 4.1. Классификация приборов вакуумной электроники

объемного заряда преодолевают более быстрые электроны, которые создают анодный ток во внешней цепи. Анодный ток в режиме пространственного заряда подчиняется "закону трех вторых":

$$I_a = U_a^{3/2} \quad (4.1)$$

При достаточно большом напряжении на аноде анодный ток становится равным току эмиссии катода. Диод выходит на режим насыщения. Вольтамперная характеристика диода приведена на рис. 4.2, б.

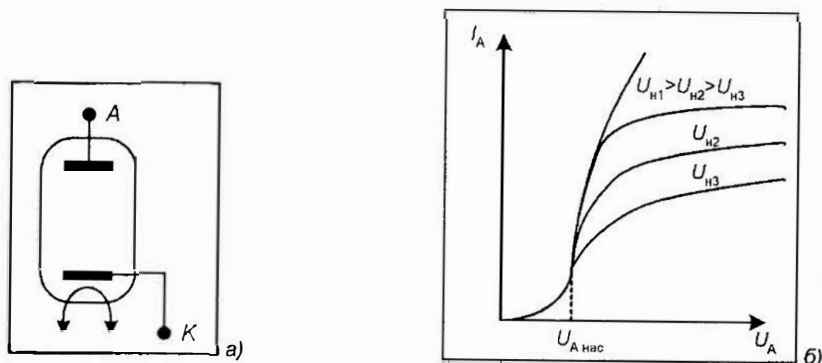


Рис. 4.2. Диоды: условное обозначение (а) и вольтамперная характеристика (б)

Основными параметрами диода являются:

□ крутизна $S = \frac{dI_a}{dU_a}$;

□ внутреннее сопротивление $R_i = 1/S$.

Односторонняя проводимость диода позволяет применять его для выпрямления переменного тока, детектирования электромагнитных колебаний и преобразования частот.

Триод — трехэлектродная вакуумная лампа с управляющей сеткой C между анодом A и катодом K (рис. 4.3, а). Изменяя потенциал сетки U_c можно управлять величиной анодного тока I_a или, другими словами, количеством электронов, проходящих через сетку от катода к аноду. Наличие сетки позволяет применять триоды для усиления и генерации электромагнитных колебаний. Различают сеточные характеристики $I_a = f(U_c)$ при постоянном анодном напряжении и анодные характеристики $I_a = f(U_a)$ при постоянном сеточном напряжении (рис. 4.3, б). Эти характеристики называются статическими.

Основными параметрами триода являются:

□ крутизна $S = \left. \frac{\partial I_a}{\partial U_c} \right|_{U_a = \text{const}}$;

□ внутреннее сопротивление $R_i = \left. \frac{\partial U_a}{\partial I_a} \right|_{U_c = \text{const}}$;

□ коэффициент усиления $\kappa = \frac{\partial U_a}{\partial U_c} \bigg|_{U_c} = \text{const}$;

так, что

$$S \cdot R_a = \kappa. \quad (4.2)$$

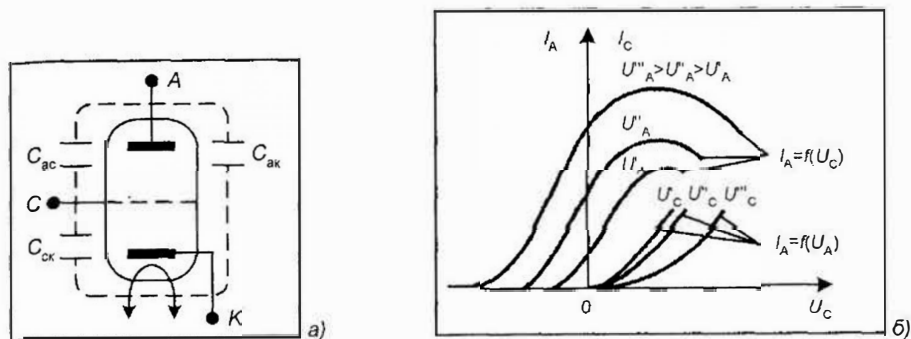


Рис. 4.3. Триод: условное обозначение (а) и вольтамперные характеристики (б)

Если в анодную цепь включить нагрузку R_a , то при изменении I_a одновременно будет меняться U_a . Такой режим и параметры соответствуют динамическому режиму. На рис. 4.4 приведена простейшая схема усилителя на триоде и сеточная динамическая характеристика триода.

При высоких частотах проходная емкость $C_{ак}$ создает большое реактивное сопротивление. Для уменьшения проходной емкости помещают экранирующую сетку $C_{э}$ между анодом и управляющей сеткой. Она экранирует электроды лампы от поля анода, ослабляя влияние U_a на I_a .

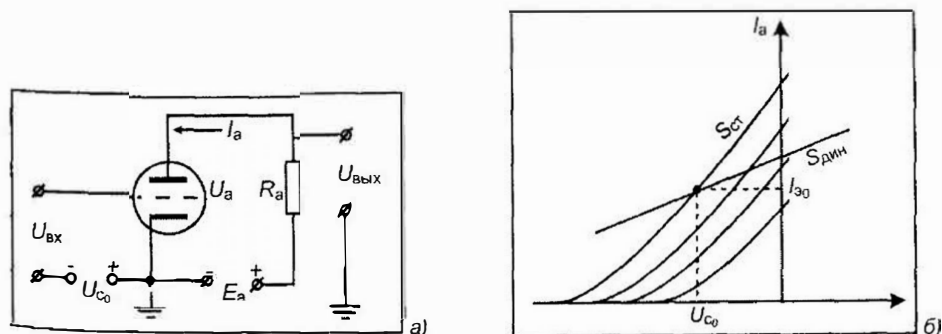


Рис. 4.4. Простейшая схема усилителя, реализованного на триоде (а) и динамическая сеточная характеристика триода (б)

Тетрод — четырехэлектродная лампа с экранирующей сеткой $C_{э}$, позволяющая снизить проходную емкость. При работе тетрода в усилительной ступени на экранирующую сетку подается положительное постоянное напряжение относительно катода $U_{c0} \sim 0,5 U_a$. Электроны, проходящие экранирующую сетку, частично ею перехватываются, формируя ток I_{c0} . При этом

проходная емкость лампы может быть уменьшена на два порядка, а статистический коэффициент усиления увеличивается.

На рис. 4.5 приведена усилительная ступень, реализованная на тетроде, и сравнительная анодная характеристика тетрода и триода. Заметен падающий участок характеристики, обусловленный динаatronным эффектом, связанным с выбиванием вторичных электронов с анода. Падающий участок уменьшает область изменения U_a и схема самовозбуждается. Для снятия динаatronного эффекта вводится еще одна сетка — антидинаatronная C_3 .

Она располагается между экранирующей сеткой и анодом и находится под потенциалом катода, либо небольшим положительным потенциалом.

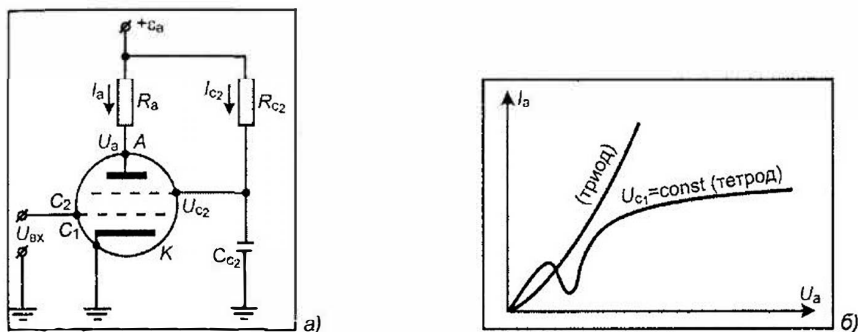


Рис. 4.5. Усилительная ступень, реализованная на тетроде (а), и сравнительная анодная характеристика тетрода и триода (б)

Пятиэлектродная лампа называется *пентодом* (рис. 4.6, а).

Благодаря защитной сетке в пространстве между анодом и экранирующей сеткой создается поле, препятствующее попаданию вторичных электронов на сетку C_2 (рис. 4.6, б). Провал на анодной характеристике пентода ликвидируется. Это позволяет воспрепятствовать проникновению электронов от анода к экранной сетке.

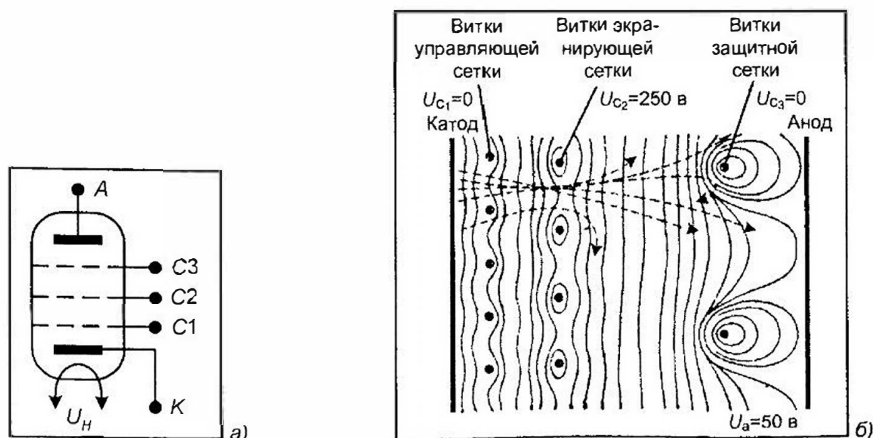


Рис. 4.6. Пентод: условное обозначение (а) и распределение электрического поля между сетками (б)

Пентоды делятся на приемоусилительные и генераторные, которые имеют положительное напряжение на сетке C_3 . В пентодах с двойным управлением сетка C_3 является второй управляющей сеткой, на которую подается отрицательный потенциал. В этом случае между сетками C_2 и C_3 образуется объемный заряд и формируется виртуальный катод. В этой области потенциал равен нулю, электроны тормозятся, создавая подобие катода. К недостаткам пентодов следует отнести большую емкость между третьей сеткой и анодом. Это ограничивает верхний предел диапазона частот при усилении электромагнитных колебаний.

Чтобы ликвидировать этот недостаток, вводят еще одну сетку. Такая конструкция получила название *гексод*. Такая электронная лампа служит в качестве смесителя частот.

Напряжение сигнала обычно подводится к первой управляющей сетке. Переменное напряжение на второй управляющей сетке изменяет токораспределение в лампе. При этом крутизна характеристики анодного тока по первой сетке изменяется с частотой гетеродина, а напряжение на ней меняется с частотой приходящего сигнала. В результате анодный ток представляет собой комбинационные колебания, в частности, колебания промежуточной частоты определяются так:

$$f_{\text{пром}} = f_s - f_c,$$

где f_s — частота гетеродина, f_c — частота сигнала.

Гексоды применяются также для усиления сигналов высокой частоты.

В радиотехнических схемах обычно применяют комбинированные лампы типа "триод-гексод".

Гептод — семиэлектродная электронная лампа, которая служит для преобразования частоты, а также может использоваться как смесительная лампа.

В цепь первой управляющей сетки включается контур гетеродина, в цепь второй сетки — катушка обратной связи. Третья и пятая сетки служат для экранирования, т. к. в гексоде одной экранирующей сетки недостаточно. Четвертая сетка также служит для управления потоком и на нее подается напряжение сигнала.

Гептод можно рассматривать как триод + тетрод.

Для триодной части имеем:

$$S_{\text{тр}} = \frac{\partial J}{\partial U} \frac{\partial_2}{\partial_1}; \quad R_{\text{тр}} = \frac{\partial U}{\partial J} \frac{\partial_2}{\partial_1}; \quad \kappa = \frac{\partial U}{\partial U} \frac{\partial_2}{\partial_1}.$$

Для тетродной части имеем:

$$S_{\text{тет}} = \frac{\partial J_a}{\partial U_a}; \quad R_{\text{тет}} = \frac{\partial U_a}{\partial J_a}; \quad \kappa = \frac{\partial U_a}{\partial U_a} = S_{\text{тет}} R_{\text{тет}}.$$

Лампа с двумя управляющими, двумя экранирующими и сеткой без витков называется *пентагрид* (*pent* — пять, *grid* — сетка).

Октод — восьмиэлектродная электронная лампа с шестью сетками, которая предназначена для работы в частотно-преобразовательных устройствах радиоприемных устройств. По существу — это усовершенствованный гептод. Шестая сетка является антидинактронной, что позволяет увеличить амплитуду выходных сигналов. Особого распространения октоды так и не получили.

Также были созданы электронные лампы с девятью и десятью электродами (*декоды*). Однако практического применения они тоже не нашли. Наибольшее распространение получили двойные диоды — триоды, двойные триоды, триоды — пентоды.

Заметим, что с позиций системного анализа все рассмотренные конструкции соответствуют предложенной модели приборов вакуумной электроники.

Генераторная лампа — электронная лампа, предназначенная для преобразования энергии источника постоянного или переменного тока в энергию высокой частоты до 10 ГГц (рис. 4.7).

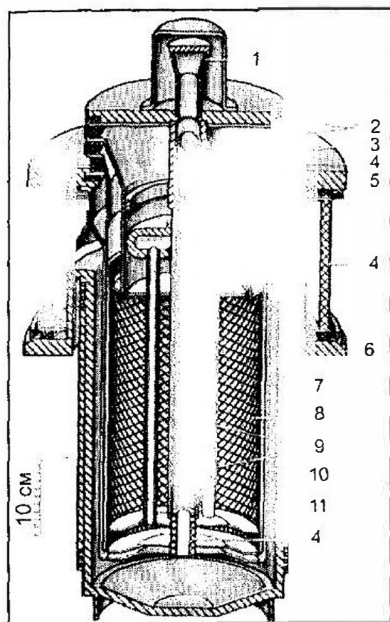
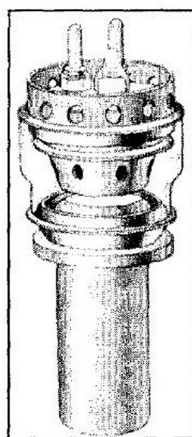
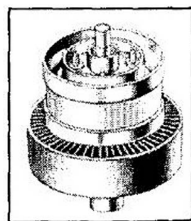


Рис. 4.7. Конструкция сверхмощного коротковолнового тетрода с испарительным охлаждением:

- 1 — штангель для откачки; 2 — вывод накала катода;
- 3 — вывод управляющей сетки;
- 4 — керамический изолятор;
- 5 — вывод экранирующей сетки; 6 — вывод анода;
- 7 — охлаждаемая поверхность; 8 — решетка катодов;
- 9 — подвод накала и крепление катода;
- 10 — управляющая сетка; 11 — экранирующая сетка



а)



б)

Рис. 4.8. Внешний вид современных генераторных ламп:
а — мощный металлостеклянный триод с водяным охлаждением;
б — металлокерамический УКВ-тетрод с воздушным охлаждением

Она может применяться в качестве управляющего элемента генератора или усилителя в радиопередатчиках для радиовещания, телевидения, радиолокации, ускорителях заряженных частиц, медицинской электроники. Различают маломощные (до 25 Вт), средней мощности (до 1 кВт), мощные (до 200 кВт) и сверхмощные лампы (> 200 кВт). На рис. 4.8 показан внешний вид современных генераторных ламп — мощного металлостеклянного триода с водяным охлаждением и металлокерамического УКВ-тетрода с воздушным охлаждением.

Они могут работать в диапазонах коротких (до 30 МГц), ультракоротких (до 300 МГц) или сверхвысокочастотных (до 10 ГГц) диапазонах длин волн.

Приемно-усилительные лампы (ПУЛ) — это электронные лампы, предназначенные для детектирования, преобразования частоты и усиления электрических сигналов на частотах до 300 МГц, а также для генерирования электрических колебаний.

4.3. Электровакуумные микролампы

Электровакуумные микролампы по своим свойствам во многом подобны полевым транзисторам. Вакуумные интегральные триоды называют также вакуумными полевыми транзисторами, а их электроды — катод, сетка, анод или же эмиттер, затвор, коллектор.

В основе работы *вакуумных интегральных схем* (ВИС) лежат те же физические явления, что и в основе работы рассмотренных электровакуумных триодов. ВИС обладают рядом уникальных характеристик, в частности, скорость электронов в них может быть намного больше, чем в любом полупроводнике, их частотные свойства лучше частотных свойств кремниевых интегральных схем и сравнимы со свойствами арсенид-галлиевых ИС. Кроме того, ВИС обладают лучшей радиационной стойкостью. При изготовлении ВИС используется хорошо отработанная технология полупроводниковых ИС.

Одной из основных проблем при создании ВИС является разработка холодных (не подогреваемых) эмиттеров (катодов). В ВИС используется, в основном, электростатическая (автоэлектронная) эмиссия.

Наиболее изученными являются интегральные триоды, которые состоят из холодного эмиттера (катода), управляющего электрода (затвор — аналог управляющей сетки) и коллектора (анода). Примеры устройств вакуумных микротриодов плоской и вертикальной конструкции приведены на рис. 4.9. Катод 1 изготовлен в виде острия из кремния. Между массивной частью катода и управляющим электродом (сеткой) 2, выполненным из металлической пленки, располагается диэлектрический слой SiO_2 . Структура выполнена на плоскости (поверхности) диэлектрической подложки и накрывается диэлектрическим пустотелым колпачком.

На рис. 4.10 показан металлокерамический триод.

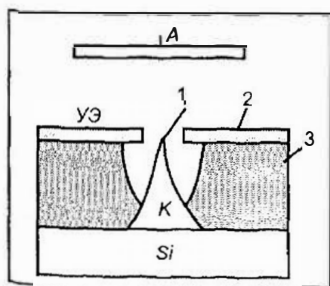


Рис. 4.9. Структура микротриода, выполненного в составе вакуумной интегральной схемы

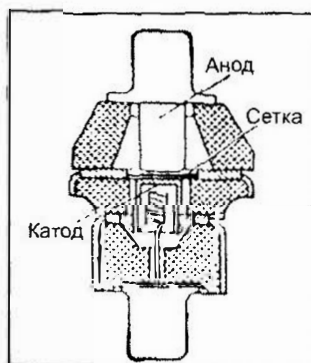


Рис. 4.10. Металлокерамический триод

4.4. СВЧ-приборы

4.4.1. Электронные лампы СВЧ

В области частот более 100 МГц на работу электронных ламп существенно влияют инерция электронов, междуэлектродные емкости и индуктивности вводов-выводов. С этой целью делают плоские электроды с междуэлектродным расстоянием 0,1—0,3 мм, выводы электродов — в виде толстых проводников, а выводы сетки и катода представляют собой кольцеобразные контакты, прижимаемые к контактам резонансных коаксиальных устройств. Вывод анода представляет собой штырь большого диаметра.

Для плоского диодного промежутка время пролета определяется как

$$\tau = \frac{d}{v} = d \sqrt{\frac{m}{eU_0}}. \quad (4.3)$$

Важным фактором является угол пролета θ , равный изменению фазы высокочастотного напряжения на электродах за время пролета τ электронами междуэлектродного расстояния d . Выражение для θ выглядит следующим образом:

$$\theta = 18 \frac{fd}{\sqrt{U_0}} \text{ (град),}$$

где f подставляется в МГц, d — в сантиметрах, напряжение на аноде U_0 — в вольтах.

На низких частотах диод представляет собой активное сопротивление R_a , а на высоких — комплексное сопротивление.

Конструкции электронных ламп СВЧ-диапазона максимально учитывают требования минимального времени пролета электронов, величин междуэлектродных емкостей и индуктивностей вывода.

Различают *генераторные* и *модуляторные лампы*.

Лампы СВЧ-диапазона включают в себя триоды, лучевые тетроды, пентоды. Для генерации метровых волн (> 6 м) служат пентоды и лучевые тетроды. В импульсном режиме для целей радиолокации используются импульсные генераторные лампы. Модуляторные триоды применяются для модуляции в радиотелефонных передатчиках и трансляционных усилителях.

4.4.2. Клистроны

Клистрон (от греческого *κλυζο* — ударять) — электровакуумный СВЧ-прибор, работа которого основана на модуляции по скорости электронного потока электрическими СВЧ-полями резонансных колебательных систем, группированием электронов в сгустки и последующим преобразованием кинетической энергии сгруппированных электронов в энергию СВЧ-колебаний.

Клистроны предназначены для генерации и усиления электромагнитных колебаний.

По способу преобразования энергии источника питания в энергию СВЧ-колебаний клистроны относятся к приборам О-типа или приборам с динамическим управлением электронным потоком.

Идея по плотностной модуляции электронов была впервые предложена советским физиком Д. А. Рожанским в 1932 году, а первые клистроны были созданы американцами В. Ханом, Г. Мектолфом и братьями Р. и З. Варианами в 1938 году.

Принцип группирования электронов в сгустки лежит в основе физики работы клистронов. На рис. 4.11, а приведена схема группирования электронов в пространстве резонатора. Электроны, для которых выполняется соотношение $U \cdot \sin \omega t > 0$, получают дополнительное ускорение. Электроны, для которых справедливо $U \cdot \sin \omega t < 0$, — замедляются. Электроны 1, 2, 3 одновременно возвращаются в резонатор и вместе с электронами от 1 до 3 образуют сгусток электронов.

Следует особо отметить тот факт, что в приборах этого типа носителями информационного сигнала являются уже не электроны, а так называемые динамические неоднородности. В данном случае это сгустки электронов.

Формирование сгустков электронов ухудшают как кулоновские силы, так и конечное время пролета электронов между сетками резонатора.

На рис. 4.11, б представлена конструкция клистрона. Пучок электронов, формируемый электронной пушкой 1, ускоряется полем электрода 2 и пронизывает узкий зазор А между стенками входного тороидального резонатора (группирователь электронов) 4 и движется в трубке дрейфа 6. В трубке дрейфа электрические поля отсутствуют, и в ней происходит преобразование скоростной модуляции потока в модуляцию по плотности. Далее поток из сгустков электронов попадает в выходной резонатор 5.

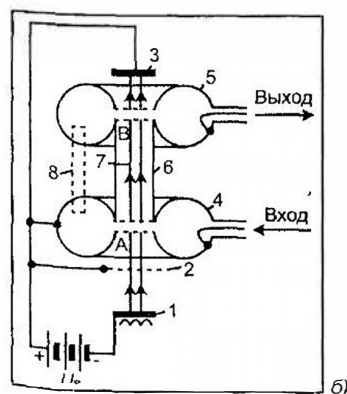
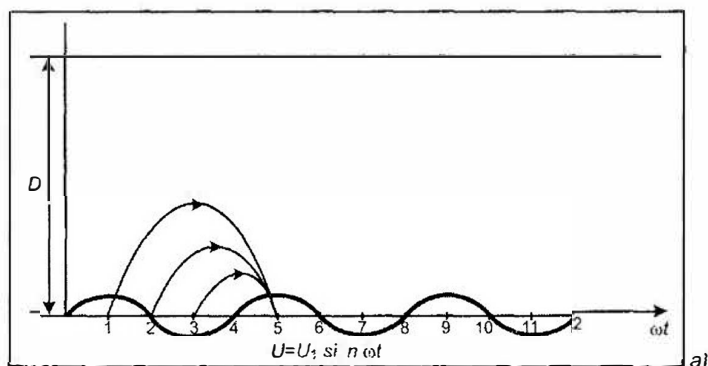


Рис. 4.11. Схема образования сгустков (а) и устройство двухрезонаторного пролетного клистрона (б): 1 — катод; 2 — ускоряющий электрод; 3 — коллектор; 4 — входной резонатор; 5 — выходной резонатор; 6 — трубка дрейфа; 7 — электронный поток; 8 — положительная обратная связь; А — 1-й зазор, В — 2-й зазор

Второй резонатор служит для отбора высокочастотной энергии из потока электронов. Частота поступлений сгустков электронов во второй резонатор равна частоте входного сигнала. Ток наводится на внутренней поверхности стенок второго (выходного) резонатора. Появляющееся между сетками резонатора электрическое поле тормозит электроны. Кинетическая энергия электронов, полученная от источника, ускоряющего напряжения преобразуется в энергию СВЧ-колебаний. Электроны, прошедшие второй зазор, попадают на коллектор и рассеиваются на нем в виде тепла.

Клистрон обладает рядом преимуществ по сравнению с приборами типа лампы СВЧ-диапазона. Среди них:

- ☐ отказ от электростатического управления электронным потоком;
- ☐ использование динамического управления, основанного на скоростной модуляции и группировке электронов;
- ☐ использование принципа наведения тока в выходном зазоре B и разделение функций выходного зазора и коллектора электронов;
- ☐ применение полых резонаторов, отвечающих требованиям СВЧ-диапазона;
- ☐ выделение катода из состава высокочастотной цепи и расположение ускоряющего промежутка перед высокочастотным управляющим зазором.

Клистрон из усилителя может быть преобразован в автогенератор путем введения положительной обратной связи между выходным и входным резонаторами. Одним из способов повышения коэффициента усиления клистрона является их каскадное соединение или создание многорезонаторных конструкций с периодической электростатической фокусировкой электронного пучка.

Клистрон отражательного типа имеет один полевой резонатор, который дважды пронизывается электронным потоком.

Резонатор играет роль группирователя электронов при первом прохождении электронов через зазор и роль выходного контура при повторном прохождении зазора.

Для того чтобы клистрон мог генерировать СВЧ-колебания, сгустки электронов должны проходить через зазор при обратном движении в те моменты, когда в нем имеется тормозящее высокочастотное электрическое поле. С этой целью одновременно регулируются напряжения на ускоряющем электроде и на отражателе.

Основным назначением отражательных клистронов является генерирование СВЧ-колебаний малой мощности. Их преимущество состоит в простоте конструкции и настройки, а также в хороших модуляционных характеристиках. Важным их достоинством является высокая механическая прочность и надежность.

Возврат электронов в зазор резонатора A обеспечивается с помощью отражателя, находящегося под отрицательным потенциалом по отношению к катоду (рис. 4.12).

При использовании отражательного клистрона в качестве усилителя ток в пучке делают меньше пускового и модулированный по скорости пучок электронов в отраженном потоке превращается в модулированный пучок по плотности. При этом в резонаторе возбуждается усиленный сигнал.

Различают отражательные клистроны с внешними и внутренними резонаторами. Широкое распространение получили клистроны с интегральной перестройкой частоты, создание которых стало возможным благодаря разработанной технологии изготовления скользящих электрических контактов в вакууме. Диапазон перестройки составляет 10—15% и КПД свыше 40%. На рис. 4.13 показан широкий спектр отечественных клистронов.

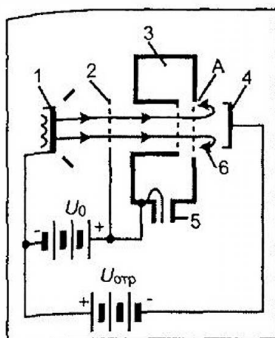


Рис. 4.12. Конструкция отражательного клистрона:
1 — катод; 2 — ускоряющий электрод; 3 — резонатор;
4 — отражатель;
5 — вывод энергии;
6 — электронный поток;
А — зазор

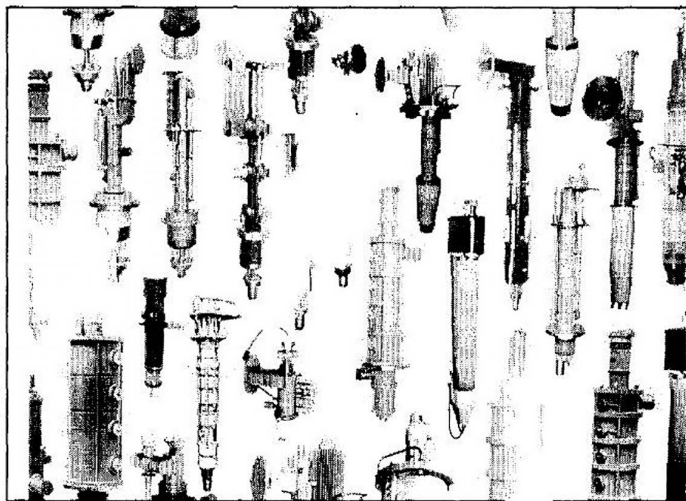


Рис. 4.13. Номенклатура отечественных клистронов

4.4.3. Лампы бегущей волны

Приборы с длительным взаимодействием электронного потока с замедленной электромагнитной волной называются *лампами бегущей волны*.

Из определения следует, что в основе работы усилительных и генераторных ламп бегущей волны лежит длительное взаимодействие электронов с бегущей электромагнитной волной, которая распространяется в нерезонансной колебательной системе. При этом происходит группировка ускоренных электронов и отдача энергии замедленных электронов полю сверхвысокой частоты.

Для длительного взаимодействия электронов с электромагнитным полем необходимо выполнение условия фазового синхронизма, при котором скорость электронов в потоке V_0 совпадает с фазовой скоростью волны V_ϕ .

Различают *лампы прямой волны* (или *лампы бегущей волны*) и *лампы обратной волны*.

Приборы, в которых направление движения электронов совпадает с направлением движения энергии по замедляющей системе, называются *лампами бегущей волны* (ЛБВ). В этих приборах электронный поток взаимодействует с прямой замедленной волной или, как ее еще называют, с положительной пространственной гармоникой.

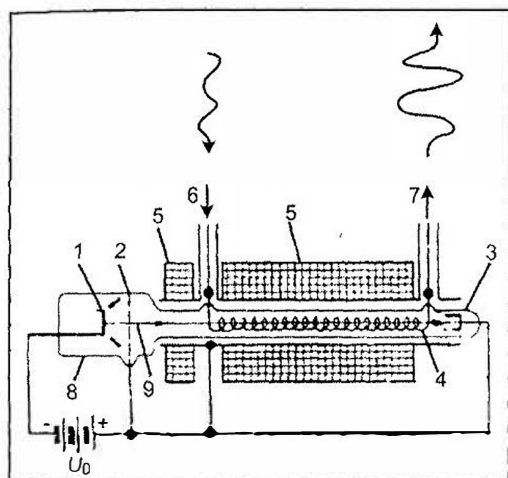
Приборы, в которых используется взаимодействие электронного потока с обратными волнами или отрицательными пространственными гармониками, получили название *ламп обратной волны* (ЛОВ). В ЛОВ электронный поток движется навстречу потоку энергии.

Лампы типа ЛБВ и ЛОВ делятся на две основные группы.

К приборам О-типа относятся лампы с продольным магнитным полем, которое служит лишь для целей фокусировки прямолинейного электронного пучка.

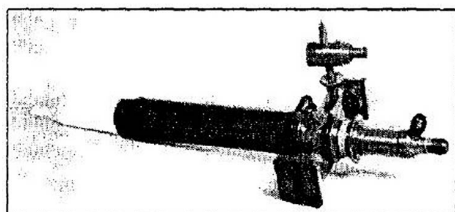
К приборам М-типа относятся все СВЧ-приборы, в которых постоянное магнитное поле является поперечным. В этом случае электроны движутся в скрещенных электрических и магнитных полях.

На рис. 4.14, а представлена ЛБВ О-типа в коаксиальной арматуре. Эмитируемые катодом электроны ускоряются напряжением U_0 , которое обеспечивает требуемое условие синхронизма между электронами и волной, замедленной до скорости $0,1 \cdot c$, где c — скорость света. Движение энергии по замедляющей системе происходит в направлении движения электронов. Фокусировка электронного потока осуществляется с помощью постоянного магнитного поля, созданного соленоидом. Электронные сгустки формируются по мере движения вдоль оси лампы и наводят в спирали ток, а также создают тормозящее высокочастотное поле. Именно это тормозящее поле обеспечивает отбор энергии от электронного потока и усиление входного сигнала.



а)

Рис. 4.14. Схема усилительной лампы ЛБВ (а) и мощная ЛБВ О-типа (б): 1 — катод; 2 — ускоряющий электрод; 3 — коллектор электронов; 4 — спираль замедляющей системы; 5 — коаксиальный соленоид; 6 — вход; 7 — выход СВЧ-энергии; 8 — стеклянная оболочка; 9 — электронный пучок



б)

Главным достоинством ЛБВ является широкая полоса усиливаемых частот. Наряду с усилительными ЛБВ разработаны преобразовательные ЛБВ. В лампах этого типа электронный поток сначала модулируется по скорости сигналом частоты Ω и поступает в пространство взаимодействия замедляющей системы. Через входное устройство подаются колебания СВЧ-диапазона с частотой $\omega \approx 10 \cdot \Omega$. Взаимодействие промодулированного двумя частотами электронного потока с полем бегущей волны приводит к образованию сложной периодической структуры сгустков электронов. Они возбуждают в замедляющей системе колебания с частотами ω и $\omega \pm m \cdot \Omega$, где m — целое число. Обычно параметры замедляющей системы оптимизируют для работы на частоте $\omega + m \cdot \Omega$. В зависимости от режима работы ЛБВ можно разделить на импульсные, непрерывного и квазинепрерывного действия.

По уровню выходной мощности ЛБВ могут быть малой (1—10 Вт), средней (10—100 Вт) и большой мощности (> 100 Вт).

4.4.4. Лампы обратной волны (ЛОВ)

ЛОВ иногда еще называют *каринотронами*. В основе их работы лежит физическое явление длительного взаимодействия электронного потока и обратной волны электромагнитного поля. Схема лампы обратной волны представлена на рис. 4.15, а.

Эмитированные катодом электроны ускоряются напряжением так, чтобы выполнялось условие синхронизма. В условиях синхронизма скорость электронов и фазовая скорость обратной гармоники в ЛОВ совпадают по направлению, а поток энергии направлен в обратную сторону от коллектора электронов к электронной пушке. Поэтому вывод энергии расположен на пушечном конце замедляющей системы. Все прямые гармоники поглощаются согласованной нагрузкой.

Кинетическая энергия электронного потока преобразуется в энергию электромагнитного поля, напряженность которого в волноведущей системе увеличивается от коллекторного конца к пушечному. Усиленные электромагнитные волны, распространяющиеся навстречу электронному потоку, взаимодействуют с электронным потоком по нарастающему эффекту. Электронный поток является как источником энергии, так и звеном, обеспечивающим положительную обратную связь в лампе.

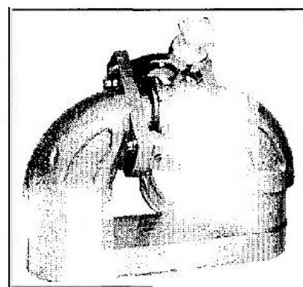
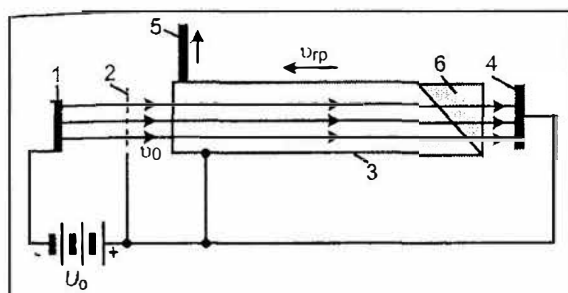


Рис. 4.15. Схема лампы обратной волны (а) и генераторная ЛОВ миллиметрового диапазона (б): 1 — катод; 2 — ускоряющий электрод; 3 — периодическая замедляющая система; 4 — коллектор; 5 — вывод энергии; 6 — согласованная нагрузка

Генерация колебаний в ЛОВ осуществляется за счет наличия внутренней обратной связи, распределенной по длине лампы. Эта связь обусловлена встречным движением энергии и волны в замедляющей системе ЛОВ и обеспечивает возможность плавной перестройки частоты генерации при изменении ускоряющего напряжения.

В ЛОВ М-типа электронный поток отдает электромагнитной волне часть своей потенциальной энергии. Электронный поток формируется в приборах этого типа в скрещенных электромагнитных волнах. Выходящая мощность лежит в пределах 0,1—1 кВт при электронной перестройке частоты. ЛОВ работают в диапазоне частот 0,5—18 ГГц, выходная мощность меняется в пределах 0,1—1 кВт при электронной перестройке частоты до 30%, а КПД лежит в пределах 5—50%.

4.4.5. Магнетроны

Магнетрон (magnetis — магнит) представляет собой коаксиальный цилиндрический диод в магнитном поле, направленном по его оси. Эффект генерирования магнетроном СВЧ-колебаний открыл в 1924 году чехословацкий инженер А. Жачек, а термин магнетрон ввел американский физик А. Халл, который в 1921 году разработал теоретические основы работы магнетрона. В СССР исследования магнетронов оригинальной конструкции были проведены в 1926 году А. А. Слущим и Д. С. Штейнбергом. Первые многорезонаторные магнетроны в СССР были разработаны Н. Ф. Алексеевым и В. М. Маляровым.

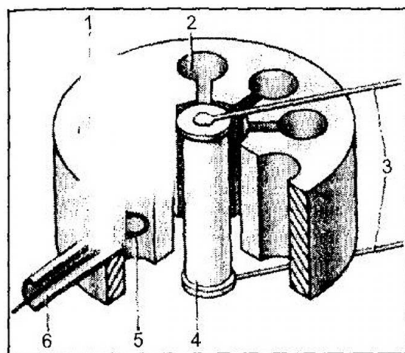


Рис. 4.16. Многорезонаторный магнетрон:
1 — анодный блок с резонаторами
типа "щель—отверстие"; 2 — резонатор;
3 — вывод подогревателя катода; 4 — катод;
5 — петля связи; 6 — выходное устройство
СВЧ-энергии

Магнетрон относится к классу генераторных электровакуумных СВЧ-приборов, в которых формирование электронного потока и его взаимодействие с электромагнитным полем СВЧ-диапазона происходят в пространстве взаимодействия, где электрические и магнитные поля скрещены.

Общий вид магнетрона представлен на рис. 4.16.

Количество резонаторов в диоде магнетронного типа всегда четное. Эмитированные катодом электроны движутся под воздействием скрещенных электрического и магнитного полей.

В пространстве между катодом и анодом происходят процессы, характерные для любого прибора СВЧ-диапазона. Управление электронным потоком, образование сгустков электронов и отдача энергии высокочастотному электрическому полю — все эти процессы происходят в одном пространстве.

Если для простоты рассмотреть пространство между анодом и катодом в виде плоскости (при $R \rightarrow \infty$), то траектория электронов будет описываться циклоидой (рис. 4.17, а).

Благодаря рассеянию полей отдельных резонаторов их колебания жестко связаны друг с другом, а система всех резонаторов представляет собой единую колебательную систему.

Высокочастотные колебания в пространстве взаимодействия магнетрона имеют вид стоячей волны (рис. 4.17, б). Стоячие волны можно интерпретировать как сумму двух бегущих волн, распространяющихся в противоположные стороны. Это достигается путем подбора значений анодного напряжения и напряженности магнитного поля. Электроны уплотняются в сгустки, имеющие вид спиц. Как видно из рис. 4.17, б, внутри спиц траектории электронов имеют сложный характер.

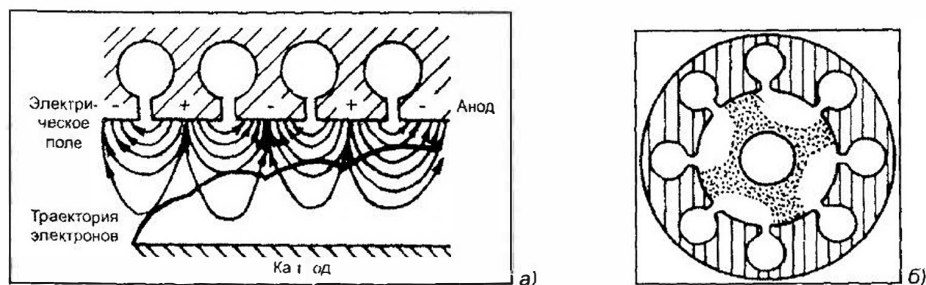


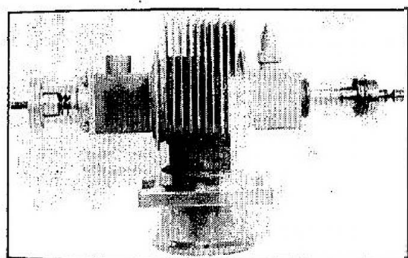
Рис. 4.17. Циклоидальные траектории электронов в плоском магнетроне (а) и образование сгустков электронов (спиц) в пространстве взаимодействия цилиндрического магнетрона (б)

Кинетическая энергия электронов при их движении в спицах определяется скоростью электронов при их движении по циклоидальной траектории. Она максимальна на вершине циклоиды, где скорость определяется как: $V_{\max} = 2 \frac{E}{B}$.

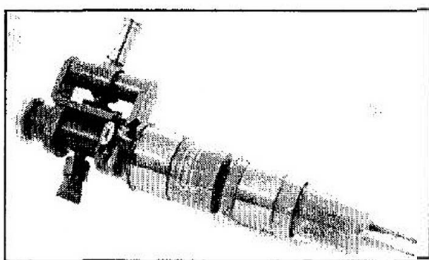
Механизм передачи энергии электронов высокочастотному полю заключается в уменьшении энергии электронов, которая максимальна у катода и близка к нулю у анода. Преобразование энергии электронов в энергию СВЧ-поля продолжается от момента их эмиссии с катода до момента достижения анода. Заметим, что часть электронов в результате сложных траекторий возвращается на катод и способствует увеличению эмиссии за счет эффекта вторичной электронной эмиссии.

Условия синхронизма выполняются при оптимальном соотношении между анодным напряжением и величиной магнитного поля.

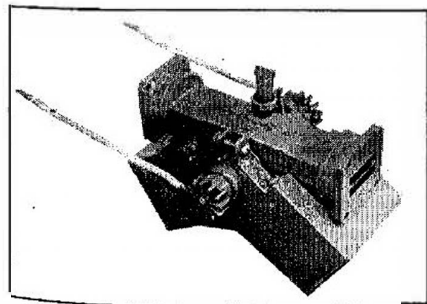
В зависимости от режима работы различают магнетроны импульсного и непрерывного действия. По конструктивному исполнению магнетроны могут быть как перестраиваемые по частоте, так и настроенные на определенную частоту (рис. 4.18).



а)



б)



в)

Рис. 4.18. Некоторые типы магнетронов:
а — мощный импульсный магнетронный генератор сантиметрового диапазона длин волн с перестраиваемой частотой;
б — мощный магнетрон миллиметрового диапазона; в — усилитель СВЧ-диапазона — амплитрон

Различают разновидности магнетронов, которые относят к приборам М-типа:

- *Митрон* — магнетрон, частота генерируемых колебаний которого изменяется в широком диапазоне и пропорциональна анодному напряжению.
- *Амплитрон* (платинотрон) — мощный усилитель обратной волны магнетронного типа с замкнутым электронным потоком.
- *Дематрон* — усилитель прямой волны магнетронного типа с распределенной эмиссией.

- *Инжекцион* — трехэлектродная импульсная модуляторная лампа, в которой для формирования электронного пучка и управления током используется магнетронная пушка.
- *Карматрон* — прибор обратной волны магнетронного типа, в котором используется взаимодействие замкнутого электронного потока с согласованной замедляющей системой.

4.5. Электронно-лучевые приборы

Электронно-лучевыми приборами называется класс электровакуумных приборов, предназначенных для преобразования информации, в которых для этих целей используется поток электронов в форме луча или пучка лучей.

Различают четыре основных типа электронно-лучевых приборов:

- сигнал—свет;
- свет—сигнал;
- сигнал—сигнал;
- свет—свет.

4.5.1. Приборы типа "сигнал—свет"

К приборам первой группы "сигнал—свет" относятся электронно-лучевые приборы, позволяющие преобразовывать электрические сигналы в световые изображения.

В соответствии с предложенной моделью электровакуумного прибора, сформированный устройством управления электронный луч в результате детектирования преобразуется в световой сигнал по различным физическим механизмам, которые и определяют прибор.

Кинескоп — электронно-лучевой прибор, предназначенный для приема электрических сигналов и преобразования их в световое изображение, например, телевизионное.

Кинескоп образован от греческих слов *kinesis* — движение и *skopeo* — смотрю. Различают кинескопы монохромные и цветные.

В основе работы кинескопа лежит явление преобразования энергии электронного пучка в световой сигнал в результате катодолюминесценции. Рассмотрим кинескоп исходя из предложенной модели электровакуумного прибора (рис. 4.19).

Управление электронным пучком осуществляется как с помощью электростатических отклоняющих систем, так и с помощью двух пар отклоняющих магнитных катушек, насаженных на горловину кинескопа.

Яркость свечения в определенной точке экрана определяется мгновенной интенсивностью пучка, управляемого принимаемым телевизионным сигналом. Электропроводность люминофоров достаточна мала.

Оседающие на экран электроны заряжают его отрицательным зарядом, и поток электронного пучка на экран может прекратиться. Однако люминофоры обладают большим коэффициентом вторичной электронной эмиссии. Это явление используется для отвода зарядов путем покрытия внутренней стороны экрана проводящим слоем и подсоединения его к аноду.

В цветном телевидении широко используются масочные кинескопы, экран которых образован узкими полосками люминофоров красного (к), зеленого (з) и синего (с) цвета свече-

ния. Этот масочный тип получил название щелевой теневой маски (рис. 4.20, а). Три электронных прожектора формируют три сходящихся у экрана электронных пучка, каждый из которых возбуждает свечение люминофора только одного цвета. При щелевой маске прожекторы располагаются в одной плоскости, а при использовании маски с круглыми отверстиями — по вершинам равностороннего треугольника. Этот тип называют также дельтаобразным расположением в одной плоскости (рис. 4.20, б).

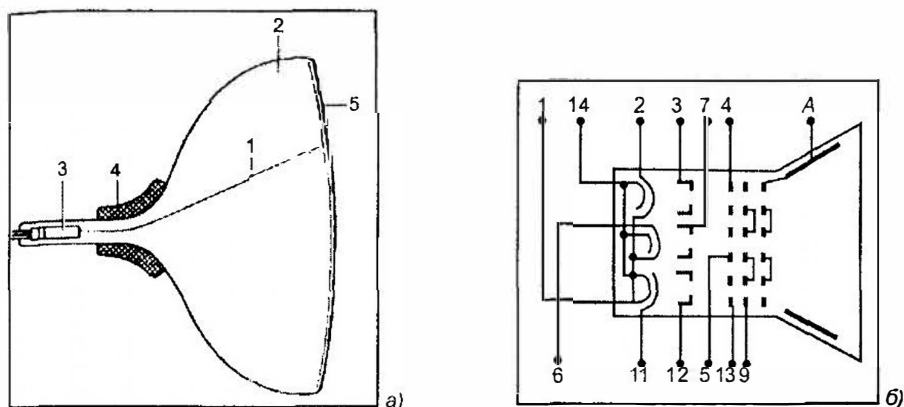


Рис. 4.19. Схема кинескопа (а) и обозначение трехлучевого кинескопа на электрических схемах (б): 1 — пучок электронов; 2 — вакуумная среда, ограниченная баллоном; 3 — электронная пушка — генератор электронов; 4 — устройство управления — отклоняющая система; 5 — детектор — люминесцентный экран

Ощущение всей гаммы цветов обеспечивается на физиологическом уровне путем сложения в глазу излучений сразу трех люминофоров. Интенсивность их возбуждения пропорциональна видеосигналу.

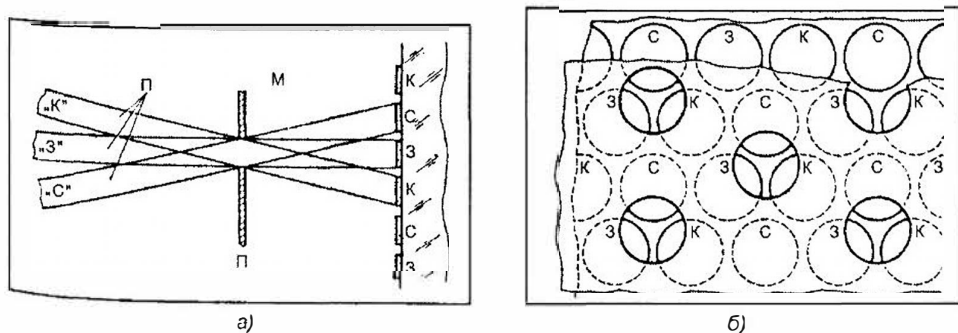


Рис. 4.20. Люминесцентный экран кинескопа: со щелевой маской (а) и маской с круглыми отверстиями (б)

Дисплей — это устройство для визуального отображения информации, как правило, на экране электронно-лучевого прибора. Происходит от английского слова *display* — показывать.

Информация на дисплей поступает непосредственно из компьютера, либо вводится оператором с клавиатуры пульта управления. В состав дисплея входит пульт управления с клавиатурой и микроконтроллер для связи с компьютером.

Осциллографическая трубка — это электронно-лучевой прибор, предназначенный для регистрации хода быстропротекающих процессов в графической форме.

Быстропротекающие процессы должны быть представлены в виде электрических сигналов. Под воздействием периодического пилообразного напряжения, подаваемого на пластины 4а, пучок перемещается с постоянной скоростью в горизонтальном направлении. Измеряемый сигнал подается на пластины 4б и вызывает вертикальное смещение пучка, которое пропорционально мгновенной величине сигнала. Сложение этих перемещений по осям x и y приводит к вычерчиванию осциллограммы физического процесса (рис. 4.21).

Основными характеристиками осциллографической трубки является полоса регистрируемых частот, чувствительность отклоняющей системы и скорость записи сигнала без его искажения.

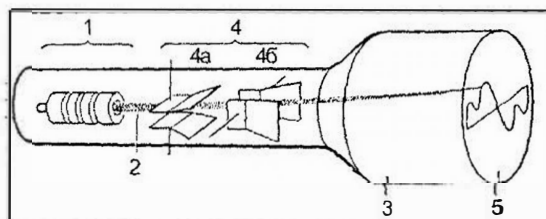


Рис. 4.21. Осциллографическая трубка:

1 — катод; 2 — электронный луч;
3 — вакуумный баллон;
4а, 4б — развертка по вертикали
и горизонтали, 5 — экран

4.5.2. Прибор типа "свет—сигнал"

Приборы типа "свет—сигнал" в основном служат для преобразования изображения в последовательность электрических импульсов с целью их передачи на расстояние.

Иконоскоп (от греческих слов *eikon* — изображение и *scopeo* — смотрю) — первый из приборов этого типа. Принцип его работы основан на накоплении электрического заряда на мозаичной светочувствительной мишени за счет процессов внешнего фотоэффекта.

Светочувствительная мишень иконоскопа представляет собой диэлектрическую подложку, на которую нанесен фотоувствительный слой. На другой стороне подложки нанесен металлический слой, являющийся сигнальной пластиной (рис. 4.22).

Вследствие фотоэффекта проецируемое изображение создает на мозаике потенциальный рельеф, соответствующий распределению освещенности объекта. Электронный луч сканирует поверхность и заряжает все элементы мозаичного экрана в соответствии с ранее накопленным зарядом. Ток в цепи сигнальной пластины становится промодулированным накопленными зарядами.

Дальнейшее развитие иконоскопа получил в *супериконоскопе*. В нем фотоувствительная мозаика заменена сплошным фотокатодом, чувствительность которого на порядок выше, и сплошной мишенью, которые разделены в пространстве.

Накопление заряда и образование потенциального рельефа происходит за счет вторичной электронной эмиссии при бомбардировке мишени фотоэлектронами в процессе переноса электронного изображения.

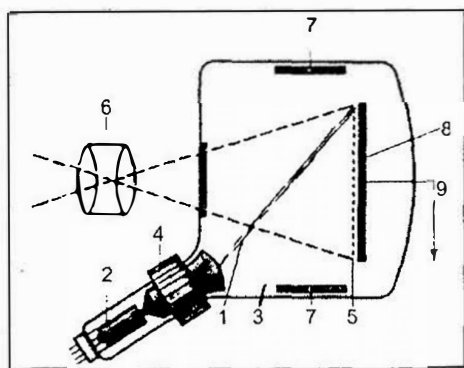


Рис. 4.22. Схема иконоскопа:
1 — электронный луч; 2 — катод —
электронный прожектор; 3 — баллон
вакуумный; 4 — отклоняющая система;
5 — мозаичная фотомишень;
6 — объектив; 7 — коллектор фото- и
вторичных электронов; 8 — сигнальная
пластина; 9 — видеосигнал

Приборы с накоплением заряда

Видикон (от латинского *video* — вижу и греческого *eikon* — изображение) — это телевизионный передающий электронно-лучевой прибор с накоплением заряда, действие которого основано на внутреннем фотоэффекте.

Изображение, которое необходимо передать по телевизионному каналу, фокусируется на мишень видикона с помощью объектива *Об* (рис. 4.23). Мишень представляет собой тонкий слой полупроводника, нанесенный на прозрачную проводящую подложку — сигнальную пластину *СП* (рис. 4.23, а). Каждый перекрываемый пучком элемент можно представить как контур из емкости и светозависимого сопротивления между облучаемой электронным пучком *1* поверхностью и сигнальной пластиной. Поверхность предварительно облучается электронным пучком до потенциала катода.

Процесс формирования изображения происходит в виде зарядовых пакетов. Элементарные емкости элементов мишени разряжаются через локальные сопротивления. Заметим, что чем выше локальная освещенность, тем меньше становится сопротивление локального участка, и соответствующая емкость разряжается сильнее, чем у менее освещенных участков. Создается так называемый зарядовый рельеф.

При очередном цикле сканирования электронным пучком происходит подзарядка конденсаторов. Ток подзарядки зависит от степени разрядки конденсатора. Таким образом, на пластине формируется видеосигнал U_c .

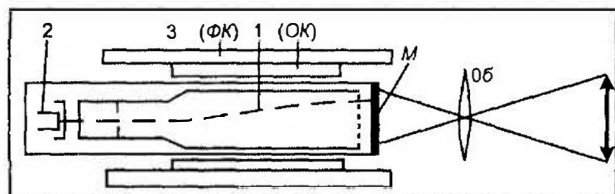
Позднее появилась разновидность видиконов, отличительной особенностью которых являлся состав мишени. Именно состав мишени во многом определял характеристики видикона.

Первые видиконы были созданы на мишени, изготовленной на основе трехсернистой сурьмы (SeS_3). Со временем появились разновидности видиконов.

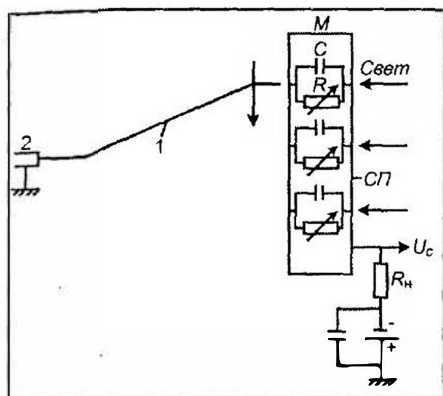
Плюмбикон — видикон, мишень которого представляет собой слой оксида свинца (PbO), нанесенный на прозрачную пленку диоксида олова (SnO_2), служащую сигнальной пластиной. Они характеризуются высокой чувствительностью к свету и малой инерционностью.

Кадмикон — видикон, мишень которого изготовлена на основе селенида кадмия (CdSe).

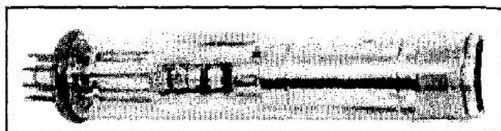
Сатикон — видикон с аморфной мишенью на основе $\text{Se} - \text{As} - \text{Te}$ (по первым буквам сат — икон).



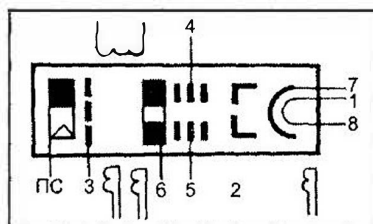
а)



б)



в)



г)

Рис. 4.23. Устройство видеокон (а), эквивалентная схема мишени (б), общий вид (в) и условное обозначение (г) видеокон ЛИ458: 1 — электронный луч; 2 — катод; 3 — устройство управления — фокусирующая катушка (ФК), отклоняющая катушка (ОК); М — мишень, U_c — видеосигнал и схемное изображение (в)

Ньювикон — видеокон с мишенью на основе соединения $ZnS - Cd - Te$.

Халникон — видеокон на основе гетероперехода селенида кадмия.

Кремникон — приборы на основе мозаики р-п-переходов в кремнии. В **суперкремниконе** используются высокоэнергетические электроны, которые ускоряются до 10 кВ.

Ребикон — видеокон с обратным электронным пучком, в котором электрический сигнал снимается с коллектора, принимающего модулированный поток вторичных электронов, усиленных вторично электронным умножителем.

Секон — передающий электронно-лучевой прибор (разновидность видеокон) с мишенью, действие которой основано на явлении вторично-электронной проводимости. Ускоряемые электроны проникают сквозь сигнальную пластину в мишень и на своем пути создают в диэлектрике вторичные электроны, образующие на мишени положительный потенциальный рельеф. Секоны отличаются высокой чувствительностью, малой инерционностью и малыми габаритами, простотой и надежностью.

Ортикон — передающий электронно-лучевой прибор с накоплением заряда на мозаичной светочувствительной мишени и считыванием изображения пучком медленных электронов. Название ортикон (от греческих слов *orthos* — прямой и *eikon* — изображение) обусловлено ортогональным падением развертывающего электронного луча на мишень.

В основу работы ортикона положено физическое явление внешнего фотоэффекта. Изображение просцируется на мишень, покрытую со стороны объекта тонким полупрозрач-

ным проводящим слоем. Это сигнальная пластина. Со стороны электронного луча наносится мелкозернистый фотоактивный слой, представляющий собой мозаичный автокатод. С помощью фототока заряжаются элементарные конденсаторы, образованные зернами мозаики и сигнальной пластиной, создавая на поверхности мишени потенциальный рельеф, соответствующий распределению освещенности по мишени.

Считывающий электронный луч, сканирующий мозаику, нейтрализует накопленный положительный заряд и формирует видеосигнал в цепи сигнальной пластины. У ортика характеристика "свет—сигнал" линейна во всем диапазоне освещенности. Однако созданы *суперортиконы*, резко увеличивающие чувствительность передаваемого сигнала.

Суперортикон — высокочувствительный передающий телевизионный прибор с накоплением заряда, переносом изображения с фотокатода на двустороннюю мишень, со считыванием изображения с мишени медленными электронами и последующим усилением видеосигнала с помощью вторично-электронного умножителя. Суперортиконы работоспособны почти в полной темноте (при освещенности автокатаода 10^{-7} — 10^{-8} лк). Недостаток этих приборов в больших габаритах, весе и большой мощности источников питания.

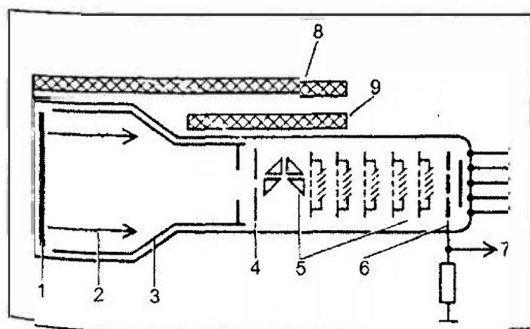
Изокон — прибор класса суперортиконов, в котором имеется система разделения обратного луча, и мишень, пропускающая рассеянные электроны.

Контрасткон — ортикон с особым усилением контраста передаваемого изображения.

Для применения в цветном телевидении используются либо три видикона с соответствующими цветными фильтрами, либо один видикон со специальной конструкцией мишени. В этой мишени вмонтированы структуры светофильтров, обеспечивающие кодирование и разделение сигналов, соответствующих трем основным цветам.

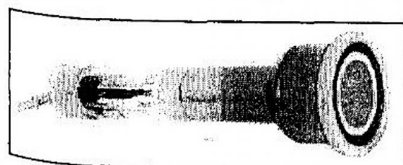
Приборы без накопления заряда

Диссекторы — передающий электронно-лучевой прибор без накопления заряда, служащий для преобразования оптического изображения в последовательность электрических сигналов (рис. 4.24).

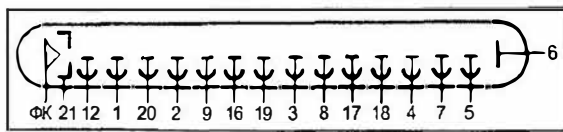


а)

Рис. 4.24. Схема устройства (а), общий вид (б) и схемное обозначение диссектора ЛД604 (в): 1 — фотокатод; 2 — фотозлектрон; 3 — ускоряющие электроды; 4 — диафрагма; 5 — диоды вторично-электронного умножителя; 6 — коллектор; 7 — выходной сигнал; 8 — система фокусировки; 9 — система отклонения; ФК — фотокатод



б)



в)

Диссектор происходит от латинского слова *dissector* — рассекать. В основе работы прибора лежит внешний фотоэффект. Поток фотоэлектронов, плотность эмиссии которых соответствует распределению освещенности на поверхности фотокатода, проходит через вырезающее отверстие (диафрагму). Эта диафрагма делит диссектор на секцию фокусировки и отклонения фотоэлектронов от секции умножения.

В секции умножения электроны поступают на нагрузочный резистор коллектора, формируя видеосигнал.

Фокусировка электронного изображения осуществляется квазиоднородным магнитным полем δ , а отклонение электронного луча — магнитным полем отклоняющих катушек 9. Электронное изображение по площади диафрагмы перемещается строка за строкой.

Отсутствие принципа накопления заряда повышает быстродействие диссекторов и возможность их использования в быстропотекающих процессах.

Отечественной промышленностью была разработана широкая номенклатура передающих электронно-лучевых приборов (рис. 4.25).

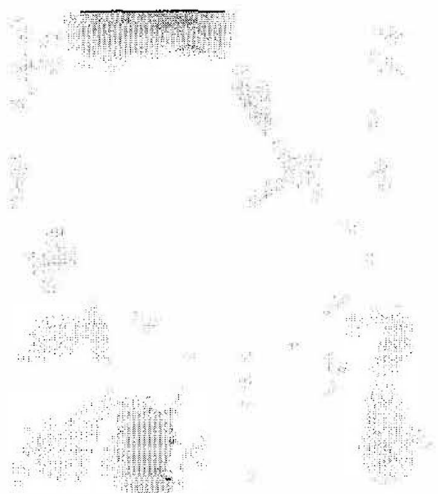


Рис. 4.25. Отечественные передающие электронно-лучевые приборы

4.5.3. Приборы типа "сигнал—сигнал"

К приборам типа "сигнал—сигнал" относятся электронно-лучевые преобразователи электрических сигналов, позволяющие преобразовать последовательность входных электрических сигналов в модифицированную последовательность выходных электрических сигналов.

Это приборы преобразования аналогового сигнала в дискретный, запоминание сигналов с последующим воспроизведением, преобразование телевизионных сигналов по различным стандартам и т. д.

Приборы этого типа делятся на *запоминающие* и *без запоминания*.

В запоминающих приборах электронный пучок модулируется входным электрическим сигналом в процессе сканирования мишени. Мишень представляет собой диэлектрический слой на поверхности металла. Электронный луч формирует на поверхности диэлек-

трика потенциальный рельеф. Считывание осуществляется тем же или другим электронным лучом. Запоминающие преобразователи позволяют многократно воспроизводить однократно записанную информацию.

К запоминающим преобразователям относятся графekon, литокон и другие преобразователи.

Графekon предназначен для преобразования электрических сигналов. Пучком быстрых электронов (~ 10 кВ) осуществляется запись информации в слое диэлектрика на металлической подложке. Кодирование сигнала осуществляется путем записи потенциального рельефа, сохраняющего долгое время свое состояние.

В процессе считывания ток вторичной эмиссии с мишени попадает на коллектор, будучи промодулирован потенциальным рельефом мишени. Большая глубина потенциального рельефа мишени (~ 10 В) позволяет многократно считывать однократно записанный сигнал. Время считывания ~ 1 минута. Стирание ненужного потенциала осуществляется записывающим электронным пучком (рис. 4.26).

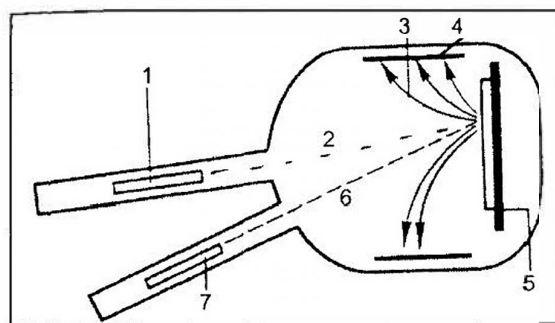


Рис. 4.26. Схема графekon:

- 1 — записывающий прожектор;
- 2 — электронный пучок записи;
- 3 — траектории вторичных электронов;
- 4 — коллектор вторичных электронов;
- 5 — экран; 6 — считывающий электронный луч;
- 7 — считывающий прожектор

Литокон — однолучевой прибор с мишенью из проводящей подложки с мозаикой из диэлектрических элементов. При считывании доля электронного пучка изменяется в соответствии с потенциальным рельефом мишени.

Вычитающие потенциалоскопы представляют собой электронно-лучевые преобразователи, у которых выходной сигнал воспроизводит разность двух последовательно записанных электрических сигналов.

Приборы с запоминанием используются для взаимного преобразования изображений, хранения данных, записи быстропротекающих процессов.

4.5.4. Приборы типа "свет—свет"

Приборы типа "свет—свет" предназначены для преобразования изображения из одной области спектра в другую, а также для усиления яркости изображений и визуализации слабо светящихся объектов, недоступных прямому наблюдению глазом.

К приборам этого типа относятся *электронно-оптические преобразователи* (ЭОП). В основе работы ЭОП лежит принцип преобразования оптического излучения в электронное, его усиление и обратное преобразование электронного изображения в оптическое (рис. 4.27). Усиление обеспечивается процессами ускорения электронов сильным элек-

трическим полем. При этом невидимое глазом изображение преобразуется в оптический спектр. Усиление оптического излучения может достигать несколько порядков.

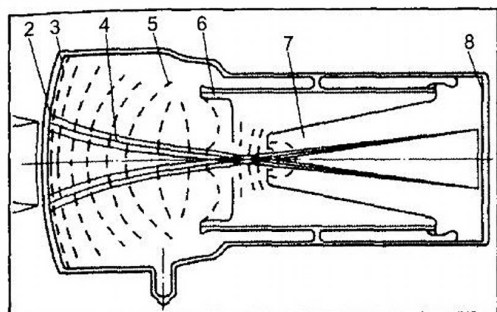


Рис. 4.27. Схема электронно-оптического преобразователя с усилителем изображения на люминесцентном экране:

1 — входное изображение; 2 — изображение на фотокатоде; 3 — фотокатод; 4 — пучок фотоэлектронов; 5 — эквипотенциальные поверхности; 6 — фокусирующий электрод; 7 — анод; 8 — люминесцентный экран

Создавая многокамерные ЭОП, можно достичь усиления до 10^7 раз. Это позволяет регистрировать каждый акт фотоэмиссии.

В качестве фотокатодов используются многощелочные сурьмяно-цезиевые или кислородно-цезиевые покрытия.

Разрешающая способность N , характеризующая четкость изображения, ограничивается абберациями электронно-оптической системы. Обычно значение N составляет порядка 25 штрихов/мм.

ЭОП широко применяются в ИК-технике, спектроскопии, медицине, ядерной физике и т. д.

В последнее время для усиления изображения используются микроканальные пластины, отличающиеся высоким коэффициентом вторичной электронной эмиссии.

4.6. Фотозлектронные приборы

Фотозлектронные приборы — элевтромагнитные приборы, преобразующие электромагнитные сигналы оптического диапазона в электрические токи, напряжения.

К вакуумным фотозлектронным приборам, прежде всего, относятся *фотозлементы вакуумные* и *фотозлектронные умножители*, в которых используется внешний фотоэффект.

4.6.1. Вакуумные фотозлементы

Фотозлемент вакуумный состоит из фотокатода K , анода A и вакуумного баллона (рис. 4.28).

Фоточувствительный слой наносится либо непосредственно на стеклянный баллон (рис. 4.28, а, позиции 1, 2, 4), либо на поверхность специальной подложки, смонтированной внутри баллона (рис. 4.28, а, позиция 3). Световой поток Φ попадает на фотокатод и стимулирует фотозлектронную эмиссию, в результате чего между фотокатодом и анодом создается поток свободных электронов.

Наибольшее распространение получили вакуумные фотозлементы с сурьмяно-цезиевым, многощелочным или кислородно-серебряно-цезиевым фотокатодами.

Применение газонаполненных фотозлементов ограничено их нестабильностью и нелинейностью их световой характеристики.

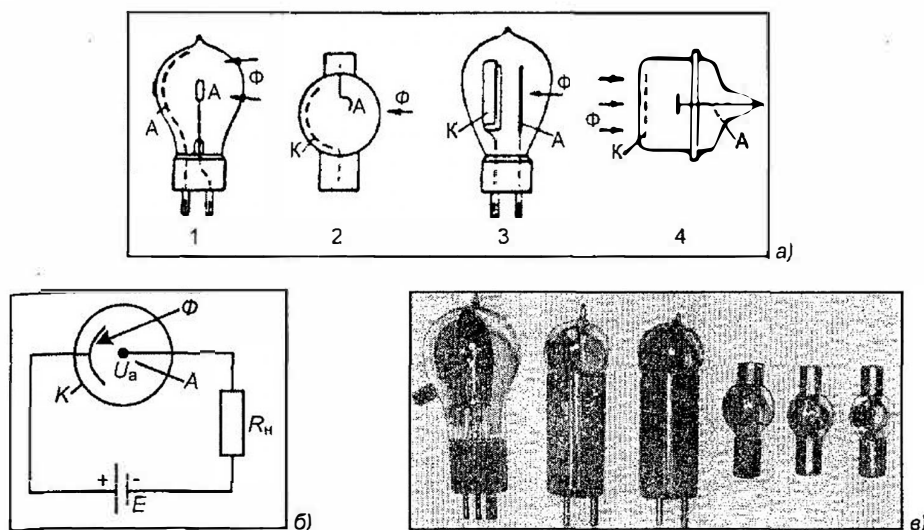


Рис. 4.28. Общий вид фотозлементов различных конструкций (а), схема их включения (б) и образцы промышленных фотозлементов (в)

4.6.2. Фотоэлектронные умножители

Фотоэлектронный умножитель (ФЭУ) предназначен для усиления слабых фототоков.

Его работа основана на эффекте вторичной электронной эмиссии. Фотоэлектронный умножитель состоит из фотокатода, каскада динодов, обеспечивающих умножение электронов за счет вторичной электронной эмиссии, анода и дополнительных электродов, помещенных в вакуумный баллон (рис. 4.29, а).

Световой поток Φ стимулирует фотоэлектронную эмиссию с фотокатода К. Электронно-оптическая система входной камеры направляет эмитированные электроны на систему динодов. В динодах благодаря эффекту вторично-электронной эмиссии происходит умножение числа электронов. На анод попадает поток вторичных электронов, умноженных каждым динодом.

Конструкции фотоумножителей весьма разнообразны, однако принцип одинаков: умножение электронов происходит в системе дискретных динодов. Они имеют корытообразную, коробчатую, торроидальную, либо жалюзийную форму с линейным или круговым расположением. Фототок, благодаря эффекту вторичной электронной эмиссии, может быть увеличен до 10^8 раз.

По функциональному назначению ФЭУ делятся на две крупные группы:

- измерители предельно малых постоянных или медленно меняющихся световых потоков;
- регистраторы слабых кратковременных световых потоков.

ФЭУ широко используется для регистраций слабых излучений вплоть до одиночных квантов, а также в различной оптической аппаратуре.

Разработаны конструкции ФЭУ для работы в разных областях спектра электромагнитного излучения.

Одноканальный электронный фотумножитель — это непрерывный диод или канал, к концам которого приложено напряжение порядка $1 \div 3$ кВ.

На внутренней стороне поверхности канала создан активный слой, обладающий вторичной электронной эмиссией и распределенным электрическим сопротивлением. Перемещение вторичных электронов происходит под действием аксиального электрического поля. Усиление в таком ФЭУ может достигать значений 10^9 .

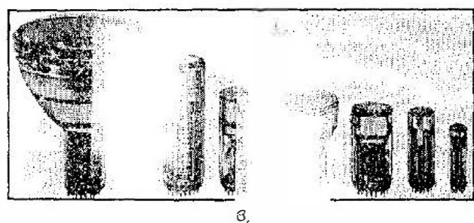
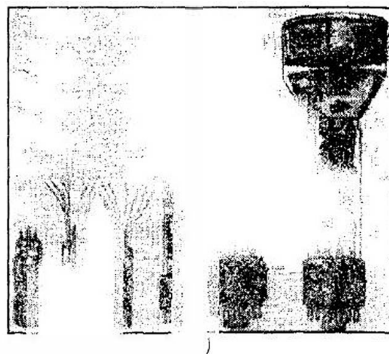
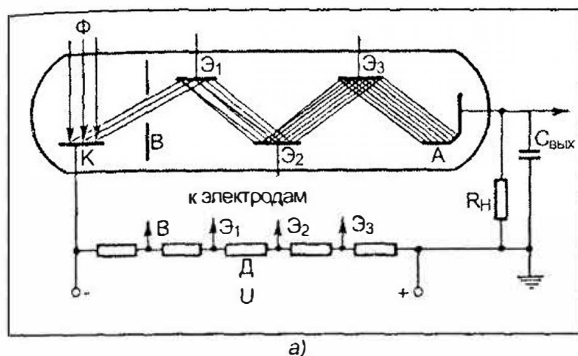


Рис. 4.29. Схема фотозлектронного умножителя (а), ФЭУ с коробчатыми диодами (б), ФЭУ с жалюзийными диодами (в):

К — фотокатод; Э — диоды;
В — входная диафрагма;
А — анод; Д — делитель напряжения

Вторично-электронный умножитель (ВЭУ) представляет собой вакуумное электронное устройство, предназначенное для умножения вторичных электронов. ВЭУ без оболочки называют открытыми и используются в условиях естественного вакуумного пространства.

ВЭУ с оболочкой или закрытого типа широко используются в различной научно-исследовательской и промышленной аппаратуре.

Задачи и упражнения

4.1. Вывод выражений для потенциала, плотности заряда и тока, напряженности поля и скорости электронов в планарном диоде

Рассмотрим диод, в котором ток ограничен пространственным зарядом. Найдите зависимости следующих величин от расстояния до катода x :

1. Потенциала V .
2. Плотности пространственного заряда ρ .
3. Напряженности электрического поля E .

4. Плотности тока J .
5. Скорости электрона v .

Вычислите значение каждого из перечисленных выше параметров в точке, находящейся на половине между анодом и катодом плоского диода. Пусть к аноду приложено напряжение 100 В, а расстояние между катодом и анодом равно 1 см.

Решение

1. В плоскопараллельном диоде при условии ограничения тока пространственным зарядом потенциал изменяется с расстоянием x от катода по закону

$$V = V_a \cdot (x/d)^{4/3}, \quad (4.1.1)$$

где d — расстояние между катодом и анодом, на катоде $V = 0$, а на аноде $V = V_a$.

Тогда

$$V = V_a \left(\frac{x}{d} \right)^{4/3} = \left(\frac{V_a}{d^{4/3}} \right) x^{4/3}.$$

2. Запишем уравнение Пуассона, дающее распределение потенциала по координате x

$$\frac{d^2 V}{dx^2} = -\frac{\rho}{\epsilon_0}, \quad (4.1.2)$$

откуда получаем

$$\rho = -\epsilon_0 \frac{d^2 V}{dx^2}.$$

Но, поскольку

$$V = \left(\frac{V_a}{d^{4/3}} \right) x^{4/3},$$

то, дифференцируя это выражение, имеем

$$\begin{aligned} \frac{dV}{dx} &= \frac{4}{3} \left(\frac{V_a}{d^{4/3}} \right) x^{1/3} \\ \frac{d^2 V}{dx^2} &= \frac{4}{9} \left(\frac{V_a}{d^{4/3}} \right) x^{-2/3} \end{aligned} \quad (4.1.3)$$

Подстановка выражения (4.1.3) в (4.1.2) дает зависимость плотности пространственного заряда от x :

$$\rho = -\epsilon_0 \frac{4}{9} \left(\frac{V_a}{d^{4/3}} \right) x^{-2/3}.$$

3. Для напряженности электрического поля имеем:

$$E = -\frac{dV}{dx} = -\frac{4}{3} \left(\frac{V_a}{d^{4/3}} \right) x^{1/3}.$$

4. Плотность тока определяется как $J = Nqv$, где $v^2 = 2qV/m$, N — плотность электронов. Подставляя в (4.1.2) и интегрируя, получим

$$V_a^{3/2} = \frac{9}{4} \frac{J}{\epsilon_0} \sqrt{\frac{m}{2q}} d^2.$$

Исходя из этого, выражение для плотности тока J примет следующий вид:

$$J = \frac{4}{9} \epsilon_0 \sqrt{\frac{2q}{m}} \frac{V^{3/2}}{x^2} = \frac{4}{9} \epsilon_0 \sqrt{\frac{2q}{m}} \left(\frac{V_a^{3/2}}{d^2} \right). \quad (4.1.4)$$

Видно, что J не зависит от x .

5. Скорость электронов определяется выражением

$$v = \sqrt{\frac{2qV}{m}},$$

где

$$V = \left(\frac{V_a}{d^{4/3}} \right) x^{4/3}. \quad (4.1.5)$$

Подставляя (4.1.5) в значение для скорости электронов, получаем

$$v = \sqrt{\frac{2q}{m}} V^{1/2} = \sqrt{\frac{2q}{m}} \left(\frac{V_a}{d^{4/3}} \right)^{1/2} x^{2/3}.$$

Расчеты

1. В выражении для потенциала

$$V = \left(\frac{V_a}{d^{4/3}} \right) x^{4/3}$$

подставляем численные значения $V_a = 100$ В, $d = 10^{-2}$ м, $x = 0,5 \times 10^{-2}$ м. Таким образом, получаем, что $V = 100 / 2^{4/3} = 100 / 2,52 = 39,7$ В.

2. Подстановка численных значений в выражение для плотности пространственного заряда:

$$\rho = -\epsilon_0 \frac{4}{9} \left(\frac{V_a}{d^{4/3}} \right) x^{-2/3}$$

дает

$$\rho = -8,85 \cdot 10^{-12} \cdot 4 / 9 \cdot 100 \cdot 2^{1/3} \cdot 10^4.$$

или

$$\rho = -6,25 \times 10^{-6} \text{ Кл/м}^3.$$

3. Подставим теперь численные значения в выражение для E :

$$E = -\frac{4}{3} \frac{V_a}{d^{4/3}} x^{1/3} = -4 / 3 \cdot (100 / 1) \cdot (1 / 2^{1/3}) \cdot 10^2 = -(400 / 3) \cdot 1,26 \cdot 10^2 = -10,6 \text{ кВ/м}.$$

4. Подставляя численные значения в выражение для J , получаем

$$J = \frac{4}{9} \epsilon_0 \sqrt{\frac{2q}{m}} \left(\frac{V_a^{3/2}}{d^2} \right) = 2,33 \cdot 10^{-6} \cdot \frac{100^{3/2}}{1^2} \cdot 10^4 = 23,3 \text{ А/м}^2.$$

5. Вычислим значение скорости электрона в точке, одинаково удаленной от анода и катода:

$$v = \sqrt{\frac{2q}{m}} \left(\frac{V_a^{1/2}}{d^{2/3}} \right) x^{2/3} = \frac{\sqrt{2} \cdot 1,76 \cdot 10^{11} \cdot 100^{1/2}}{2^{2/3}} = \frac{5,93 \cdot 10^5 \cdot 10}{1,588} = 3,74 \cdot 10^6 \text{ м/с}.$$

4.2. Движение электрона в триоде с положительной сеткой

Имеется плоская трехэлектродная система, состоящая из катода K , сетки G и анода A (рис. 4.30, а). Сетка состоит из параллельных проволочек диаметром $\varnothing = 10^{-4}$ м, натянутых на расстоянии $\Delta = 1,25 \times 10^{-3}$ м друг от друга. Потенциал сетки $U_g = 200$ В, потенциал анода $U_a = -80$ В. Расстояния d_{kg} и d_{ag} равны 6×10^{-3} м и 10^{-2} м, соответственно. Установить характер движения электронов, определить, сколько колебаний в среднем совершит один электрон, и вычислить период его колебаний.

Решение

При решении задачи пренебрегаем действием пространственного заряда, образуемого электронами. Тогда распределение потенциала в системе будет аналогично тому, которое изображено на рис. 4.30, б.

Электрон, эмитированный из катода, совершает колебательные движения вокруг положительной сетки. Его левая точка возврата будет находиться у поверхности катода, а правая точка возврата — в том месте, где потенциал пространства становится равным нулю.

1. Геометрическая прозрачность сетки составляет

$$q = \frac{\Delta - \varnothing}{\Delta} = 92\%.$$

Полагаем, что при однократном прохождении электронного потока сквозь сетку проходит 92% всех электронов, а захватывается 8%. Таким образом, за первое колебание сквозь сетку пройдет n_0 электронов, за второе — $n_0 q$, за третье — $n_0 q^2$ и т. д.

$$n_m = n_0 q^{m-1}.$$

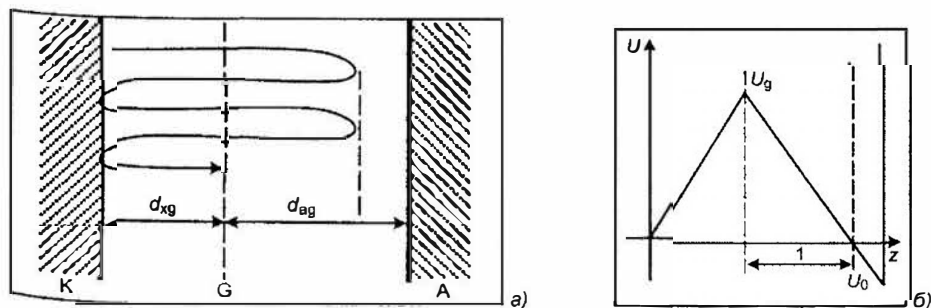


Рис. 4.30. Схема триода (а) и распределение потенциала (б)

Суммируя эту геометрическую прогрессию для $m = \infty$ и разделив сумму на число электронов, получим, что один электрон совершает в среднем N проходов сквозь сетку:

$$N = \frac{\sum_{m=0}^{\infty} n_m}{n_0} = \frac{1}{1-q} = 12,5 \text{ раз.}$$

2. Период колебания электрона равен

$$T = 2(\tau_{kg} + \tau_{ga}),$$

где τ_{kg} и τ_{ga} — время пролета электрона в пространстве "катод—сетка" и "сетка—анод", соответственно.

$$\tau_{kg} = \frac{2d_{kg}}{\left(2\frac{q}{m}U_g\right)^{1/2}}, \quad \tau_{ga} = \frac{2l}{\left(2\frac{q}{m}U_g\right)^{1/2}}.$$

Расстояние l определяем из простых геометрических соображений.

$$l = d_{ga} \frac{U_g}{U_g - U_a},$$

$$\begin{aligned} T &= \frac{2 \cdot 2}{\left(2\frac{q}{m}U_g\right)^{1/2}} \left(d_{kg} + d_{ga} \frac{U_g}{U_g - U_a} \right) = \\ &= \frac{4}{6 \cdot 10^{-4} \sqrt{200}} \left(6 \cdot 10^{-3} + 10^{-2} \frac{200}{200 + 80} \right) = 6,25 \cdot 10^{-3} = 6,25 \times 10^{-9} \text{ с.} \end{aligned}$$

4.3. Отражательный клистрон

Электрон, имеющий скорость v_0 , пролетает через сетку и попадает в область тормозящего однородного поля.

Вычислите время пролета электрона между первым и вторым прохождением его через сетку. С помощью графиков обсудите способ, позволяющий использовать упомянутый выше процесс в отражательном клистроне для генерации СВЧ-мощности.

Решение

Как только электрон попадает в тормозящее поле, он начинает замедляться и его ускорение имеет отрицательный знак:

$$a = -\frac{q}{m} E.$$

Время пролета электрона между первым и вторым прохождением через сетку определяется из уравнения движения с ускорением

$$S = vt + \frac{1}{2} at^2.$$

В этом случае $S = 0$, $v = v_0$, $a = -(q/m) \cdot E$. Следовательно,

$$v_0 t = \frac{1}{2} \frac{q}{m} E t^2,$$

откуда

$$t = \frac{2v_0 m}{qE}.$$

На рис. 4.31 показано устройство отражательного клистрона. Полный резонатор служит как входной, так и выходной цепью. Электроны вылетают с катода и ускоряются напряжением анода.

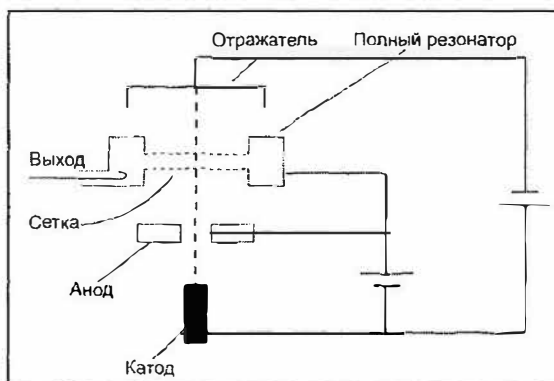


Рис. 4.31. Схема отражательного клистрона

При прохождении обеих сеток резонатора скорости электронов изменяются (происходит модуляция скорости) под действием входного переменного напряжения, приложенного между двумя сетками резонатора. Затем электроны, изменив направление скорости в поле отражателя, который имеет значительный отрицательный потенциал по отношению к катоду, направляются обратно к резонатору и вновь пролетают через его сетки.

Оба эти процесса, происходящие одновременно, создают эффект группирования электронов, который можно пояснить с помощью рис. 4.32. Линией Б (рис. 4.32, а) обозначена траектория электрона, который проходит через зазор между сетками, когда потенциал между ними равен нулю.

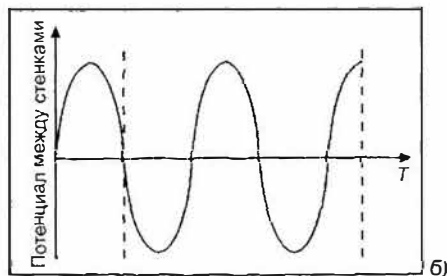
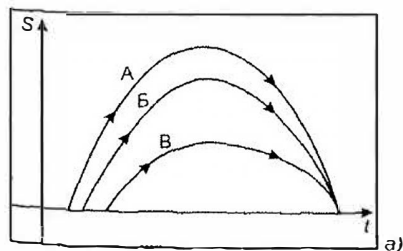


Рис. 4.32. Схема группирования электронов

атем электрон движется по направлению к отражателю, замедляется и, в конечном счете, охватывается к сетке. Электрон, движущийся по траектории A , попадает в зазор, когда потенциал между сетками положителен. Поэтому электрон проходит дальше в направлении отражателя до тех пор, пока не повернется к сетке. Электрон, движущийся по траектории B , входит в зазор, когда потенциал между сетками отрицательный, поэтому электрон пройдет меньшее расстояние по направлению к отражателю, прежде чем он повернется к зазору резонатора. Следовательно, как видно из графиков, каждый из этих электронов может снова вернуться в зазор резонатора в один и тот же момент (происходит группировка электронов).

Проходя через зазор, сгруппированные электроны вследствие индукции заставляют полюсы резонировать, и таким образом энергия может передаваться во внешнюю цепь. Напряжение на отражателе определяет фазу сгустков, тогда как частота сигнала, передаваемого во внешнюю цепь, определяется размерами объемного резонатора.

Можно сконструировать отражательные клистроны, работающие до частот порядка 00 ГГц, при этом диапазон выходной мощности простирается от нескольких киловатт на низких частотах до нескольких ватт на высоких.

4.4. Система электростатического отклонения в электронно-лучевой трубке

Выведите выражение для чувствительности к электростатическому отклонению электронно-лучевой трубки. Изложите требования для получения высокой чувствительности.

Рассчитайте чувствительность к электростатическому отклонению электронно-лучевой трубки из следующих данных: потенциал на последнем от катода аноде равен 2 кВ; отклоняющая система представляет собой пару параллельных пластин длиной 20 мм, расположенных на расстоянии 5 мм; расстояние от точки, расположенной в середине между пластинами, до экрана составляет 0,25 м.

При какой частоте переменного напряжения на отклоняющих пластинах чувствительность к отклонению будет равна нулю?

Решение

Электрон влетает в отклоняющую систему электронно-лучевой трубки из точки O со скоростью v_0 (рис. 4.33). Показано, что электрон притягивается к положительной пластине. Заметим, что горизонтальная компонента скорости не меняется. Полное расстояние, пройденное электроном в вертикальном направлении (ось y) за время t , в течение которого электрон находится между пластинами, определяется из уравнения

$$y = \frac{1}{2} at^2,$$

где $a = (q/m) \cdot E_y$, $t = (x/v_0)$. Следовательно,

$$y = \frac{1}{2} \frac{q}{m} E_y \frac{x^2}{v_0^2}. \quad (4.4.1)$$

Отклонение y_1 между отклоняющими пластинами в точке L получается из уравнения (4.4.1) подстановкой $x = L$:

$$y_1 = \frac{1}{2} \frac{q}{m} E_y \frac{L^2}{v_0^2}. \quad (4.4.2)$$

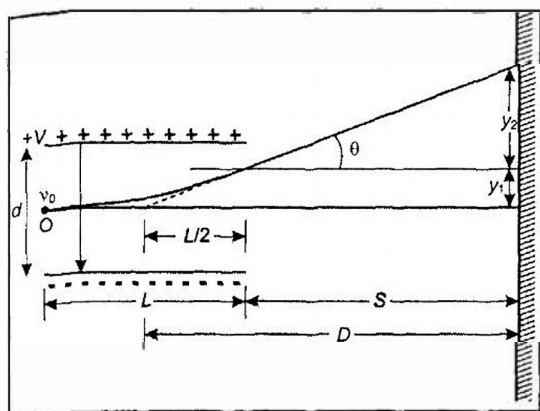


Рис. 4.33. Траектория электрона, движущегося между двумя параллельными пластинами, и последующая траектория электрона при движении к экрану

Из рис. 4.33 следует, что

$$y_2 = S \cdot \operatorname{tg} \theta \quad (4.4.3)$$

и что при $x = L$

$$\operatorname{tg} \theta = \frac{dy}{dx}.$$

Из уравнения (4.1.4) имеем величину тока, ограниченного пространственным зарядом

$$I = \frac{4}{9} \epsilon_0 \left(\frac{2q}{m} \right)^{1/2} V_a^{3/2} \frac{A}{d^2}, \quad (4.4.4)$$

где A — площадь электродов, d — расстояние между ними.

Дифференцируя уравнение (4.4.4) по x , получаем

$$\frac{dx}{dy} = \frac{e}{m} E_y \frac{x}{v_0^2},$$

откуда находим

$$\operatorname{tg} \theta = \frac{e}{m} E_y \frac{L}{v_0^2}.$$

Следовательно,

$$y_2 = S \frac{e}{m} E_y \frac{L}{v_0^2}. \quad (4.4.5)$$

Полное отклонение

$$y_{\text{полн}} = y_1 + y_2 = \frac{1}{2} \frac{e}{m} E_y \frac{L^2}{v_0^2} + S \frac{e}{m} E_y \frac{L}{v_0^2} = \frac{e}{m} \frac{E_y}{v_0^2} L \left(\frac{L}{2} + S \right). \quad (4.4.6)$$

Из рисунка видно, что расстояние от центра конденсатора до экрана равно

$$(L/2) + S = D,$$

а

$$E_y = V/d,$$

где V — разность потенциалов между пластинами, d — расстояние между ними.

Начальная кинетическая энергия электрона равна $\frac{1}{2}mv_0^2$. Далее энергия приобретает в результате ускорения электрона за счет потенциала конечного анода V_a :

$$\frac{1}{2}mv_0^2 = qV_a. \quad (4.4.7)$$

Выражение (4.4.6) можно переписать в виде

$$y_{\text{полн}} = \frac{VLD}{2V_a d}.$$

Чувствительность к электростатическому отклонению определяется как отклонение на 1 В:

$$S_{\text{откл}} = \frac{y_{\text{полн}}}{V}.$$

Следовательно,

$$S_{\text{откл}} = \frac{1}{2} \frac{LD}{V_a d}.$$

При выводе этих формул мы предполагали, что электрон влетает в зазор между пластинами в горизонтальном направлении и что поле однородно.

Довольно очевидно, что если величины L и D настолько велики, насколько позволяет конструкция трубки, а V_a и d достаточно малы, то чувствительность будет высокой. Однако чтобы получить максимально возможную яркость пятна, значение ускоряющего потенциала V_a должно быть большим. Следовательно, выбирая практические значения этих параметров, приходится идти на компромисс между противоречивыми требованиями.

Согласно условию задачи, $V_a = 2 \times 10^3$ В, $L = 20 \times 10^{-3}$ м, $D = 0,25$ м и $d = 5 \times 10^{-3}$ мм. Следовательно,

$$S_{\text{откл}} = \frac{1}{2} \cdot \frac{20 \cdot 10^{-3} \cdot 0,25}{2 \cdot 10^3 \cdot 5 \cdot 10^{-3}} \text{ м/В} = 0,25 \text{ мм/В}.$$

Чувствительность к отклонению равна нулю, если время пролета электрона между пластинами равно периоду переменного напряжения, приложенного к пластинам. Время пролета $t = L / v_0$, а из соотношения (4.4.6) находим

$$v_0 = \left(\frac{2qV_a}{m} \right)^{1/2}.$$

Нулевая чувствительность к отклонению получается, когда период переменного напряжения равен L / v_0 . При этом частота определяется формулой

$$f = \frac{v_0}{L} = \frac{1}{L} \left(\frac{2qV_a}{m} \right)^{1/2}.$$

Следовательно,

$$f = \frac{1}{20 \cdot 10^{-3}} \left(2 \cdot 1,76 \cdot 10^{11} \cdot 2 \cdot 10^3 \right)^{1/2} = 1,33 \cdot 10^9 \text{ Гц} = 1,33 \text{ ГГц}.$$

Контрольные вопросы

1. Что такое электронные лампы?
2. Какие параметры триода вы знаете?
3. Что такое клистрон? Опишите конструкцию клистрона.
4. Перечислите преимущества клистрона перед лампами СВЧ-диапазона.
5. Что такое лампа бегущей волны? Как она устроена?
6. Что такое лампа обратной волны? Как она устроена?
7. Что такое приборы М-типа и чем они отличаются от приборов О-типа?
8. Что такое магнетрон? Опишите конструкцию магнетрона.
9. Какие электронно-лучевые приборы вы знаете?
10. Что такое кинескоп и как он устроен?
11. Что такое иконоскоп и как он устроен?
12. Что такое видикон, и какие разновидности видиконов вы знаете?
13. Что такое ЭОП и как он устроен?
14. Что такое фотоумножитель, и какие явления положены в основу его работы?

Рекомендуемая литература

1. Анаев Н. А., Шишкин Г. Г. Электронные приборы. — М.: МАИ, 1996.
2. Добрецов Л. Н., Гомоюнова М. В. Эмиссионная электроника. — М.: Наука, 1966.
3. Лебедев И. В. Техника и приборы СВЧ. — М.: Высшая школа, 1972.
4. Линч П., Николаидес А. Задачи по физической электронике. Под ред. проф. Г. В. Скроцкого. — М.: Мир, 1975.
5. Левитский С. М. Сборник задач и расчетов по физической электронике. — Изд. Киевского университета, 1964.
6. Соболева Н. А., Мсламид А. Е. Фотоэлектронные приборы. — М.: Высшая школа, 1974.
7. Шука А. А. Вакуумная электроника. Учебное пособие. — М.: МИРЭА, 2002.

5. Плазменная электроника

5.1. Основные понятия

Плазменная электроника — раздел электроники, в котором изучаются процессы коллективного взаимодействия потоков заряженных частиц с плазмой и ионизированным газом, приводящие к возбуждению в системе волн и колебаний, а также использование эффектов такого взаимодействия для создания приборов и устройств электронной техники.

Плазма (от греческого *plasma* — выявленное, оформленное) представляет собой частично или полностью ионизированный газ, в котором плотности положительных (ионов) и отрицательных зарядов (электронов) практически одинаковы. Это название было предложено американским физиком Ленгмюром в 1923 году.

Под *ионизированным газом* понимается газ, в котором значительная часть атомов потеряли или приобрели по одному или несколько электронов и превратились в ионы. В зависимости от количества потерянных или приобретенных ионов различают различные степени ионизации газа.

Существуют различные механизмы ионизации газа, основные из которых: термическая ионизация при нагреве газа, ионизация светом — фотоионизация, бомбардировка газа заряженными частицами. Ионизированный газ превращается в четвертое агрегатное состояние вещества. Другими словами, пространственные заряды электронов и положительных ионов должны компенсировать друг друга и тогда плазма считается квазинейтральной. Самый простой способ получения плазмы — зажигание газового разряда.

5.2. Электрический разряд в газах

Электрический разряд в газах возникает при прохождении электрического тока через ионизированные газы.

В зависимости от характера приложенного поля, от давления газа, формы и расположения электродов различают различные типы разрядов. При небольшом напряжении (≈ 100 В) через газ наблюдается прохождение тока ($\approx 10^{-15}$ А). Видимо, этот ток обусловлен слабой ионизацией газа космическими лучами. Разряды в постоянном поле возникают при прохождении тока в газе и могут быть представлены в виде вольтамперной характеристики (рис. 5.1).

При увеличении напряжения ток резко возрастает (участок 1—2 кривой на рис. 5.1) и достигает насыщения (участок 2—3). При небольших давлениях ($\approx 10^{-1}$ — 10 Торр) и большом ограничительном сопротивлении R внешней цепи при токе $\approx 10^{-6}$ А зажигается *темный* или *таунсендовский разряд* (участок 2—3).

Этот процесс характеризуется полным вытягиванием всех образовавшихся в газе зарядов.

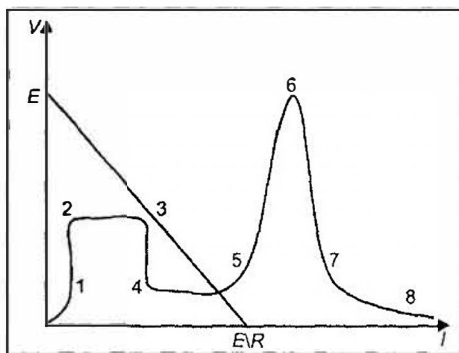


Рис. 5.1. Вольтамперная характеристика газовых разрядов

Двигаясь под действием электрического поля в газовой среде, электроны производят на своем пути ионизацию молекул газа. Мерой такой ионизации является степень ионизации α , определяемая как

$$\alpha = A p \cdot \exp\left[-B/(E/p)\right], \quad (5.1)$$

где A [м^{-1}], B [$\text{В/м}\cdot\text{мм рт. ст.}$], E/p [$\text{кВ}\cdot\text{м}^{-1}\cdot\text{мм рт. ст.}$] — константы, имеющие определенные табличные значения для различных газовых сред.

Вторичные электроны, появившиеся в результате ионизации газа, также движутся в электрическом поле и ионизируют газ. Лавинообразное размножение электронов приводит к тому, что поток в конце разрядного промежутка значительно больше потока электронов в его начале. Коэффициент k такого газового усиления в однородном поле определяется как

$$k = \exp(\alpha d), \quad (5.2)$$

где d — ширина разрядного промежутка.

Положительные ионы, возникшие в результате ионизации, выбивают на отрицательном электроде дополнительное количество электронов. Вклад вторичной ионно-электронной эмиссии увеличивает значение k :

$$k = \frac{\exp(\alpha d)}{1 - \gamma \cdot (\exp(\alpha d) - 1)}, \quad (5.3)$$

где γ — коэффициент вторичной ионно-электронной эмиссии материала катода, значения которого для различных газов и материалов можно найти в справочниках.

Обращение в нуль знаменателя в формуле (5.3) означает, что небольшой поток первичных электронов способен возбудить значительный разрядный ток и зажечь самостоятельный разряд.

В зависимости от типа газа, конфигурации электродов и расстояния между ними при дальнейшем небольшом увеличении напряжения наблюдается возникновение самостоятельного разряда. Этот тип разряда не зависит уже от внешних факторов и не нуждается во внешнем источнике ионизации (участок 3—4). При этом резко падает напряжение разряда. В этом случае говорят об *электрическом пробое* газового промежутка. При этом возникают новые электроны, которые под действием поля набирают энергию и ионизируют газ (участок 4—5).

Критерием возникновения самостоятельного разряда является условие

$$\gamma \cdot (\exp(\alpha d) - 1) = 1. \quad (5.4)$$

Пробойным потенциалом или потенциалом зажигания является разность потенциалов, при котором выполняется условие (5.4), и определяется он из соотношения

$$U_x = Bpd / \ln(Apd) / C, \quad (5.5)$$

где константы A и B уже упоминались, а константа $C = \ln(1 + 1/\gamma)$.

При небольшом внешнем сопротивлении R возникает *тлеющий разряд* (участок 3—6) при токе $\approx 10^{-6}$ — 10^{-1} А и напряжении до 10^3 В. Характерный поперечный размер разрядной трубки составляет ≈ 1 см. При этом образуется однородный светящийся столб. Это так называемый *положительный столб разряда*, который представляет собой плазму тлеющего разряда. Она является неравновесной с невысокой степенью ионизации, и ее электронная температура составляет $T_e \approx 10^4$ К.

Если повышать напряжение дальше, то после газового пробоя зажигается *дуговой разряд*. Для него характерно высокое значение тока и достаточно низкое напряжение (участок 7—8). Дуговой разряд визуально представляет собой ярко светящийся газовый столб.

В дуговом разряде плазма равновесия такая, что температура электронов примерно равна температуре ионной составляющей $T_e \approx T_i \approx 10^4$ К при достаточно высокой степени ионизации.

При резко неоднородном поле возникает *коронный разряд*. Ионизация и возбуждение молекул газа в коронном разряде наблюдается только в области, где поверхность электродов искривлена и возникает наибольшая напряженность электрического поля.

Для цилиндрической системы электродов возникновение коронного разряда происходит при напряженности поля, определяемой эмпирической формулой

$$E_{кр} = 31p \left[1 - \frac{0,38}{\sqrt{pr_0}} \right], \quad (5.6)$$

где $E_{кр}$ — напряженность электрического поля, кВ/см; p — давление газа, атм.; r_0 — радиус коронирующего электрода.

На рис. 5.1 изображена нагрузочная прямая $E - E/R$, с помощью которой путем подбора сопротивления нагрузки R в нагрузочной прямой можно получить определенный тип газового разряда.

Разряды можно классифицировать по состоянию плазмы разряда и частоте электрического поля.

По характеру ионизационного состояния газа можно различать:

- ☐ пробой газа;
- ☐ поддержание электрическим полем неравновесной плазмы;
- ☐ поддержание равновесной плазмы.

В свою очередь электрические поля делятся по частотам:

- ☐ постоянные, включая низкочастотные;
- ☐ высокочастотные (ВЧ) с $f \approx 10^5$ — 10^8 Гц;

- сверхвысокочастотные (СВЧ) с $f \approx 10^9 - 10^{11}$ Гц;
- оптические (от ультрафиолета до инфракрасного диапазона).

Все эти комбинации электрического поля применяются в плазменной электронике.

Если разрядный промежуток достаточно велик, то основную его часть занимает плазменный столб. Это относится как к тлеющему, так и к дуговому разряду. Параметры тлеющего столба зависят от условия стационарности, которое заключается в том, что должно соблюдаться равенство средней частоты ионизации и обратного времени жизни заряженных частиц.

Ионизация происходит под действием электронных ударов, а распределение электронов по энергии экспоненциально спадает. Средняя энергия электронов в столбе составляет порядка нескольких эВ.

При малых давлениях и токах величина температуры ионов порядка комнатной. В таком случае плазма неравновесная. Для такой плазмы характерны неустойчивости, которые выражаются в формировании *страт* — чередующихся светлых и темных полос в разряде, а также в шнуровании тока, что приводит к отрыву плазмы от стенок. Неустойчивости можно избежать, варьируя напряженность электрического поля и его частоту. Для создания электронных плазменных приборов важным моментом является получение стационарного газового разряда с равновесной плазмой.

5.3. Процессы в плазме

Плазму можно получить при разогреве вещества, находящегося в газообразном или твердом состоянии. Плазму называют четвертым состоянием вещества. Состояние равновесной плазмы, как и всякого газа, определяется ее составом, концентрацией компонент и температурой. Обозначим парциальные концентрации компонент плазмы n_α , придавая индексу α значения a (нейтральные частицы), i (ионы), q (электроны). Вообще говоря, плазма может содержать различные виды (сорты) атомов и ионов. В большинстве случаев будем рассматривать так называемую простую плазму, состоящую из нейтральных частиц одного сорта, однозарядных ионов того же сорта и электронов. Тогда степень ионизации η можно определить как отношение концентрации ионов и нейтральных атомов:

$$\eta = n_i / (n_i + n_a). \quad (5.7)$$

Температуру плазмы T будем выражать в энергетических единицах. Она связана с обычно используемым определением температуры соотношением:

$$T = kT_K, \quad (5.8)$$

где T_K — температура, К; k — постоянная Больцмана.

В равновесной плазме параметры концентрации и температуры полностью характеризуют ее состояние. Температура такой плазмы определяет не только среднюю энергию, но и распределение частиц по скоростям (максвелловское распределение). По концентрации и температуре плазмы можно найти степень ионизации, концентрацию ионов, возбужденных атомов, фотонов и т. п. Однако далеко не всегда плазму можно считать равновесной. В частности, газоразрядная плазма, получаемая обычно в лаборатории, далека от равновесия. В некоторых случаях встречается так называемое частичное равновесие, при котором распределение скоростей заряженных и нейтральных частиц является максвелловским.

Температуры, определяющие это распределение для электронов и тяжелых частиц, различны. Для такой неизотермической плазмы можно ввести отдельно электронную и ионную температуры — T_e , T_i . В общем случае распределение скоростей заряженных частиц в неравновесной плазме может существенно отличаться от максвелловского. Однако здесь мы будем также говорить о температуре компонент плазмы, определяя ее как меру средней энергии хаотического движения частиц. Разумеется, для получения полной информации о поведении неравновесной плазмы сведений о средних энергиях (температуре) компонент недостаточно, т. к. необходимо еще знать функцию распределения частиц по скоростям.

Характерной особенностью плазмы является ее макроскопическая нейтральность, поддерживаемая вследствие взаимной компенсации пространственного заряда положительных ионов и электронов. Однако такая компенсация имеет место лишь в достаточно больших объемах и на протяжении больших интервалов времени. Поэтому говорят, что плазма — квазинейтральная среда. Размеры областей и промежутки времени, в пределах которых может нарушаться компенсация объемного заряда, называют *пространственным и временным масштабами разделения зарядов*.

При нарушении квазинейтральности плазмы в объеме возникают пространственные электрические поля и пространственные заряды. Как правило, возникают процессы, приводящие к восстановлению квазинейтральности плазмы.

Степенью ионизации плазмы называется число α , определяемое в условиях термодинамического равновесия формулой Саха:

$$\alpha = \frac{1}{\sqrt{1+K}},$$

где величина K определяется следующим образом

$$K = N_\lambda \cdot \exp \frac{I}{kT}, \quad (5.9)$$

где I — энергия ионизации, эВ; $N_\lambda = n\lambda_e^3$ — число всех частиц в кубе с ребром

$$\lambda_e = h \cdot \sqrt{2\pi / m_e q k T},$$

где k — постоянная Больцмана, h — постоянная Планка, T — температура плазмы.

В зависимости от величины α говорят о слабо, сильно или полностью ионизированной плазме. Различают высокотемпературную плазму с $T \geq 10^6$ — 10^8 К и низкотемпературную плазму с $T \leq 10^5$ К.

Существует два важнейших показателя, согласно которым свойства плазмы отличаются от свойств нейтрального газа:

1. Взаимодействие частиц в плазме определяется кулоновскими силами притяжения или отталкивания, а не только температурными процессами. Такое взаимодействие электрически заряженных частиц является коллективным.
2. Электрические и магнитные поля сильно действуют на плазму, формируя в ней электрические заряды и токи.

Квазинейтральность плазмы соблюдается в том случае, если линейные размеры области плазмы много больше дебаевского радиуса экранирования.

Физический смысл дебаевского радиуса экранирования заключается в том, что он является пространственным масштабом в плазме (или полупроводниках), на котором экранируется поле заряженной частицы. Причиной экранирования какого-либо заряда является процесс преимущественного группирования заряженных частиц противоположного знака. Если заряженная частица с зарядом Zq создаст электрический потенциал Φq , то дебаевский радиус экранирования r_D определяется из выражения

$$Zq = \Phi q / r \cdot \exp(-r / r_D).$$

При этом происходит нейтрализация заряда системы на расстоянии $\sim r_D$. В этом случае плазму можно рассматривать как квазинейтральный коллектив, содержащий большое число заряженных частиц и занимающий область с линейными размерами $L \gg r_D$.

$$r_D = \sqrt{\frac{kT_e \cdot T_i}{4\pi q_e q_i (n_e T_e + n_i T_i)}}, \quad (5.10)$$

где q_e и q_i — заряды электронов и ионов, n_e и n_i — электронная и ионная плотности, T_e и T_i — температура электронной и ионной составляющей плазмы, соответственно.

Плазма называется *идеальной*, если потенциальная энергия взаимодействия частиц мала по сравнению с тепловой энергией.

В плазме возникают продольные колебания пространственного заряда — ленгмюровские волны, угловая частота которых определяется соотношением:

$$\omega_p = \sqrt{\frac{4\pi n q_e^2}{m}}, \quad (5.11)$$

где ω_p — плазменная частота, q_e , m — заряд и масса электрона.

Наличие собственных колебаний и волн — характерное свойство плазмы. Электроны и ионы в плазме движутся по спиралям: электроны вращаются по часовой стрелке, а ионы против часовой стрелки.

Магнитные моменты μ круговых токов равны

$$\mu = mv^2 / 2 \cdot B, \quad (5.12)$$

где m — масса, v — масса и скорость частиц, B — вектор индукции магнитного поля.

Как и всякое диамагнитное вещество, плазма выталкивается из области сильного магнитного поля в область слабого поля.

Важной характеристикой плазмы является длина свободного пробега частицы L :

$$L = 1 / n\chi,$$

где величина $\chi = \frac{v}{nV}$, v — частота столкновений, n — число частиц со скоростью V . Частота столкновений определяется по формуле

$$v = \frac{1}{\tau},$$

где τ — время между столкновениями.

Следует особо отметить, что длина свободного пробега в плазме существенно отличается от длины свободного пробега в газах.

В плазме различают три времени взаимодействия при столкновении:

1. τ_{ee} — максвелловское распределение для электронов.
2. τ_{ii} — максвелловское распределение для ионов: $\tau_{ii} = \tau_{ee} \sqrt{\frac{m_i}{m_e}}$.
3. τ_{ei} — максвелловское распределение для плазмы: $\tau_{ei} = \tau_{ee} \frac{m_i}{m_e}$.

Различие этих времен обязывает ввести понятие температуры для ионной T_i и электронной T_e составляющей плазмы. Если процессы в плазме не завершены, то $T_e \neq T_i$.

Столкновения частиц определяют диссипативные свойства квазинейтральной плазмы.

Электропроводность σ полностью ионизированной плазмы не зависит от плотности плазмы и пропорциональна $T^{3/2}$.

При $T \approx 1,5 \times 10^7$ К электропроводность плазмы превосходит электропроводность серебра. При увеличении температуры свойства плазмы приближаются к свойствам сверхпроводника.

Одним из методов описания плазмы является движение частиц. В плазме заряженная частица летит со скоростью V_n вдоль магнитной силовой линии, вращаясь по спирали с плазменной частотой $\omega_p = \sqrt{\frac{4\pi n q_e^2}{m}}$, где q и m — масса и заряд частицы, n — плотность частиц (рис. 5.2).

В общем случае на частицу зарядом q и массой m , дрейфующую в плазме, действует сила F , которая определяется выражением:

$$F = mg + qE - \mu \nabla B - \bar{n} \frac{m v_n^2}{R} + m v_l.$$

В этом выражении первое слагаемое является составляющим гравитационного поля, где g — ускорение свободного падения. Второе слагаемое характеризует дрейф в электрическом поле E . Третье слагаемое — диамагнитная составляющая в поле с индукцией \vec{B} и градиентом ∇ по координатам и магнитной проницаемостью μ . Четвертое слагаемое обусловлено центробежной силой в искривленном поле по нормали \bar{n} . Пятое слагаемое связано с изменением электрического дрейфа и называется *поляризационной силой*.

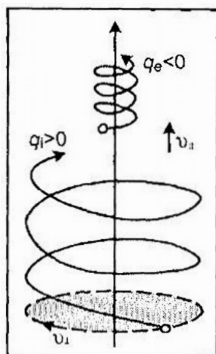


Рис. 5.2. Схема вращения электрона q_e и положительного иона q_i по ларморовской спирали

Если $E = 0$, $g = 0$, то на частицу действуют только диамагнитные и центробежные силы. В этом случае частица дрейфует по бинормали, а траектории медленных частиц представляют собой поверхности, сечения которых имеют форму бананов.

5.4. Излучение плазмы

Спектр излучения низкотемпературной плазмы состоит из отдельных спектральных линий.

В газосветных трубках, например, наряду с процессами ионизации происходят процессы рекомбинации ионов с электронами. В результате рекомбинационных процессов формируется рекомбинационное излучение со спектром в виде широких полос.

Спектр излучения высокотемпературной плазмы более разнообразен.

Тормозное излучение с непрерывным рентгеновским спектром возникает при столкновении электронов с ионами в электрическом поле.

Интенсивность тормозного излучения пропорциональна квадрату ускорения заряженной частицы. В свою очередь ускорение обратно пропорционально массе частицы. Поэтому в одном и том же поле тормозное излучение электрона гораздо мощнее тормозного излучения другой частицы, например протона.

Интенсивность тормозного излучения пропорциональна квадрату атомного номера Z ядра, в поле которого тормозится электрон, что следует из закона Кулона.

Спектр тормозного излучения непрерывен и ограничен максимально возможной энергией фотонов.

Угловое распределение тормозного излучения зависит от кинетической энергии электрона T_e . В нерелятивистском случае, когда $T_e < m_e c^2$ (m_e — масса электрона), угловое распределение излучения подобно угловому распределению электрического диполя, перпендикулярного к плоскости траекторий электрона. При ультрарелятивистских энергиях электрона, когда $T_e \gg m_e c^2$, тормозное излучение направлено по траектории движения электрона и концентрируется в пределах конуса с углом

$$\varphi \approx \frac{m_e c^2}{T_e} \text{ [рад].}$$

Это свойство используется для получения интенсивных пучков фотонов высокой энергии и γ -квантов.

В магнитном поле ларморовское вращение релятивистских электронов в плазме приводит к появлению *магнитотормозного* или *синхротронного излучения*. Синхротронное излучение обладает практически непрерывным спектром и сосредоточено в направлении мгновенной скорости частицы в узком конусе с углом раствора

$$\psi = m_e c^2 / E,$$

где m_e — масса электрона, E — ее энергия.

Синхротронное излучение лежит в ультрафиолетовой области спектра, а также в областях мягкого рентгеновского излучения. Синхротронное излучение имеет хорошие перспективы использования в рентгенолитографических процессах микроэлектроники.

Аналогичное излучение нерелятивистских частиц, движущихся по круговым или спиральным траекториям, называется *циклотронным излучением*. Излучение отдельной час-

тицы в общем случае эллиптически поляризовано, большая ось эллипса поляризации расположена в проекции магнитного поля.

Плазма может генерировать *корпускулярное излучение*, формируемое за счет быстрых частиц, вылетающих из неравновесной плазмы в результате развития различного типа неустойчивостей.

5.5. Диагностика плазмы

Диагностика плазмы заключается в определении параметров плазмы, которые характеризуют ее состояние. Диагностика плазмы весьма непростая процедура, поскольку плазма является многокомпонентной неравновесной, неоднородной системой.

Среди параметров, определяемых в процессе диагностики, следует отметить:

- форму и месторасположение плазмы;
- плотность электронов n_e , ионов n_i , атомов n_a , фотонов;
- статистические распределения этих частиц по скоростям;
- интенсивность излучения;
- коэффициент поглощения;
- частоту столкновений компонентов;
- коэффициент диффузии.

Помещение датчика в плазму искажает ее параметры. Как правило, используются бесконтактные методы, в которых поле и излучение несут информацию о состоянии плазмы.

Используются пассивные и активные методы измерения параметров плазмы.

Пассивные методы основаны на регистрации излучений и потоков частиц из плазмы или измерения характеристик окружающих полей.

Активные методы основываются на измерении характеристик внешнего зондирующего излучения при его прохождении через плазму и на отклике плазмы на зондирующий луч.

Целенаправленное создание в плазме возмущений и исследование динамики их релаксации позволяют определить локальные характеристики плазмы. Динамика плазмы может быть исследована с помощью скоростной оптической развертки.

Спектроскопические исследования позволяют по излучению плазмы оценить значения концентраций электронов n_e , ионов n_i и т. д.

Широко используются методы зондирования плазмы когерентным электромагнитным полем. Это может быть как излучение СВЧ-диапазона, так и лазерное излучение в оптическом или ИК-диапазонах. В этом методе используются интерферометрические методы измерений в схемах типа интерферометров Маха — Цендера, Майкельсона.

Зондирование плазмы является самостоятельной областью научных исследований.

Задачи и упражнения

5.1. Число актов ионизации, совершаемое электроном на 1 м дрейфового движения

Электрон движется в неоне под действием электрического поля напряженностью $E = 5 \times 10^3$ В/м. Давление газа $p = 0,25$ мм рт. ст. Определить α — число актов ионизации, совершаемых электроном на пути движения в 1 м.

Решение

Сделаем ряд допущений:

- вероятность ионизации аппроксимируется зависимостями: $w_i = 0$ при $U < U_i$, $w_i = 1$ при $U > U_i$;
- скорость хаотического движения электронов не учитывается;
- при каждом соударении электрон полностью теряет приобретенную им энергию.

У электрона, движущегося в газе, вероятность свободного пробега длиной более чем l будет пропорциональна $\exp(-l/\lambda)$. Поэтому из N свободных пробегов лишь $N \cdot \exp(-l/\lambda)$ будут по длине превосходить l .

Двигаясь вдоль поля E , электрон на пути l приобретает энергию qEl .

Если $l \geq U_i/E$, то энергия, приобретенная на этом пути, будет достаточна для ионизации молекулы газа, и, согласно первому допущению, ионизация будет иметь место. Таким образом, из N свободных пробегов электрона лишь $N \cdot \exp(-U_i/E\lambda)$ будут приводить к ионизации.

Из этих допущений вытекает, что движение электрона будет прямолинейным. На пути в l м он совершит Q соударений с молекулами газа; из них только $Q \cdot \exp(-U_i/E\lambda)$ будут приводить к ионизации. Тогда

$$\alpha = Q \cdot \exp(-U_i/E\lambda) = Q_0 \cdot \exp(-Q_0 p U_i/E) = \\ = 1520 \cdot 0,25 \cdot \exp[(21,5 \cdot 1520 \cdot 0,25) / 5000] = 74 \text{ м}^{-1}.$$

5.2. Вольтамперная характеристика газового разряда.**Коэффициенты первичной и вторичной ионизации Таунсенда**

Катод плоскопараллельного вакуумного диода с однородным полем облучается слабым ультрафиолетовым светом. Начертите вольтамперную характеристику от очень низкого напряжения вплоть до напряжения пробоя и кратко обсудите физические механизмы, которыми определяется форма полученной кривой.

Предполагая, что плотность электронного тока на катоде равна J_e , выведите выражение для плотности тока положительных ионов на катоде. Коэффициент первичной ионизации Таунсенда обозначьте через α . При достаточно высоком напряжении между электродами самостоятельный разряд становится самостоятельным, и в этом случае говорят, что наступает пробой. Выведите условие пробоя через коэффициент ионизации α , длину межэлектродного пространства d и коэффициент вторичной эмиссии γ .

Решение

Соответствующая вольтамперная характеристика приведена на рис. 5.3.

На ВАХ можно выделить четыре различные области.

1. Поскольку катод облучается ультрафиолетовым светом, испускаются фотоэлектроны и I_a растет с увеличением V_a .
2. Область насыщения, т. е. все эмитированные электроны собираются анодом. Значение тока насыщения I_0 зависит от интенсивности света. Поскольку освещение слабое, это значение невелико.
3. При $V_a \gg V_i$ наступает ионизация, и ток увеличивается (V_i — потенциал ионизации газа).

4. Дальнейший рост тока связан с появлением вторичных электронов, выбиваемых из катода в результате его бомбардировки положительными ионами.

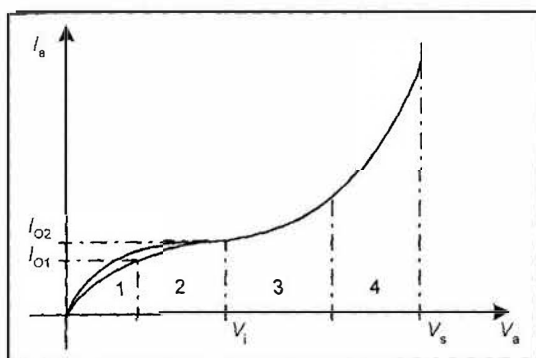


Рис. 5.3. Вольтамперная характеристика газового разряда при УФ облучении

В конечном счете, этот механизм при напряжении зажигания самостоятельного разряда V_s приводит к пробую.

При $V_a = V_s$, как показано на рис. 5.4, а, ионизация может происходить в результате соударений на аноде.

При $V_a > V_s$ такие соударения могут происходить не только у самого анода, но и в глубине межэлектродного промежутка. При $V_a > 2V_s$ внутри лампы может иметь место вторичная ионизация и, таким образом, начнется увеличение тока. Эти два этапа показаны на рис. 5.4, б, в, соответственно.

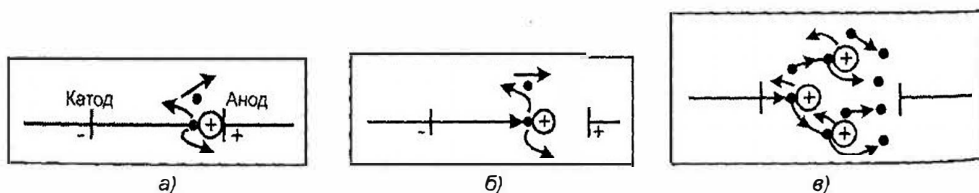


Рис. 5.4. Схемы столкновений в процессах ионизации

Рассмотренный механизм носит название электронной лавины Таунсенда.

Пусть J_e — плотность электронного тока на катоде и α — коэффициент первичной ионизации Таунсенда, т. е. число новых электронов, создаваемых на единице длины траектории первичного электрона.

Пусть n_e — число электронов, покидающих катод за 1 с, а n — число электронов, пересекающих за это же время сечение разрядной трубки в точке x (рис. 5.5).

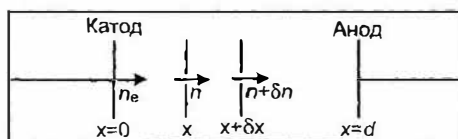


Рис. 5.5. Схема процесса ионизации в промежутке "катод—анод"

На последующем малом интервале δx каждый электрон создаст $\alpha \cdot \delta x$ новых электронов. Их число запишется в виде

$$\frac{\delta n}{n} = \alpha \delta x.$$

Интегрирование этого уравнения даст

$$\int_{n_0}^n \frac{dn}{n} = \int_0^x \alpha dx.$$

Откуда

$$\ln(n / n_0) = \alpha x$$

или

$$n = n_0 \cdot \exp(\alpha \delta x).$$

На аноде ($x = d$) число электронов равно

$$n_a = n_0 \cdot \exp(\alpha d). \quad (5.2.1)$$

Пусть n_i — число положительных ионов, достигших катода. Тогда оно определяется следующим выражением

$$n_i = n_a - n_0 = n_0 \cdot \exp(\alpha d) - n_0 = n_0 \cdot (\exp(\alpha d) - 1).$$

Вообще говоря, электронный ток $I = qn$, а плотность тока

$$J = qn / A,$$

где A — площадь поперечного сечения разрядной трубки.

Следовательно, мы можем записать плотность тока положительных ионов на катоде в виде

$$J_i = I_i \cdot (\exp(\alpha d) - 1).$$

Пусть γ — коэффициент вторичной эмиссии, обусловленной бомбардировкой катода положительными ионами, равный числу электронов, вылетающих из катода под действием одного положительного иона. Число ионов, достигающих катода, есть $n_i = n_a - n_0$.

Они создают $\gamma \cdot (n_a - n_0)$ вторичных электронов за 1 с.

При лавинном процессе $n_a = n_0 \cdot \exp(\alpha d)$.

Поэтому общее число электронов, эмитированных в секунду, определяется выражением

$$n_e = n_0 + \gamma \cdot n_i \cdot (\exp(\alpha d) - 1),$$

$$n_e \cdot [1 - \gamma \cdot (\exp(\alpha d) - 1)] = n_0.$$

Отсюда находим

$$n_e = \frac{n_0}{1 - \gamma \cdot (\exp(\alpha d) - 1)},$$

$$n_a = \frac{n_0 \exp(\alpha d)}{1 - \gamma \cdot (\exp(\alpha d) - 1)}.$$

Следовательно, анодный ток

$$I_a = \frac{I_0 \exp \alpha d}{1 - \gamma \cdot (\exp \alpha d - 1)}, \quad (5.2.2)$$

где I_0 — ток в области насыщения (рис. 5.4). Уравнение для плотности анодного тока имеет вид

$$J_a = \frac{J_0 \exp \alpha d}{1 - \gamma \cdot (\exp \alpha d - 1)}. \quad (5.2.3)$$

Когда $1 - \gamma \cdot (\exp(\alpha d) - 1) \geq 0$, то $J_a \rightarrow \infty$. Для этого должно выполняться условие

$$\gamma \cdot (\exp \alpha d - 1) \geq 1,$$

$$\exp \alpha d = \frac{1 + \gamma}{\gamma},$$

откуда получаем условие перехода несамостоятельного разряда в самостоятельный:

$$\alpha d = \ln \frac{1 + \gamma}{\gamma}.$$

5.3. Механизмы ионизации и кривая Пашена

Перечислите процессы, вызывающие ионизацию в газах, и укажите, какое воздействие оказывает каждый из этих процессов на величину потенциала зажигания в разрядной трубке с двумя большими плоскими электродами, содержащей газ при низком давлении.

Рассмотрите роль электродов, а также укажите, какими свойствами должен обладать материал электродов, для того чтобы потенциал зажигания был максимальным или, наоборот, минимальным.

Начертите типичную кривую Пашена для рассматриваемого газа и объясните ее поведение.

Решение

Ионизация — это процесс, при котором один или несколько электронов либо высвобождаются из атомов или молекул нейтрального газа, либо присоединяются к ним.

Когда частица (атом или молекула) нейтрального газа приобретает один или несколько электронов, она становится отрицательным ионом. При этом число свободных электронов в газе уменьшается, и поскольку подвижность образовавшихся отрицательных ионов много меньше подвижности свободных электронов, электропроводность газа понижается. Отрицательные ионы могут образоваться таким способом не во всех газах. Так, в частности, в водороде, азоте и инертных газах отрицательные ионы таким путем не образуются. С другой стороны, кислород, хлор и пары воды имеют высокое электронное сродство, и в них легко образуются отрицательные ионы. Незначительные следы этих последних трех газов сильно влияют на электропроводность. При низких давлениях вероятность образования отрицательных ионов очень мала.

Молекула или атом газа может потерять электрон и стать, таким образом, положительным ионом с помощью одного из следующих трех процессов:

- ☐ при неупругом соударении с электроном, ионом или атомом нейтрального газа;
- ☐ при соударении с передачей энергии от одного атома к другому;
- ☐ под действием электромагнитного излучения (этот процесс называют фотоионизацией).

Неупругие соударения. Вероятность ионизации, происходящей в результате неупругих соударений с другим атомом, чрезвычайно мала; и ее можно не учитывать, за исключением области очень высоких температур (в электродуговом разряде). Неупругое соударение между электроном и атомом газа или молекулой приведет к образованию иона только в том случае, если кинетическая энергия электрона больше потенциала ионизации (т. е. энергии, необходимой для ионизации газа) атома или молекулы.

Если суммарная энергия сталкивающихся частиц меньше потенциала ионизации одного из сталкивающихся атомов, то произойдет возбуждение этого атома. Показано, что средняя длина свободного пробега электрона обратно пропорциональна давлению газа. Поэтому при низком давлении газа средняя длина свободного пробега велика, т. е. очень мала вероятность столкновений, приводящих к ионизации. Однако вероятность таких столкновений, приводящих к ионизации, сильно зависит от скорости электрона (или потенциала анода) и для определенной области скоростей она имеет максимум.

Передача энергии. Этот процесс обычно наблюдается в смешанных газах, например, в аргоне и неоне. Атом аргона имеет потенциал ионизации 15,7 эВ, а первый потенциал возбуждения у атома неона равен 16,5 эВ (энергии образования первого метастабильного состояния). Таким образом, атом неона может передавать энергию атому аргона, и в результате ионизации последнего образуется ион и свободный электрон и выделяется небольшое количество энергии.

Фотоионизация. Этот процесс играет весьма существенную роль в газовых разрядах, особенно в газах низкого давления. Под действием сильного электромагнитного излучения с энергией кванта больше потенциала ионизации газа ($h\nu > V_i$) образуется большое число свободных электронов и положительных ионов. Это, в свою очередь, приводит к уменьшению потенциала пробоя газа между электродами.

На рис. 5.6 показана характерная зависимость степени ионизации α/p от давления и расстояния между электродами. По оси ординат отложено число ионов α , возникающих на длине 1 м при давлении $p = 1$ мм рт. ст., по оси абсцисс — величина напряженности электрического поля на единицу давления.

Степень ионизации определяется формулой

$$\frac{\alpha}{p} = f_i\left(\frac{V_a}{pd}\right),$$

а для потенциала пробоя (зажигания) справедлив закон Пашена:

$$V_s = f \cdot (pd).$$

Поэтому потенциал зажигания разряда в газе между электродами непосредственно зависит от давления газа.

Ниже потенциала зажигания V_s более важную роль в процессах ионизации играют свойства катода. Если катод имеет высокую фотоэмиссионную способность, то значение V_s понизится; если же катод не обладает фотоэмиссионными свойствами, потенциал V_s будет выше.

С другой стороны, электроды с высокой фотоэмиссионной способностью, например изготовленные из щелочных металлов, имеют низкие коэффициенты вторичной эмиссии, необходимые для поддержания разряда.

Электроды должны обладать способностью рассеивать большое количество тепла, которое выделяется вследствие протекания больших токов после зажигания разряда.

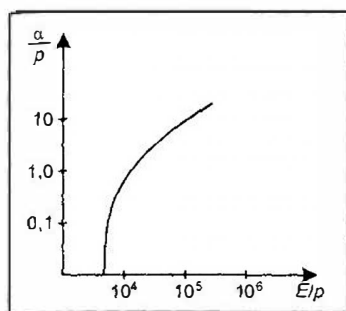


Рис. 5.6. Зависимость степени ионизации от давления

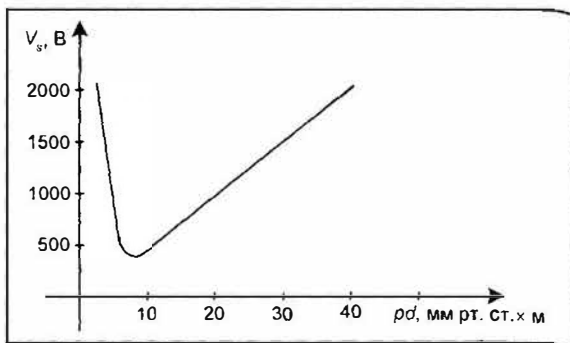


Рис. 5.7. Кривая Пашена

Типичная кривая Пашена для воздуха показана на рис. 5.7. Таким образом, напряжение пробоя в газе V_s зависит только от произведения давления на расстояние между электродами pd .

5.4. Термическая ионизация газа

Определить концентрацию свободных электронов в изотермической плазме, возбужденной в ксеноне, если температура плазмы $T = 6500$ К, а давление газа составляет $p = 1,96 \times 10^6$ Н/м².

Решение

1. Воспользуемся уравнением Саха:

$$\frac{\alpha^2}{1 - \alpha^2} \approx \frac{2}{g_2} \left(\frac{2\pi m}{h^2} \right)^{3/2} \frac{(kT)^{5/2}}{p} \exp\left(-\frac{qU_i}{kT}\right),$$

где α — доля ионизированных атомов; g_1 и g_2 — кратность вырождения иона и нейтрального атома. Для простоты будем считать, что $g_1 = g_2$, h — постоянная Планка. Поскольку ожидается, что $\alpha^2 \ll 1$, можно записать

$$\begin{aligned} \alpha &= \sqrt{2} \left(\frac{2\pi m}{h} \right)^{3/4} \frac{(kT)^{5/4}}{p^{1/2}} \exp\left(-\frac{eU_i}{kT}\right) = \\ &= \sqrt{2} \left(\frac{2 \cdot 3,14 \cdot 9 \cdot 10^{-31}}{4,4 \cdot 10^{-67}} \right)^{3/4} \frac{(1,38 \cdot 10^{-23} \cdot 6,5 \cdot 10^3)^{5/4}}{(1,96 \cdot 10^6)^{1/2}} \times \\ &\times \exp\left[-\frac{1,6 \cdot 10^{-19} \cdot 12,1}{2 \cdot 1,38 \cdot 10^{-23} \cdot 6,5 \cdot 10^3}\right] = 2,2 \cdot 10^{-4}. \end{aligned}$$

2. Определяем абсолютное значение концентрации свободных электронов:

$$n_e = \alpha n_a - \alpha \frac{p}{kT} = \frac{2,2 \cdot 10^{-4} \cdot 1,96 \cdot 10^4}{1,38 \cdot 10^{-22} \cdot 6,5 \cdot 10^3} \approx 4,98 \cdot 10^{21} \text{ м}^{-3}.$$

5.5. Высокочастотный разряд в газе

К разрядному промежутку с плоскопараллельными электродами прикладывается высокочастотное напряжение с амплитудой порядка 500 В. Оценить частоту, начиная с которой разряд можно считать высокочастотным.

Газ — водород, давление газа 5 мм рт. ст., ширина разрядного промежутка $d = 1,5$ см.

Решение

Специфика высокочастотного разряда заключается в том, что электроны, возникающие в межэлектродном промежутке, совершают под действием высокочастотного поля колебательные движения, амплитуда которых значительно меньше межэлектродного расстояния. Поэтому электроны, создавая лавинную ионизацию, не уводятся полем на электроды, а накапливаются в разрядном промежутке.

Критерием высокочастотного разряда можно считать условие $A < d/2$, где A — амплитуда колебаний электронов.

1. Вычисляем величину A . Движение электронов под действием электрического высокочастотного поля считаем дрейфовым

$$v = \mu_e E_m \sin \omega t,$$

$$A = \frac{\mu_e E_m}{\omega} = \frac{\mu_{e0} U_m}{2\pi f p d}.$$

2. Предельная частота f_{np} , при которой перестает выполняться вышеприведенный критерий, равна

$$f_{np} = \frac{\mu_{e0} U_m}{\pi p d^2} = \frac{37 \cdot 500}{3,14 \cdot 5 \cdot 2,25 \cdot 10^{-4}} = 5,2 \times 10^6 \text{ Гц}.$$

Теперь проверим, будет ли движение электронов дрейфовым. Условием дрейфового характера движения электронов является неравенство $\omega \ll v$.

В нашем случае $\omega = 3,3 \times 10^7 \text{ с}^{-1}$.

$$v = v_0 p = 6 \cdot 10^9 \cdot 5 = 3 \cdot 10^{10} \text{ с}^{-1}.$$

Движение электронов действительно будет дрейфовым.

5.6. Условие пробоя в газе и потенциал зажигания разряда

Дайте определение коэффициентов первичной α и вторичной γ ионизации Таунсенда для разряда в газе и покажите, что если пренебречь рекомбинацией между электронами и ионами, то условие пробоя в газе определяется выражением $\gamma [\exp(\alpha d) - 1] \geq 1$. С помощью этого выражения или другим способом покажите, что для данного газа выполняется закон Пашена, т. е. потенциал зажигания пропорционален произведению $p \cdot d$, где p — давление газа, а d — расстояние между электродами. Отметьте и объясните все основные отклонения от этого закона.

Для камеры, наполненной аргоном при низком давлении и предназначенной для катодного распыления, определите максимальное расстояние между электродами, при котором можно избежать электрического разряда.

Известно, что для аргона в рассматриваемых условиях коэффициенты Таунсенда равны $\alpha = 150$ и $\gamma = 2$.

Решение

Коэффициент первичной ионизации Таунсенда α можно определить как число новых электронов, возникающих на единице длины вследствие столкновений, т. е. каждый первичный электрон на единице длины своей траектории создает в газе α новых электронов. Коэффициент вторичной ионизации Таунсенда γ можно определить как число электронов, выбиваемых из катода каждым положительным ионом, попадающим на катод.

Заметим, что α имеет размерность м^{-1} , тогда как γ — безразмерная величина. Условие пробоя в газе имеет вид

$$\gamma[\exp(\alpha d) - 1] \geq 1. \quad (5.6.1)$$

Закон Пашена (см. задачу 5.3) гласит, что потенциал зажигания разряда в газе является функцией произведения расстояния между электродами и давления газа.

Существенное отклонение от этого закона наблюдается для значений $p \cdot d < 0,005$ мм рт. ст., при которых потенциал зажигания начинает резко увеличиваться. Причина этого отклонения состоит в том, что средняя длина свободного пробега электрона становится сравнимой с межэлектродным расстоянием, вследствие чего уменьшается вероятность столкновений.

Пусть коэффициент умножения

$$\beta = \frac{I}{I_0} = \frac{\exp(\alpha d)}{1 - \gamma[\exp(\alpha d) - 1]},$$

откуда находим, что

$$\exp(\alpha d) = \frac{\beta(1 + \gamma)}{1 + \beta\gamma} = \frac{1 + \gamma}{(1/\beta) + \gamma}.$$

При $\beta \rightarrow \infty$ имеем

$$\exp(\alpha d) = \frac{1 + \gamma}{\gamma}. \quad (5.6.2)$$

Следовательно, при $\beta \rightarrow \infty$ расстояние между электродами определяется выражением

$$d = \frac{1}{\alpha} \ln \frac{1 + \gamma}{\gamma}.$$

Подставляя сюда исходные данные, получаем значение максимально допустимого расстояния между электродами:

$$d = 1 / 150 \cdot \ln 3 / 2 = 2,7 \text{ мм.}$$

Заметим, что выражение (5.6.2) может быть также получено непосредственно из условий пробоя.

5.7. Дебаевский радиус экранирования

Найти выражение для дебаевского радиуса экранирования r_D в случаях изотермической и неизотермической плазмы. Оценить r_D в неизотермической плазме при $T = 3 \times 10^2$ К и концентрации зарядов $n = 1 \times 10^{16} \text{ см}^{-3}$. Сколько заряженных частиц содержится в объеме дебаевской сферы? Рассмотрение провести в условиях слабого электрического поля и бoльшмановского распределения плотности плазмы.

решение

В неизотермической плазме ($T_e \gg T_i$)

$$n_{ie} = n \cdot \exp(q\phi/kT_e),$$

$$n_{ii} = n \cdot \exp(q\phi/kT_i).$$

При условии, что $q\phi \ll kT_e, kT_i$, имеем

$$n_{ii} - n_{ie} = -qn/k[(T_e + T_i)/T_e T_i].$$

Тогда дебаевский радиус экранирования в соответствии с (5.10)

$$r_D = (k/4\pi \cdot q^2 n)^{1/2} \cdot [T_e \cdot T_i / (T_e + T_i)]^{1/2} \approx (kT_i / 4\pi \cdot q^2 n)^{1/2} = 6,9 \cdot (T_i / n)^{1/2} = 1,2 \times 10^{-3} \text{ см.}$$

В изотермической плазме

$$T_e = T_i = T,$$

и тогда

$$r_D = (k/8\pi \cdot q^2 n)^{1/2}.$$

Число частиц, содержащихся в объеме дебаевской сферы, равно

$$N = 4/3 \cdot \pi \cdot r_D^3 \cdot n \approx 70.$$

Контрольные вопросы

1. Что такое плазменная электроника?
2. Какие типы газового разряда вы знаете?
3. Что такое плазма?
4. Что такое степень ионизации плазмы?
5. Какими свойствами отличается плазма от нейтрального газа?
6. В чем физический смысл дебаевского радиуса экранирования?
7. Что такое идеальная плазма?
8. Какие силы действуют на частицу, дрейфующую в плазме?
9. Что такое тормозное излучение плазмы?
10. Что представляет собой синхротронное излучение плазмы?
11. Какие параметры плазмы определяются в процессе диагностики?
12. Что представляют собой пассивные и активные методы измерения параметров плазмы?

Рекомендуемая литература

1. Арцимович Л. А. Элементарная физика плазмы. — М.: Атомиздат, 1969.
2. Левитский С. М. Сборник задач и расчетов по физической электронике. — Изд. Киевского университета, 1964.
3. Линч П., Николаидес А. Задачи по физической электронике. Под ред. проф. Г. В. Скоцкого. — М.: Мир, 1975.
4. Основы физики плазмы. В 2 т. Под ред. А. А. Галесва, Р. Судана. — М.: Энергоатомиздат, 1984.
5. Чен Ф. Введение в физику плазмы. Пер. с англ. — М.: Мир, 1987.

6. Приборы и устройства плазменной электроники

6.1. Ионные приборы

Ионные или газоразрядные приборы представляют собой класс электровакуумных приборов, которые наполнены каким-либо инертным газом, парами ртути или водородом, и действие которых основано на прохождении электрического тока через образованную в межэлектродном пространстве газоразрядную плазму.

По типу газового разряда, зажигающегося в приборе, по природе электронной эмиссии, по роду газа и его плотности различают ионные приборы самостоятельного дугового разряда, самостоятельного дугового разряда, тлеющего разряда, искрового разряда.

Основными носителями тока являются электроны. Их подвижность значительно больше подвижности ионов, и поэтому они играют решающую роль. Роль же ионов в газовом разряде сводится к компенсации объемного заряда электронов, с помощью которой обеспечиваются высокие токи в ионных приборах.

Простейшим ионным прибором является диод с накаливаемым или холодным катодом. Стеклый или керамический баллон заполняется инертным газом или парами ртути. При подаче на электроды напряжения электроны ионизируют газ, образуя газовый разряд. Пространственный заряд электронов компенсируется положительными ионами. Такой ионный диод в отличие от вакуумного имеет малое внутреннее сопротивление и способен пропускать токи $\sim 10^3$ — 10^4 А. В таких диодах может быть использована плазма дугового тока.

Двухэлектродный неуправляемый ионный прибор получил название *газотрон*. В нем используется самостоятельный дуговой или тлеющий разряд.

Тиратрон представляет собой газоразрядный прибор с сеточным управлением моментом зажигания самостоятельного дугового или тлеющего разряда.

В тиратроне используются как накаливаемые, так и холодные катоды. На рис. 6.1 представлена конструкция тиратрона (а) и характеристика его зажигания (б).

В тиратроне дугового разряда разряд зажигается при определенном соотношении напряжения на аноде U_a от напряжения на сетке U_c .

По способу управления анодным током различают тиратроны с отрицательной и положительной характеристиками зажигания.

Тиратроны с отрицательной характеристикой используются в схемах выпрямления или релейных устройствах (рис. 6.1, б). Тиратроны с положительной характеристикой зажигания чаще используются в импульсных схемах.

Тиратроны относятся к приборам силовой электроники и не являются предметом подробного рассмотрения. Тиратроны выпускаются в стеклянном, металлостеклянном, а также в металлокерамическом исполнении.

К газоразрядным приборам самостоятельного разряда относятся *ионитроны*, *ртутные вентили*, *экситроны*, *аркотроны*. В этих приборах используется дуговой разряд. Это

мощные приборы силовой электроники, которые применяются в электросварочной и коммутационной аппаратуре.

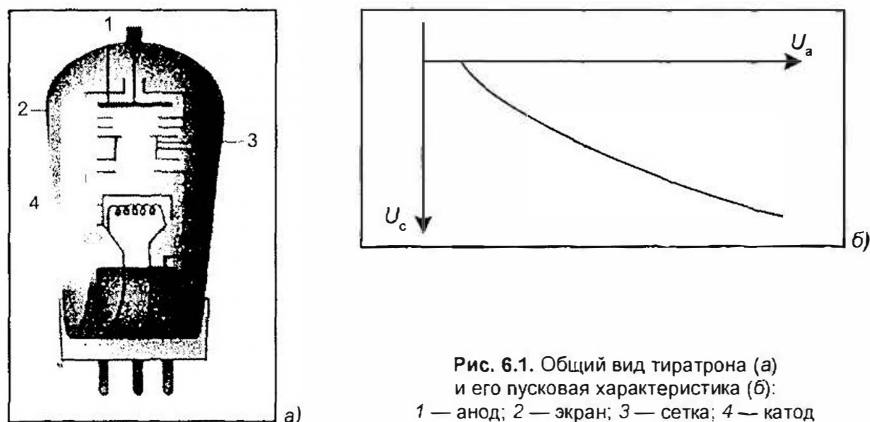


Рис. 6.1. Общий вид тиратрона (а) и его пусковая характеристика (б): 1 — анод; 2 — экран; 3 — сетка; 4 — катод

Тлеющий разряд используется в газоразрядных приборах типа *декатронов*, *стабилизаторов*. *Декатрон* представляет собой многоэлектронный газоразрядный прибор тлеющего разряда, предназначенный для индикации электрических сигналов. Его действие основано на направленном переносе тлеющего разряда с одного электрода на другой с целью формирования индикаторного знака. С появлением полупроводниковых приборов интерес к этому типу ионных приборов значительно снизился.

Искровой разряд представляет собой кратковременный дуговой или тлеющий разряд. Обычно это электрическая искра. Приборы искрового разряда применяются для защиты радиоустройств.

Газоразрядных приборов существует более 50 типов. Одним из них является газоразрядные лазеры (см. часть III).

6.2. Ионные приборы обработки и отображения информации

В ионных приборах обработки и отображения информации используется в основном тлеющий разряд. Использование тлеющего разряда в ионных приборах позволяет применять ненакаливаемые (холодные) катоды. В плазме тлеющего разряда возникает УФ-излучение, под действием которого светятся люминофоры различных цветов.

Именно это явление свечения при прохождении электрического тока через возбужденный газ лежит в основе работы ионных, газоразрядных приборов. Свечение связано с процессами возбуждения атомов ударами электронов с последующим возвратом атомов в невозбужденное состояние и одновременным выделением квантов света, либо с процессом рекомбинации положительных ионов с электронами в объеме или на стенках прибора.

Излучение может лежать в видимом или УФ-диапазоне спектра. В этом случае для преобразования излучения в видимое используются фотолюминофоры. К приборам этого типа относятся знаковые индикаторы. Они предназначены для отображения информации в

виде изображений цифр, букв и различных символов. Конструкция знаковых индикаторов состоит из одного или нескольких сетчатых анодов и набора катодов в форме отображаемых символов. Схема коммутации катодов обеспечивает включение нужного катода, соответствующего отображаемой информации. Подбором тока на анод обеспечивается режим равномерного свечения катода, который является индицируемым символом.

Шкальные индикаторы предназначены для отображения как цифровой, так и аналоговой информации. В качестве индикаторного элемента используется газовый промежуток "анод—катод". В зависимости от способа подачи управляющего импульса информация отображается в виде светящегося столбика, либо в виде светящейся точки, перемещающейся вдоль шкалы (рис. 6.2). Число индикаторных элементов не превышает нескольких сотен при шаге дискретности 0,5—1,5 мм.

В качестве индикаторов могут быть использованы тиратроны тлеющего разряда. На основе индикаторных тиратронов созданы различные матричные индикаторы для отображения буквенно-цифровой информации. Некоторые конструкции таких тиратронов легко сопрягаются с микросхемами и поэтому могут оперативно ими управляться.

Однако наибольший интерес вызывают газоразрядные индикаторные панели (ГИП).

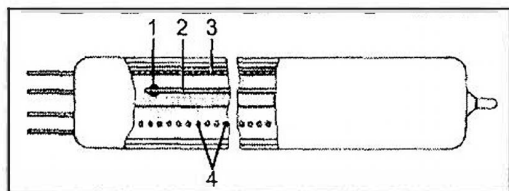


Рис. 6.2. Схема шкального индикатора:

- 1 — керамическая втулка крепления катода по оси;
2 — проволочный катод; 3 — экран со щелью;
4 — индикаторные аноды

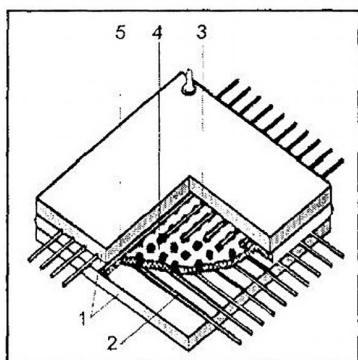


Рис. 6.3. Схема газоразрядной индикаторной панели: 1 — стеклянные пластины;

- 2 — катоды; 3 — аноды; 4 — перфорированная изолирующая пластина; 5 — герметик

Газоразрядная индикаторная панель (ГИП) представляет собой конструкцию, содержащую большое число светоизлучающих элементов отображения информации.

Такие индикаторы обладают большой информационной емкостью. Эти элементы формируют столбы и строки, объединенные в одном корпусе. Светоизлучающие элементы образуются в местах взаимного пересечения систем электродов: анодов и катодов (рис. 6.3).

Число светоизлучающих элементов равно произведению числа анодов на число катодов. При подаче на взаимно пересекающиеся электроды высокого напряжения образуется свечение. Это происходит вследствие собственного излучения ионизированного газа, либо свечения люминофоров, возбуждаемых УФ излучением плазмы разряда. Зазор между стеклянными пластинами заполняется неоном или смесями газов на его основе. Обычно это аргон, криптон или ксенон, которые способствуют понижению напряжения разряда, а также изменению спектральной характеристики с целью получения нужного спектра из-

лучения. Гелий в газовой смеси позволяет ослабить температурную зависимость за счет увеличения теплопроводности смеси. В некоторых смесях используются также пары ртути.

Путем подбора определенного сочетания светящихся точек и модуляции их яркости удается формировать сложные информационные картины.

Конструктивно ГИП делятся на ГИП постоянного тока, ГИП переменного тока, ГИП с плазменно-электронным возбуждением люминофоров.

Коммутация или адресация электродов может быть автономной или встроеной в панель.

Для адресации элементов индикации используются два метода. В методе совпадений локализация и возникновение разряда происходит при одновременном приложении поля к скрещенным электродам ячейки.

В методе зарядовой связи возникновение заряда в заданном промежутке стимулируется разрядом в другом промежутке. Метод зарядовой связи позволяет получить направленный перенос газового разряда.

ГИП с плазменно-электронным возбуждением люминофора используется не только для отображения буквенно-цифровой информации, но и для воспроизведения телевизионных изображений. В этом типе ГИП возбуждается газовый разряд и создается плазменный катод. Электроны плазмы ускоряются системой электродов и возбуждают люминофор, что соответственно повышает световую эффективность приборов этого типа.

Различают ГИП индивидуального, группового и коллективного пользования, отличающиеся разрешающей способностью и размером панели. Так ГИП индивидуального пользования имеют размер $\sim 1 \text{ м}^2$ с разрешением более 20 ячеек на сантиметр. ГИП коллективного пользования имеют размеры до 10 м^2 .

Плазменные панели дают самое высокое качество изображения по сравнению с известными экранами и дисплеями. Угол видимости достигает 160° . Считается, что альтернативы плазменным панелям на сегодняшний день нет. Недостатком ГИП является их довольно низкая экономичность.

Задачи и упражнения

6.1. Расчет газоразрядного стабилизатора напряжения (стабилитрона)

Стабилитрон наполнен аргоном при давлении 10 мм рт. ст. Площадь катода составляет 10 см^2 , материал катода — алюминий.

Чему равно нормальное падение напряжения на стабилитроне? Каков верхний предел рабочего тока стабилитрона?

Решение

Стабилитрон представляет собой газоразрядную лампу, работающую в режиме нормального тлеющего разряда. Наибольший ток, при котором стабилитрон продолжает еще работать в нормальном режиме, получается, когда плотность тока на всей поверхности катода равна нормальной. При меньших токах разряд покрывает лишь поверхность катода, но плотность тока на этой работающей части продолжает оставаться нормальной. Нормальной остается также и величина катодного падения напряжения. Поэтому напряжение на стабилитроне мало меняется при изменении разрядного тока в довольно широких пределах.

1. Вычислим величину катодного напряжения для тлеющего разряда исходя из условия (5.5)

$$U_K = 6B/2A \cdot \ln(1 + 1/\gamma) = 6 \cdot 2,35 \cdot 10^{-4} \cdot 2,3 \cdot \lg(1 + 1/0,11) / 2 \cdot 1360 = 120 \text{ В.}$$

Значения констант взяты из справочника.

Поскольку падение напряжения на анодных частях разряда невелико, можно считать, что все напряжение, прикладываемое к стабилитрону, равно U_K .

2. Вычислим нормальную плотность тока тлеющего разряда. В нормальном режиме

$$C_2 J_{\text{норм}} = 0,67,$$

где

$$C_2 = \ln(1 + \frac{1}{\gamma}) / \epsilon_0 A B^2 \mu p^2 \cdot (1 + \gamma),$$

где μ — подвижность ионов газа при $p = 1$ мм рт. ст.

Тогда

$$J_{\text{норм}} = 0,67 \cdot \epsilon_0 A B^2 \mu p^2 (1 + \gamma) / \ln(1 + 1/\gamma) = \\ = 0,67 \cdot 8,85 \cdot 10^{-12} \cdot 1,36 \cdot 10^3 \cdot 5,5 \cdot 10^8 \cdot 0,16 \cdot 10^{-2} \cdot 1,11 / 2,3 \cdot \lg(1 + 1/0,11) = 34 \text{ А / м}^2.$$

3. Наибольший ток, при котором стабилитрон продолжает еще работать в нормальном режиме, равен произведению нормальной плотности тока на площади катода

$$I_m = J_{\text{норм}} \cdot S_k = 34 \times 10^{-3} = 3,4 \times 10^{-2} \text{ А.}$$

6.2. Коэффициенты первичной ионизации и умножения

В газоразрядной трубке находится газ под давлением. Электроды расположены на расстоянии $d = 5$ мм друг от друга. Трубка помещена в однородное электрическое поле. Пробой наступает при 400 В.

Определить коэффициент первичной ионизации α и коэффициент умножения β при напряжении 200 В и расстоянии между электродами 2,5 мм в том же самом приборе, если коэффициент вторичной эмиссии $\gamma = 0,02$.

Решение

Электрический пробой наступает при выполнении условия

$$\gamma \cdot \exp(\alpha d) = 1$$

или

$$0,02 = \exp(0,5\alpha) = 1,$$

откуда

$$0,5\alpha = \ln 50 \text{ и } \alpha = 7,82 \text{ см.}$$

При изменении условий напряженность электрического поля осталась той же самой и, следовательно, сохраняется значение α .

Коэффициент умножения

$$\beta = \frac{I}{I_0} = \frac{\exp(\alpha d)}{1 - \gamma \cdot \exp(\alpha d)} = \frac{\exp(7,82 \cdot 0,25)}{1 - 0,02 \cdot \exp(7,82 \cdot 0,25 - 1)} = 8,03.$$

Контрольные вопросы

1. Что такое нониусные приборы?
2. Что такое тиратрон и как он устроен?
3. Что такое декатрон? Как он устроен?
4. Что такое газоразрядная панель и как она устроена?
5. Какие типы газоразрядных панелей вы знаете?

Рекомендуемая литература

1. Арцимович Л. А. Элементарная физика плазмы. — М.: Атомиздат, 1969.
2. Левитский С. М. Сборник задач и расчетов по физической электронике. — Изд. Киевского университета, 1964.
3. Терехов В. А. Сборник задач по электронным приборам. — М.: Энергоатомиздат, 1983.
4. Яблонский Ф. М. Газоразрядные приборы для отображения информации. — М.: Энергия, 1979.
5. Яблонский Ф. М., Троицкий Ю. В. Средства отображения информации. — М.: Высшая школа, 1985.

7. Введение в зондовую микроскопию

7.1. Туннельная микроскопия

Вакуумная электроника в настоящее время является одним из наиболее интенсивно развивающихся разделов науки и техники. Существенную роль при этом играют фундаментальные исследования в области физики поверхности твердого тела, разнообразных электронных, атомных и молекулярных процессов, происходящих на поверхности твердых тел.

Заветным желанием ученых, особенно разработчиков электровакуумных приборов, на протяжении многих столетий было непосредственное наблюдение за поведением отдельных атомов на поверхности твердого тела, изучение процессов с участием одиночных или небольших групп атомов.

Методы исследования поверхности твердых тел, основанные на увеличении деталей поверхности с применением оптических микроскопов, относятся к концу XVII столетия и совершенствуются до настоящего времени. Современные оптические микроскопы позволяют увеличить детали поверхности до 10^3 раз и получить двумерное изображение. Оптические методы относятся к неразрушающим методам, поскольку образец не подвергается разрушению. При использовании лазера в качестве источника света удалось достигнуть увеличения до 10^4 раз. Предел разрешения деталей поверхностей в оптических методах определяется в основном длиной волны излучения и составляет порядка 0,5 мкм. Оптические методы не требуют особых условий, и диагностика с их помощью проводится на воздухе или в жидкости.

В середине XX века появились методы, связанные с использованием электронов в качестве источника излучения. Длина волны электронов зависит от ускоряющего напряжения и может достичь нескольких ангстрем. Разработанные на основе электронных потоков сканирующие электронные микроскопы позволяют получить двумерные изображения с увеличением порядка 10^6 . При этом исследуемый образец необходимо помещать в высокий вакуум. К настоящему моменту разработаны и хорошо развиты методы сканирующей электронной микроскопии (SEM, scanning electron microscopy) и просвечивающей электронной микроскопии (TEM, transmission electron microscopy). Они относятся к числу разрушающих методов контроля поверхности.

Для формирования электронного пучка обычно используются электронные пушки на основе термоэлектронных катодов. Разработаны и модели электронных микроскопов с использованием автокатодов в качестве источника электронов. Разрешающая способность растровых и просвечивающих микроскопов позволяет наблюдать молекулярную структуру объектов исследуемой поверхности. Однако образец в процессе исследований разрушается, и, следовательно, результаты искажаются.

В 1936 году немецким физиком Эрвинном Вильгельмом Мюллером был изобретен *электронный проектор* (или автоэлектронный микроскоп), позволяющий получить разре-

ние на атомном уровне (25 Å). Лучшее разрешение получить не удавалось из-за эффекта близости электронов и их дифракции. Улучшение разрешения до 3—5 Å удалось получить в *ионном микроскопе* (ионном проекторе), в котором для увеличения изображения поверхности исследуемого объекта используется пучок ионов. Следует заметить, что такое высокое разрешение можно получить только с вершины металлического острия в условиях сверхвысокого вакуума. Этот прибор не может быть использован для исследования достаточно протяженной поверхности.

В 80-е годы прошлого столетия был разработан мощный метод топографии и анализа поверхности твердого тела — *сканирующая туннельная микроскопия*. Идея сканирующего туннельного микроскопа родилась в 1978 году. В научной дискуссии между сотрудником фирмы IBM Г. Рорером и аспирантом Франкфуртского университета Г. Биннингом последний высказал мысль, что вакуумное туннелирование может быть подходящим методом для исследования поверхности. Для осуществления своей идеи он пригласил Рорера в лабораторию фирмы IBM в Цюрих.

Исследования показали, что туннельный ток с кончика острия ограничен областью значительно меньшей по сравнению с площадью проекции кончика острия. Этот экспериментальный факт лег в основу работы предложенного ими туннельного микроскопа. Весной 1981 года был построен прибор с разрешающей способностью 2 Å. Прибор позволял также разрешать детали структуры вблизи отдельного атома. Аналогов ему не было. В 1986 году Г. Биннинг и Г. Рорер были удостоены Нобелевской премии по физике за создание сканирующего туннельного микроскопа, с помощью которого удалось получить разрешение 0,02 Å по нормали к исследуемой поверхности и 2 Å в плоскости образца. Одним из достоинств этого метода является то, что в процессе исследования образец не подвергается разрушению.

7.2. Физические явления на границе "острие—поверхность"

7.2.1. Туннельный ток

Туннельный эффект представляет собой физическое явление проникновения микрочастиц сквозь энергетические барьеры, величина которых превосходит собственную энергию микрочастиц.

С помощью туннельного эффекта объясняются такие физические явления, как холодная эмиссия электронов из металлов в сильном электрическом поле, распад радиоактивных ядер, протекание электрического тока через тонкие диэлектрики, туннельные токи в сверхпроводниках и многое другое.

В настоящее время явление туннелирования электронов широко используется в спектроскопии элементарных возбуждений металлов, в биофизике, химии, экспериментальной физике.

Вакуумное туннелирование — явление прохождения тока через контакт двух проводников, разделенных вакуумным зазором.

Обычно это — металлический зонд в виде остро заточенной иглы и поверхность исследуемого вещества (рис. 7.1). В первых экспериментах использовались вольфрамовые иглы и платиновые пластины. В дальнейшем круг исследуемых поверхностей существенно расширился.

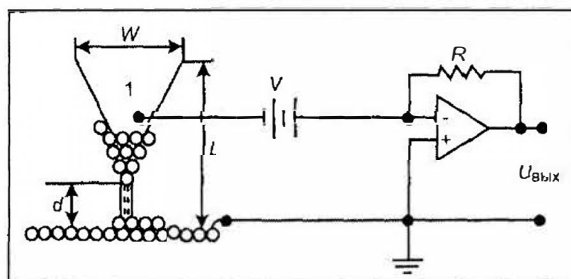


Рис. 7.1. Исследование туннельного тока в промежутке "игла—поверхность":
1 — игла; 2 — поверхность;
 W , L , d — геометрические параметры зонда;
 R — резистор операционного усилителя

Коэффициент прохождения D частиц массой M с энергией E через потенциальный барьер высотой Φ и шириной d определяется соотношением:

$$D = 4k2\eta2 / (k2 - \eta2) \cdot \sinh 2l / \eta + 4k2\eta2, \quad (7.1)$$

где $\eta = 1/\hbar \cdot \sqrt{2M(\Phi - E)}$, $k = 1/\hbar k \cdot \sqrt{2ME}$.

Туннельный ток через вакуумный зазор шириной d определяется следующей зависимостью:

$$I \approx \exp(-A\sqrt{\Phi}d), \quad (7.2)$$

где $A = \sqrt{\frac{8M}{\hbar^2}} = 1,025 \text{ эВ}^{1/2} \cdot \text{А}^{-1}$, Φ — средняя работа выхода веществ острия и поверхности. Из формулы (7.1) следует, что изменение величины D на несколько ангстрем, т. е. на величину моноатома, приводит к изменению туннельного тока на три порядка. Именно такая резкая зависимость туннельного тока от величины вакуумного зазора позволяет получить сверхвысокое разрешение. При этом величина напряжения на зазоре поддерживается постоянной.

На рис. 7.2 приведена зонная диаграмма туннельного контакта двух проводников и огибающие волновые функции электронов в металле и барьере в приближении эффективной массы.

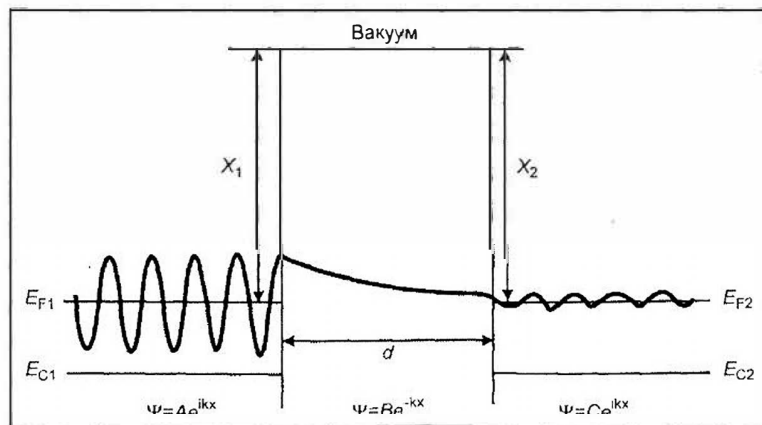


Рис. 7.2. Зонная диаграмма туннельного контакта двух проводников

В потенциальном барьере волновая функция электронов затухает экспоненциально и туннельное сопротивление R зазора оказывается очень чувствительно к изменению d :

$$R \approx R_0 \exp(1,025 d \Phi^{1/2}), \quad (7.3)$$

где значение величины $\Phi = \frac{(X_1 + X_2)}{2}$ является усредненной работой выхода, а напряжение на зазоре $V \ll \Phi/e$.

На рис. 7.3 приведена зависимость величины туннельного тока от величины зазора. Величина зазора сравнима с величиной межмолекулярных расстояний и длиной пробега атомов в воздухе. Поэтому можно считать, что туннельный ток в зазоре протекает как в условиях вакуума. Напряженность электрического поля в зазоре достигает значений $\sim 10^6$ В/см.

Если с помощью обратной связи туннельный ток, а следовательно, сопротивление R поддерживать постоянным, то напряжение будет меняться в соответствии с рельефом поверхности. В этом случае можно получить информацию о топографии поверхности при условии постоянства работы выхода.

Разрешающая способность индикации поверхности зондом с радиусом закругления r определяется соотношением:

$$\delta \approx 3\sqrt{2r/(A\Phi^{-1/2})} \approx 3\sqrt{r}, \quad (7.4)$$

где r измеряется в ангстремах.

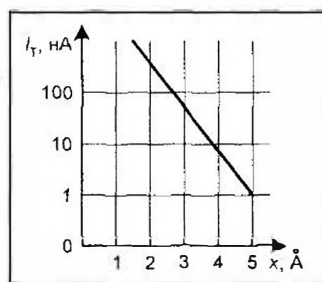


Рис. 7.3. Зависимость туннельного тока от величины зазора

Заметим, что для получения высокого разрешения необходимо, чтобы кончик иглы состоял из 1–2 атомов. Ясно, что высокое разрешение δ можно получить при использовании острия с радиусом закругления $r \leq 100$ Å. Это является типичным размером игл для эмиссионных микроскопов.

Существует два основных режима исследования поверхности: режим постоянного тока и режим постоянного положения пьезотрубки.

В режиме постоянного тока $I = \text{const}$, т. е. величина туннельного тока поддерживается постоянной в процессе сканирования поверхности. В этом случае топография исследуемой поверхности будет пропорциональна амплитуде напряжения сигнала подаваемого на пьезодатчик (рис. 7.4, а).

В этом случае высота неоднородности на поверхности будет пропорциональна изменению положения зонда при условии поддержания постоянного значения туннельного тока.

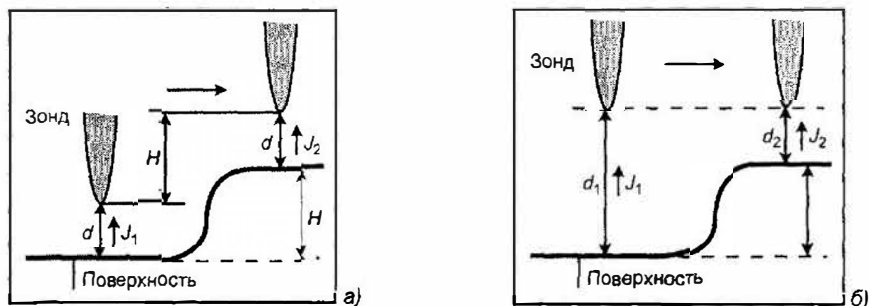


Рис. 7.4. Методы исследования поверхности: режим постоянного тока (а); режим постоянного неподвижного зонда (б)

В режиме постоянного положения пьезотрубки $Z \approx \text{const}$ (рис. 7.4, б). В этом случае происходит регистрация туннельного тока, величина которого $J_1 - J_2$ пропорциональна неоднородности исследуемой поверхности ΔH .

Если первый метод позволяет получить точную информацию о рельефе поверхности, то второй метод позволяет получить изображение в реальном масштабе времени. Периодическое изменение параметров вызывает осцилляции туннельного тока. Эта идея легла в основу модуляционных методик исследования поверхности.

7.2.2. Силы атомных связей

Между атомами вещества, а также между любыми сближающимися телами действуют силы межмолекулярного взаимодействия. На рис. 7.5 показаны характерные зависимости сил межмолекулярного взаимодействия F (а) и потенциала взаимодействия (б). Пунктиром показана зависимость силы Ван-дер-Ваальса от расстояния.

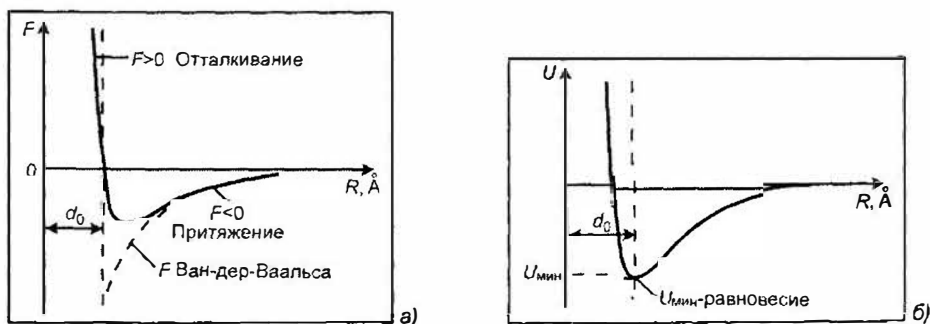


Рис. 7.5. Зависимость межмолекулярных сил взаимодействия (а) и потенциала взаимодействия (б) от расстояния

Реальная сила обращается в нуль при $R = d_0$, где d_0 — расстояние между центрами взаимодействующих молекул в положении равновесия. Другими словами, в положении равновесия силы межмолекулярного взаимодействия равны нулю, а потенциальная энергия принимает минимальное значение.

Потенциал взаимодействия достаточно хорошо описывается формулой Леннарда-Джонса

$$U(R) = 4U \left[\left(\frac{d_0}{R} \right)^{12} - \left(\frac{d_0}{R} \right)^6 \right], \quad (7.5)$$

где параметр U имеет размерность энергии.

На больших расстояниях $R \geq d_0$ происходит притяжение между молекулами с силой пропорциональной R^{-7} , а на малых расстояниях наблюдается отталкивание с силой пропорциональной R^{-13} .

Силы атомных связей действуют также между зондом — острием и исследуемой поверхностью. При приближении зонда к поверхности образца на него начинают действовать силы Ван-дер-ваальсового притяжения. Эта сила достаточно дальнедействующая и ощущается уже на расстоянии десятков ангстрем. Во влажном воздухе на поверхности образца присутствуют молекулы воды, и капиллярные силы увеличивают силу притяжения.

При дальнейшем приближении зонда к образцу возникает процесс перекрытия электронных оболочек атомов зонда и атомов поверхности исследуемого образца. Это приводит к появлению сил отталкивания. При небольших расстояниях отталкивающая сила становится доминирующей.

На поверхности образца всегда имеется слой адсорбированных атомов, состоящий из воды, компонентов воздуха, а также веществ, с которыми образец контактировал или из которых изготавливался. Обычно толщина такого слоя составляет 2—50 нм.

При соприкосновении зонда с адсорбированным слоем возникают силы притяжения капиллярной природы. Эффект капиллярного притяжения сильно сказывается при отведении зонда от образца. Случается, что сила атомного взаимодействия иглы и образца может быть меньше при его приближении, чем при удалении.

Для типичных значений геометрических размеров иглы сила Ван-дер-Ваальса определяется как

$$F_{\text{вдв}} = 10^{-19} \cdot R d^{-2} \approx 10^{-9} - 10^{-8} \text{ Н},$$

а сила капиллярного притяжения

$$F_{\text{км}} = 4\pi R \cdot \cos\theta / (1 + d/D) \approx 10^{-8} \text{ Н},$$

где γ — поверхностное натяжение пленки адсорбата, θ — контактный угол для зонда и образца, D — глубина погружения зонда в образец. Следовательно, силы взаимодействия образца с иглой по порядку величин совпадают с силами межатомного взаимодействия.

Для различных материалов зонда и поверхности различаются и константы адсорбции, что существенно влияет на характер их взаимодействия.

7.3. Системный подход к созданию зондовых сканирующих микроскопов

В зависимости от используемого физического эффекта взаимодействия зонда — иглы с исследуемой поверхностью различают *сканирующие туннельные* микроскопы (СТМ) и *атомно-силовые* микроскопы (АСМ). Также развиваются и другие конструкции, например, *близкопольные* сканирующие микроскопы. Все они объединены общим названием — *сканирующие зондовые микроскопы*.

Таким образом, сканирующий зондовый микроскоп (СЗМ) представляет собой прибор, который предоставляет возможность исследовать свойства поверхности в субмикронном диапазоне и на атомарном уровне.

Известные конструкции сканирующих зондовых микроскопов состоят из следующих основных элементов:

- сенсор;
- пьезоэлектрические двигатели и (или) датчики;
- устройство управления — электронная цепь обратной связи;
- устройство обработки и индикации информации.

Приведем краткую характеристику этих элементов.

7.3.1. Сенсоры

Основными типами сенсоров являются туннельный и атомно-силовой.

Туннельный сенсор представляет собой зонд — металлическое острие с малым радиусом закругления на основе одиночных атомов или даже с последним атомом на острие (рис. 7.6).

Когда такое острие подводится к поверхности на расстояние 10 \AA , то при приложении между острием и образцом небольшого напряжения смещения V_s ($0,01$ — 10 В), через вакуумный промежуток d начинает протекать туннельный ток I , порядка 10^{-9} нА . Электроны из образца туннелируют через туннельный промежуток в иглу или, наоборот, в зависимости от знака приложенного напряжения смещения. Полагая, что электронные состояния (орбитали) локализованы на каждом атомном участке, при сканировании поверхности образца в направлении x или y с одновременным измерением выходного сигнала в цепи z можно получить картину поверхностной структуры на атомном уровне. Эта структура может быть отображена в двух режимах:

- измеряя туннельный ток и поддерживая расстояние d от острия до поверхности образца;
- измеряя изменения в положении острия (т. е. расстояние до поверхности образца) при постоянном туннельном токе (этот режим используется чаще).

Таким образом, туннельный ток используется как механизм для получения картины исследуемой поверхности. Для его возникновения необходимо, чтобы образец и игла были проводниками, либо полупроводниками.

В качестве микроострия могут выступать отдельные атомы на зонде (атомно-силовой сенсор). В связи с этим такой сенсор является идеальным зондом для изучения рельефа поверхности в атомарном масштабе.

В атомно-силовом микроскопе используется механический зонд *кантилевер* — V - или L -образная консоль, на конце которой укреплено пирамидальное острие. Кантилеверы производятся из кремния (жесткие) или из нитрида кремния (мягкие).

В процессе перемещения кантилевера вдоль поверхности на него действуют силы, которые отклоняют кантилевер от положения равновесия (рис. 7.7). В соответствии с законом Гука соотношение между действующей на кантилевер силой F_u и отклонением x задается следующим выражением:

$$F_u = -kx,$$

где k — коэффициент жесткости ($k \sim 1 \text{ Н/м}$).

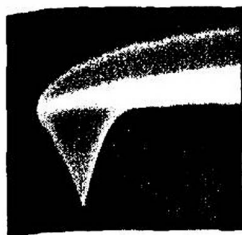


Рис. 7.6. Игла на свободном конце кантилевера

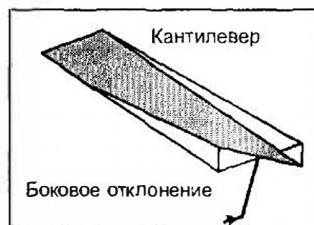


Рис. 7.7. Схема деформации кантилевера в процессе сканирования

Для регистрации малых механических перемещений используется оптический датчик смещений, состоящий из полупроводникового лазера и квадрантного (четырёхсекционный) фотодиода (рис. 7.8). Излучение полупроводникового лазера с длиной волны 650—670 нм фокусируется объективом в пятно диаметром ~ 5 мкм на отражающей поверхности кантилевера. Отраженный луч попадает на квадрантный фотодетектор.

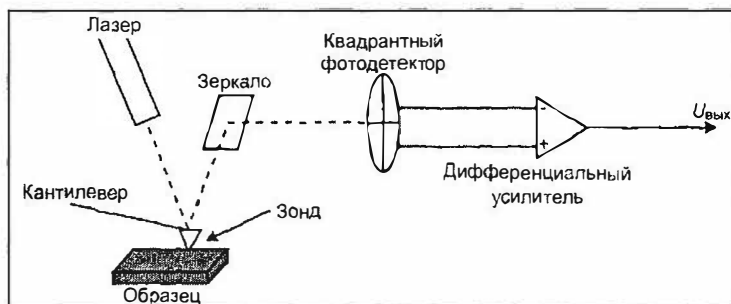


Рис. 7.8. Принцип работы силового сенсора с оптическим датчиком смещений

Вертикальное отклонение регистрируется по разностному сигналу $(A + C) - (B + D)$, а крутильная деформация формирует сигнал $(A + B) - (C + D)$.

7.3.2. Пьезоэлектрический двигатель

Контролируемое перемещение зонда на сверхмалых расстояниях осуществляется с помощью пьезоэлектрических двигателей линейных перемещений Π_x , Π_y , Π_z (рис. 7.9).

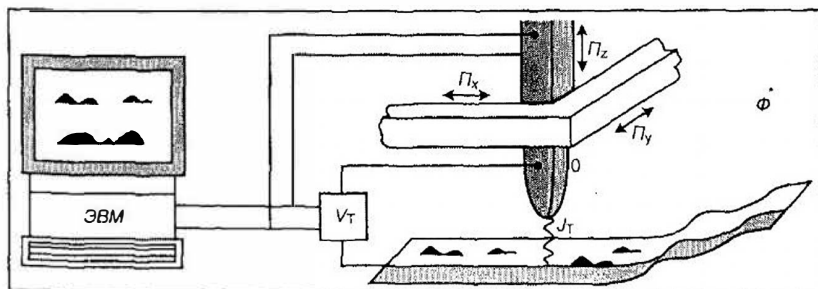


Рис. 7.9. Пьезоэлектрические двигатели линейных перемещений по поверхности (Π_x , Π_y) и по глубине Π_z и цепь обратной связи

В качестве пьезоэлектрических преобразователей наиболее широко применяются цирконат титанат свинца — ЦТС (PZT).

Материалы имеют пьезоэлектрический коэффициент от 0,1 до 300 нм/В, что позволяет перемещать зонд на 0,01 нм при приложении напряжения 100 мВ.

Пьезодвигатели исполняют также функцию точного позиционного зонда в пределах $10 \times 10 \text{ мм}^2$ с точностью $\sim 0,1 \text{ \AA}$.

7.3.3. Устройство управления (обратная связь)

Управление процессом позиционирования зонда, а также получения информационных сигналов о состоянии исследуемой поверхности осуществляется с помощью цепи отрицательной обратной связи.

При приближении зонда к поверхности сигнал сенсора V_x возрастает. Сигнал сенсора сравнивается с опорным напряжением V_{om} и в компараторе вырабатывается корректирующий сигнал $V_{кор}$. Этот сигнал поступает на управляющий пьезодвигатель, который отводит зонд от исследуемой поверхности.

При исследовании туннельного сенсора приближение иглы к поверхности приводит к увеличению туннельного тока. Сигнал рассогласования подается на пьезопривод, который отводит зонд и, благодаря этому, сохраняется заданное расстояние от зонда до поверхности.

В соответствии с теорией автоматического регулирования для поддержания некоторого параметра системы на заданном уровне V_s , например, сигнала на выходе электронного сенсора, необходимо выработать сигнал коррекции V_c , представляемый в виде суммы трех компонент:

$$V_c = V_p + V_i + V_d,$$

где

$V_p = P(V(t) - V_s)$ — пропорциональная компонента отклика;

$V_i = I \int_{t-\tau}^t (V(t) - V_s) dt$ — интегральная компонента отклика;

$V_d = D - \frac{d}{dt}[V(t) - V_s]$ — дифференциальная компонента отклика;

P — константа пропорциональности;

I — интегральная константа;

D — дифференциальная константа.

Пропорциональная компонента отклика V_p обеспечивает фиксацию достаточно резкого изменения сигнала рассогласования $V_p = V(t) - V_s$. Для того чтобы достоверно изображать мелкие детали поверхности, значение величины P должно быть достаточно большим.

Интегральная компонента V_i ответственна за контроль отклика системы в области низких частот, другими словами, именно ее фиксация обеспечивает обработку крупных деталей поверхности.

Дифференциальная компонента V_d играет роль стабилизирующего параметра. Увеличение этого параметра способствует уменьшению осцилляций и увеличению скорости сканирования.

Существуют цифровые и аналоговые системы обратной связи. Правильно отрегулированная обратная связь позволяет подобрать оптимальные параметры P , I и D .

7.3.4. Устройство обработки и индикации информации

Управление и обработка данных осуществляется с помощью компьютера. Работой микроскопа управляет микропроцессор, совместимый с IBM PC. Индикация осуществляется на экране монитора в специальном многоканальном режиме.

7.4. Возможности зондовой микроскопии

Последнее десятилетие в электронике интенсивно развиваются принципиально новые методы изучения поверхностей с нанометровым и атомарным пространственным разрешением и, прежде всего, сканирующая зондовая микроскопия. Данный термин относится к любым типам микроскопов, в которых изображение формируется за счет перемещения (сканирования) острого микрозонда (иглы) над исследуемой поверхностью. Родоначальником таких приборов был сканирующий туннельный микроскоп (СТМ).

Главным достоинством современных СЗМ являются:

- ☐ возможность получения достоверных данных о высоте микрорельефа;
- ☐ отсутствие дополнительных промежуточных процедур (напыление, изготовление реплик), снижающих достоверность результатов;
- ☐ возможность получения нанометрового, а иногда и ангстремного разрешения на воздухе.

СЗМ может применяться для проведения специальных воздействий и нанолитографических операций на нанометровом уровне — токовых, электрополевых воздействий, механических, химических и пр. СЗМ может быть использован для исследований в биологии и медицине, в материаловедении, в исследованиях различных покрытий и тонких пленок, полимерных и наноструктурированных материалов, в химии и химической промышленности, физике и т. д. (рис. 7.10).

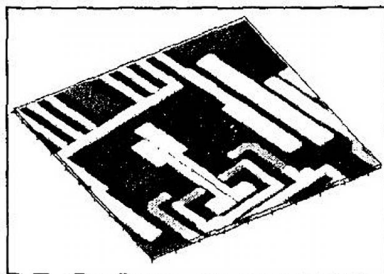


Рис. 7.10. Фрагмент интегральной схемы, полученный на микроскопе SOLVER S-MDT. Поле сканирования 100×100 мкм

СЗМ имеет широкое применение:

- ☐ в науке (в области образования, общей метрологии, материаловедении, в исследованиях структур металлов и сплавов, в исследованиях полупроводниковых приборов и интегральных схем, свойств тонких пленок, в разработках запоминающих сред, в том числе для терабитной памяти, для манипуляций на нанометровом уровне);

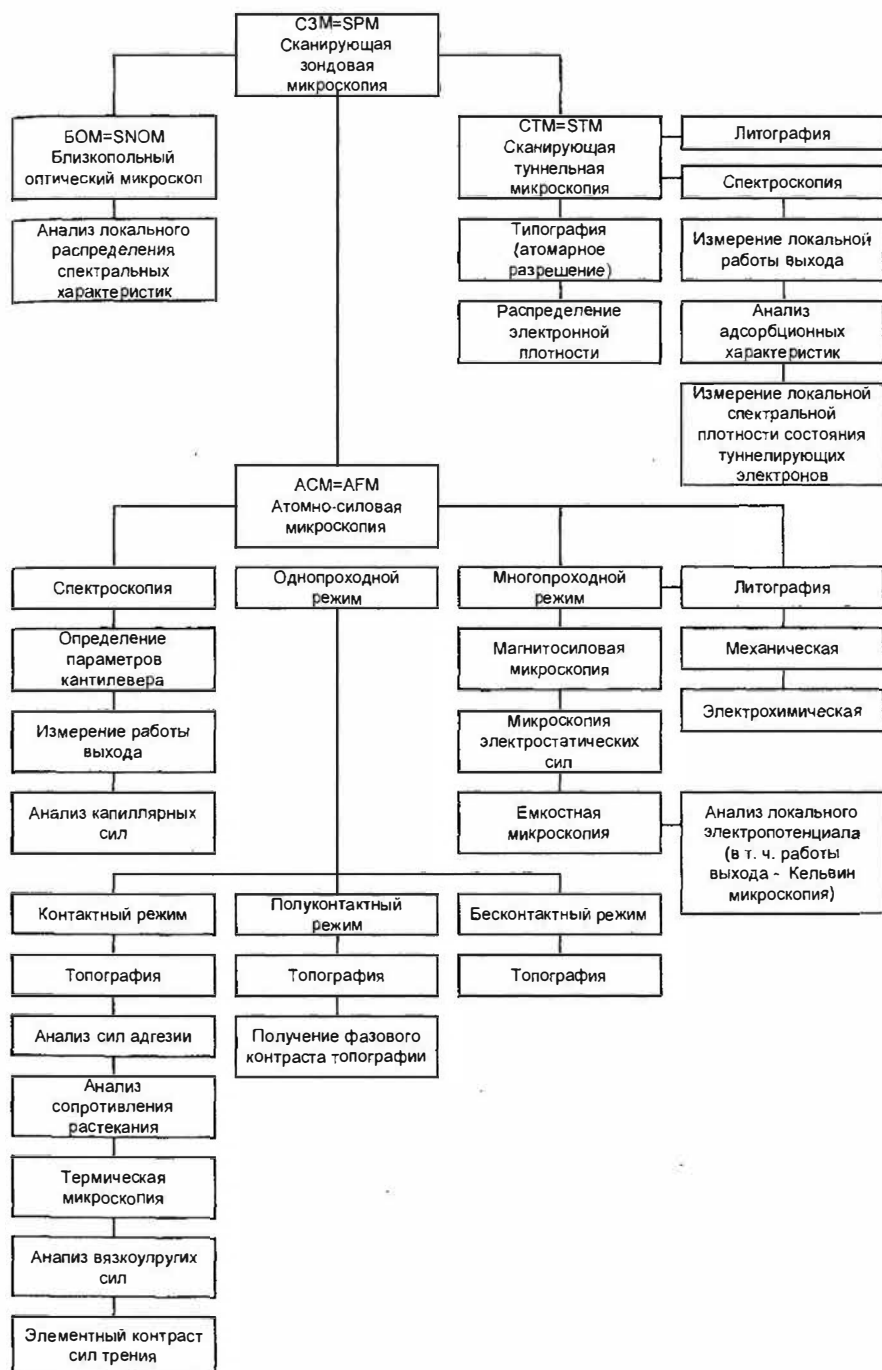


Рис. 7.11. Классификация основных методик СЗМ

- *в промышленности* (в металлургии и металлообработке, оптической промышленности, при анализе качества поверхности материалов, медицине и медицинской промышленности, в производстве порошковых материалов, красок, защитных покрытий, в микроэлектронике, в производстве компакт-дисков, накопителей и устройств записи-считывания для ЗУ сверхбольшой емкости).

На рис. 7.11 приведены основные методики сканирующей зондовой микроскопии.

Фундаментальные достижения в области вакуумной электроники являются основой для развития новых направлений в науке. Например, в сканирующей зондовой микроскопии, нанoeлектронике и нанотехнологии широко используются достижения вакуумной электроники. И это не последнее, что взято из annалов вакуумной электроники.

Контрольные вопросы

1. Опишите явление вакуумного тунелирования.
2. Что такое силы Ван-дер-Ваальса?
3. Что такое кантилевер и как он устроен?
4. Каковы требования к зонду туннельного микроскопа?
5. Опишите структурную схему атомного силового микроскопа.
6. Какие методики исследований с помощью сканирующего зондового микроскопа вы знаете?

Рекомендуемая литература

1. Арутюнов П. А., Толстихина А. Л. АСМ в задачах проектирования приборов микро- и нанoeлектроники. — Микроэлектроника, т. 28, № 6, 1999.
2. Быков В. А. Микромеханика для сканирующей зондовой микроскопии и нанотехнологии. — Микросистемная техника, № 1, 2000.
3. Неволин В. К. Физические основы туннельно-зондовой нанотехнологии. Учебное пособие (рукопись). — М.: МИЭТ, 2000.
4. Свистунов В. М., Белоголовский М. А., Дьяченко А. И. Вакуумная туннельная микроскопия и спектроскопия. — УФН, т. 154, в. 1, 1988.

Заключение

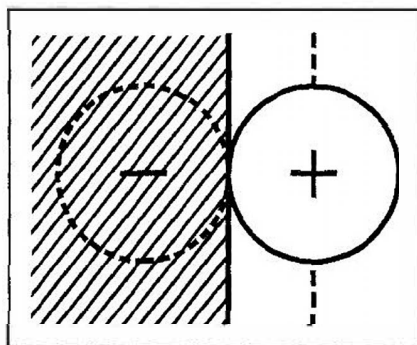
Вакуумная электроника возникла в начале прошлого века и стала базовой для развития твердотельной электроники (включая полупроводниковую электронику и микроэлектронику), квантовой электроники, функциональной электроники, оптоэлектроники.

На своем более чем вековом пути развития вакуумная электроника стимулировала развитие исследований в области термоэлектронной, фотоэлектронной, автоэлектронной (туннельной), вторично-электронной и других видов эмиссии, физики и техники высокого и сверхвысокого вакуума, в области формирования потоков электронов и ионов, технологии обработки поверхностей и других областей. Достижения вакуумной электроники, ее технология были широко использованы как в смежных, так и, на первый взгляд, отдаленных областях науки и техники. Примеров тому великое множество.

На протяжении столетия совершенствовались характеристики электронных приборов, постепенно приборы вакуумной электроники были вытеснены приборами микроэлектроники, квантовой электроники, функциональной электроники и др. Например, устройства вычислительной техники первого поколения были реализованы только на электронных лампах, а современные компьютеры только на приборах микроэлектроники.

Вместе с тем и сегодня позиции приборов вакуумной и плазменной электроники остались незыблемы в электронике больших мощностей, радиовещании, радиолокации, ускорительной технике, дальней связи, плазменной технологии и многими другими областями техники.

Наука по-прежнему на марше!



ЧАСТЬ II

Микроэлектроника

Введение

1. Микроэлектроника вчера и сегодня
2. Физика полупроводниковых структур
3. Интегральные транзисторные структуры
4. Элементная база интегральных схем
5. Технология производства интегральных схем
6. Интегральные схемы и их классификация
7. Логические элементы интегральных схем
8. Запоминающие устройства
9. Триггеры
10. Цифровые устройства на основе триггеров
11. Микропроцессоры и компьютеры
12. Аналоговые интегральные схемы
13. Базовые матричные кристаллы
14. Интегральные схемы СВЧ-диапазона
15. От микро- к нанoeлектронике

Заключение

[REDACTED]

[REDACTED]

Введение

Курс "Микроэлектроника" является базовым курсом при подготовке специалистов по направлению "Электроника и микроэлектроника". Существующая учебная литература по специальностям этого направления не полностью удовлетворяет требованиям Государственного стандарта.

Не умаляя полезности существующих учебников и учебных пособий по микроэлектронике, настоящая книга может стать фундаментом для начального образования по микроэлектронике, своеобразным мостиком для освоения новых областей электроники, таких как наноэлектроника, микросхемотехника и функциональная электроника.

Материал этой части, как и всего учебника, излагается с позиции системного анализа. Автор посчитал важным привести краткий исторический раздел, в котором изложены основные вехи развития полупроводниковой электроники и микроэлектроники, имена ученых, включая наших отечественных, которые внесли решающий вклад в развитие этого раздела электроники.

В каждом разделе приведены задачи, которые снабжены решениями. Это классические, ранее апробированные задачи. Некоторые из них решены не только традиционными методами расчета, но и с помощью компьютеров с использованием стандартных программ. Эти задачи могут быть использованы для проведения индивидуальных занятий, а также для самостоятельной работы студентов. Знакомясь с решениями задач, студент может самообразовываться, самостоятельно разбираться в отдельных вопросах курса лекций. Контрольные вопросы могут быть использованы при составлении экзаменационных билетов.

1. Микроэлектроника вчера и сегодня

1.1. Краткая историческая справка

По количеству выпускаемых промышленностью единиц, по точности технологии и чистоте производства, по количеству областей применения нет прибора, который хоть в какой-то мере мог приблизиться к транзистору. На основе транзисторов созданы интегральные схемы — совершенный продукт технологического прогресса XX столетия. Как же и когда возник транзистор и интегральная схема — чудо-приборы?

В основе работы транзисторов лежат физические явления и эффекты в полупроводниках, или, как их называли более 100 лет назад, в "плохих" проводниках.

В 1833 году Майкл Фарадей столкнулся с первой загадкой полупроводников. При исследовании сернистого серебра он обнаружил, что с повышением температуры электропроводность образца возрастает. А всем уже было известно, что металлические проводники линейно увеличивали свое сопротивление с ростом температуры. И только сопротивление образцов из семейства "плохих" проводников, в данном случае сернистого серебра, наоборот, уменьшалось по экспоненциальному закону. Физики не могли тогда дать ответ на эту загадку "плохих" проводников.

В 1851 году Александр Эдмон Беккерель обнаружил, что при освещении "плохого" проводника светом появляется электродвижущая сила. Возникновение фото-ЭДС или фотогальванический эффект стал второй загадкой в ряду неразгаданных.

В 1874 году немецкий физик Карл Фердинанд Браун обнаружил, что переменный ток, проходя через контакт свинца и пирита, выпрямляется. Пирит, или колчедан, был самым распространенным минералом, который в то время всесторонне исследовался. Так вот, сопротивление контакта не подчинялось известному и почитаемому всеми физиками закону Ома. Более того, свойства контакта определялись величиной и знаком приложенного напряжения. Это была третья загадка "плохих проводников".

Эффект выпрямления тока был обнаружен и в других веществах. К ним были отнесены сульфиды и оксиды металлов, кремний, закись меди и др. Этот класс веществ стали называть полупроводниками. Природу выпрямления тока тогда объяснить так и не удалось. Большинство ученых отнесли этот эффект к термическому.

Американский физик Эдвин Герберт Холл в 1879 году открыл явление возникновения электрического поля в проводнике с током, помещенном в магнитное поле, которое направлено перпендикулярно направлению тока и магнитному полю. В одних полупроводниках возникающее электрическое поле направлено в одну сторону, в других — в противоположную. Вышло так, что существуют отрицательно заряженные частицы — электроны и какие-то, в то время неизвестные, положительно заряженные, которые и определяют направление этого поля. Открытое явление стало четвертой загадкой "плохих" проводников.

Эффекты выпрямления электрического тока и фотопроводимость полупроводников стали использовать для практических целей. Уже в 1876 году В. Адамсон и Р. Дей создали

фотоэлемент на основе селена, а в 1883 году С. Фритте изготовил первый твердотельный выпрямитель электрического тока. Однако в то время объяснить их работу никто так и не смог.

Большинство ученых считали, что эти "загадки" могут быть объяснены какими-то еще неизвестными термическими эффектами. Теория теплоты тогда могла объяснить многое, но не это.

Все это свидетельствовало об электрической, а не термической природе явлений. Особо подчеркнем, что созданная в 1860—1865 годах Дж. Максвеллом теория электромагнитного поля не объясняла ни одну из четырех загадок.

Только лишь в 1906 году Г. Пирс доказал электрическую природу выпрямления тока.

Пока лучшие теоретики искали отгадки, инженеры все шире применяли полупроводники. В начале нашего века ученых захватили исследования в области беспроводной связи. Они создавали приемники радиоволн, которые могли детектировать сигналы. В первых конструкциях радиоприемников использовались контакты полупроводникового материала и металла, например, карбид кремния, теллур, окись цинка, селен совместно с металлической пружиной. Кристаллические полупроводниковые детекторы излучения позволяли выпрямлять радиочастотные сигналы, но усиливать их не могли.

Наш соотечественник, выдающийся радионинженер Олег Владимирович Лосев, изучая в 1922 году свойства кристаллического детектора, обнаружил на вольтамперной характеристике кристалла участок с отрицательным дифференциальным сопротивлением, и на основе этого кристалла получил генерирующий детектор. Советский радионинженер впервые в мире построил полупроводниковый прибор, способный усиливать и генерировать электромагнитные колебания. Основой его служила контактная пара: металлическое острие — полупроводник (кристалл цинкита). Этот прибор вошел в историю полупроводниковой электроники как "кристаллин" Лосева. Позднее, в 1927 году, Лосев открыл также явление свечения кристалла карборунда при прохождении тока через точечный контакт. По существу это был первый светодиод. Открытия Лосева вызвали большой, но кратковременный интерес. Четкого объяснения открытых явлений дано не было, и в то же время воспроизводимость опытов оставляла желать лучшего. Промышленное внедрение так и не состоялось. Тем более что те далекие 30—40-е годы прошлого века были порой расцвета электровакуумных ламп.

Идеи электронной лампы попытались воплотить в твердотельный аналог. В 1925 году профессор физики Юлиус Лилиенфельд получил несколько патентов на конструкцию усилителя с использованием сульфида меди. Спустя десять лет в Англии был выдан патент немецкому изобретателю Оскару Хейлю на полевой триод. В нем предлагалось использовать управляющий электрод для регулирования тока через слой полупроводника.

В 1926 году советский физик Яков Ильич Френкель ввел понятие о дырках в кристаллах, о дефектах кристаллической решетки, представляющих собой дырку в междоузлии — так называемые *дефекты по Френкелю*. В дальнейшем теорию дырок развил английский физик Поль Адриен Морис Дирак и немецкий физик Вернер Карл Гейзенберг в середине тридцатых годов.

В 1932 году советский физик Игорь Евгеньевич Тамм предсказал существование на поверхности кристаллов особых состояний электронов — *уровней Тамма*.

В 1938 году немецкий физик Роберт Вихард Поль разработал конструкцию кристаллического усилителя, действие которого основано на использовании проволоочной сетки для управления потоком электронов через нагретый кристалл бромида калия. Прибор позво-

лял усиливать сигналы и доказывал возможность создания кристаллических полупроводниковых приборов.

Ненадежные в те времена полупроводниковые приборы не могли конкурировать с электронными вакуумными лампами. Известно, что хорошая теория быстро продвигает технические идеи на уровень внедрения. А в полупроводниковой электронике четыре загадки так и остались неразгаданными с прошлого века:

1. Почему сопротивление полупроводников падает с ростом температуры?
2. Почему сопротивление полупроводников уменьшается при их освещении?
3. Почему сопротивление контакта двух полупроводников зависит от полярности прикладываемого напряжения?
4. Почему в полупроводнике с током, помещенном в магнитное поле, могут существовать отрицательно и положительно заряженные носители заряда?

На все эти и другие "почему" никто не мог дать ответ почти 100 лет.

1.1.1. Одна отгадка на все разгадки

К началу 30-х годов проводились интенсивные исследования полупроводниковых структур, предпринимались энергичные поиски природных и синтезированных полупроводников, интерметаллических соединений.

В 1928—1930 годах американский физик Феликс Блох и французский физик Леон Бриллюэн заложили основы зонной теории твердых тел.

И вот в 1931 году английский физик Алан Хэррис Вильсон построил квантовую теорию полупроводников, привлекая математический аппарат квантовой механики. В соответствии с этой теорией в твердом теле энергетические состояния электронов образуют так называемые зоны, разделенные промежутками запрещенных значений энергий. Верхняя зона, в которой находятся свободно перемещающиеся заряды, получила название *зоны проводимости*, а нижняя, в которой заряды находятся в связанном состоянии, была названа *валентной зоной*. Промежуточное расстояние называли *запрещенной зоной*.

Исходя из представлений о зонной структуре электронного спектра, Вильсон провел деление кристаллов на металлы, полупроводники и диэлектрики. Если ширина запрещенной зоны велика, то в твердом теле с такой энергетической характеристикой электропроводность отсутствовала. Такие вещества называются диэлектриками. Если же ширина запрещенной зоны невелика, то существуют различные возможности возбуждения электронов. Это *полупроводники*. Например, при разогреве твердого тела происходит тепловое возбуждение электронов, повышается их энергия, и они из валентной энергетической зоны переходят в более высокоэнергетическую зону проводимости. При этом вещество с такой энергетической зонной структурой обладает большей электропроводностью, а значит, меньшим сопротивлением. С ростом температуры число возбужденных электронов увеличивается, стало быть, сопротивление полупроводников падает. Возможен и другой механизм перевода электронов из валентной зоны в зону проводимости. Речь идет о возбуждении электронов квантами света. Термически или оптически возбужденные электроны становятся свободными электронами проводимости. Таким образом, теория Вильсона объяснила две первые загадки: почему сопротивление полупроводников падает при нагревании и освещении. Впоследствии выяснилось, что процесс электропроводности полупроводников значительно сложнее. Квантово-механическое представление структуры твердого тела подсказывало исследователям, что освободившиеся от электронов места в процессе их перехода в зону проводимости образуют вакансии или *дырки* и

становятся свободными носителями положительного заряда. Поведение дырок аналогично поведению возбужденных электронов. Они обладают подвижностью, эффективной массой и могут образовывать электрический ток, направление которого противоположно току электронов. Стал понятен эффект Холла, когда в одних полупроводниках преобладают отрицательные носители заряда, а в других — положительные.

Стало ясно, что существуют полупроводники с *электронным* типом проводимости (n-типа), для которых эффект Холла имеет отрицательное значение во всей области температур. Вместе с тем, есть полупроводники с положительным значением эффекта Холла, которые имеют *дырочный* тип проводимости (p-тип). Первые были названы донорными, а вторые — акцепторными. Зонная теория "заработала". Открытия посыпались, как из рога изобилия.

А вот разгадать загадку контактного сопротивления долго не удавалось. Ни теория Вильсона, ни дальнейшее ее развитие не могли объяснить этот феномен полупроводников. А между тем на практике широко использовались полупроводниковые выпрямители электрического тока. Хорошие выпрямители одновременно выполняли функции хороших фотоэлементов. Выяснилось также, что величина термо-ЭДС в полупроводниках на несколько порядков выше, чем в металлах. Все эти экспериментальные факты нуждались в обобщении и объяснении. Теория Вильсона не могла все объяснить, в частности механизм выпрямления тока. Оставалась одна загадка — "загадка века".

Шли напряженные исследования. В конце 30-х годов три физика — советский Александр Сергеевич Давыдов, английский Невилл Фрэнсис Мотт и немецкий Вальтер Шоттки независимо друг от друга предложили теорию контактных явлений. Оказывается, в полупроводниковых материалах вблизи границы дырочного и электронного типов полупроводников имеет место область обеднения носителями заряда. В этом месте возникает эффективный электронно-дырочный барьер для равновесных электронов, который не позволяет электронам и дыркам свободно "гулять" по полупроводнику. Через такую систему ток проходит свободно в одном направлении, а в другом — плохо. Электрическое сопротивление системы зависит от величины и направления приложенного напряжения. Например, при приложении электрического поля в прямом направлении высота барьера снижается, и наоборот. Неосновные носители заряда (дырки в электронном и электроны в дырочном полупроводнике) играют определяющую роль. Соответствия между теоретическими и экспериментальными данными были только качественными. Свойства твердого тела очень чувствительны к структуре, к присутствию в ней дефектов. К таким чувствительным свойствам относятся электропроводность, фотопроводность, люминесценция, механическая прочность. Наличие и природа рассматриваемых эффектов, их роль в формировании свойств твердого тела были обнаружены в период становления квантовой механики. Одновременно с развитием теории обрабатывались технологические процессы получения полупроводников с заданным типом проводимости. В результате многочисленных экспериментов удалось изготовить образец, вырезанный из цилиндрического слитка, включающий границу перехода между двумя типами проводимости. Так, впервые был создан *p—n-переход*, ставший важнейшим элементом современной полупроводниковой электроники.

В 1941 году советский физик Вадим Евгеньевич Лашкарев впервые обнаружил *p—n-переход* в записи меди. Вскоре он открыл биполярную диффузию неравновесных носителей тока, построил общую теорию фото-ЭДС в полупроводниках.

Итак, лишь в 40-е годы, почти через 100 лет, удалось разгадать все основные загадки "плохих" проводников.

В эти предвоенные и военные годы быстрыми темпами развивалась радиолокация, мощные источники высокочастотной энергии. Нужны были электронные приборы дециметро-

вого и сантиметрового диапазона длин волн. Тут электронные лампы были вне игры. В этой области эффективными были именно твердотельные детекторы. Так германиевые и кремниевые детекторы или СВЧ-диоды могли использоваться в частотном диапазоне до 4 ГГц.

В СССР эти работы велись в основном в НИИ-108 и НИИ-160 под научным руководством Сергея Григорьевича Калашникова и Александра Викторовича Красилова, соответственно.

В послевоенный период были созданы германиевые выпрямители с коэффициентом полезного действия (КПД), равным 98%.

Разработки отечественных и зарубежных ученых в те годы шли примерно на одном уровне. Дело в том, что руководители страны считались с мнением ученых и интенсивно развивали научные и прикладные исследования по широкому спектру проблем радиоэлектроники. Организовывались новые НИИ и КБ, заводы и лаборатории при них. Они оснащались современным уникальным оборудованием.

В области полупроводников в те годы стояла нерешенная проблема — протекание тока через контакт "металл — полупроводник". Широко применялись методы зондовых исследований поверхностного потенциала вокруг точечного контакта. Был выявлен эффект управления током одного из точечных контактов с помощью рядом расположенного другого контакта.

Однако интерпретация была разной. Наши исследователи были очень осторожны, не поверили в эффект усиления сигнала и заставляли сотрудников более тщательно чистить поверхность (Калашников С. Г., Пенин Н. А.). А вот американские физики увидели новое явление и создали точечный транзистор.

1.1.2. Эпоха транзисторизации

Наступил 1948 год. На одной из страниц газеты "Нью-Йорк таймс" за 1 июля, посвященной развитию радио и телевидения, было помещено скромное сообщение. В нем говорилось о том, что фирма Bell Telephone Labs разработала новый электронный прибор. Этот прибор был способен работать в устройствах вместо "незаменимой" в те времена электронной лампы. Новый прибор фирма намеревалась использовать вместо телефонных реле, надежность и быстродействие которых не удовлетворяли запросам того времени.

Успех пришел к сотрудникам отделения твердого тела фирмы Bell Telephone Labs. Группа в составе Джона Бардина, Уолтера Браттейна, Джеральда Пирсона, Уильяма Шокли, Ричарда Гибни, Гордона Мура, в которой никто ранее не работал в области физики твердого тела, заинтересовалась работами по созданию кристаллического детектора для радаров типа "кошачий ус". Эти полупроводниковые элементы были изготовлены из германия или кремния. К тому времени уже знали, как сделать так, чтобы они были чистыми, а также было известно, как и чем их легировать для получения материалов p- или n-типов.

Уильям Шокли предполагал сделать полупроводниковый усилитель по принципу полевого эффекта. В основе его конструкции была тонкая полупроводниковая пленка, проводимость которой должна была меняться при наложении поперечного электрического поля. Расчеты, проведенные Шокли, показывали, что усиление электрического сигнала может быть получено при условии формирования индуцированного заряда из подвижных носителей (электронов и дырок). Однако экспериментально обнаружить эффект усиления не удалось. К работам подключился Дж. Бардин, выдвинувший гипотезу, согласно которой индуцированный заряд образуется электронами, захваченными в состояниях у поверхности. Это экранирует внутрен-

ную область от приложенного поля. Даже несколько атомов на поверхности полупроводника способны защитить или экранировать его от поверхности поля. Когда же плотность поверхностных состояний не слишком велика, поле контактного потенциала может увеличить или уменьшить слой пространственного заряда. Это также стало ключом к отгадке тайны выпрямляющих ток контактов. Эффект выпрямления объясняется электростатической разностью потенциалов между внутренней частью полупроводника и его поверхностью.

Гипотеза поверхностных состояний была весьма плодотворной и позволила сделать несколько предположений, которые можно было экспериментально проверить. По одной из них слой пространственного заряда может существовать у свободной поверхности. В этом случае заряды в слое компенсированы поверхностными состояниями. В тонком слое, прилегающем к поверхности, может возникнуть инверсный слой с типом проводимости, противоположным тому, который характерен для объема материала. У. Браттейн экспериментально подтвердил эту гипотезу.

Справедливости ради отметим, что на существование поверхностных состояний на свободной поверхности твердого тела указали советский физик Тамм (1932 г.) и немецкий физик Шоттки (1939 г.). Была выполнена серия экспериментов по подтверждению положений теории поверхностных состояний, подтверждено существование электрического поля. Носители, генерированные при освещении перехода, рассасывались этим полем в противоположных направлениях в зависимости от знака заряда. Также менялось и значение контактной разности потенциалов.

А тем временем работы по изучению поверхностных состояний продолжались. С ними были связаны идеи по созданию усилителя на полевом эффекте. Усилиями У. Браттейна и Р. Гибни было выяснено, что поверхностные состояния можно не учитывать, если в контакте с поверхностью находится электролит, а поле направлено перпендикулярно поверхности. В то же время Дж. Пирсон показал, что подвижность носителей в пленках намного ниже, чем в объемном материале. Эти экспериментальные результаты навели У. Браттейна и Дж. Бардина на мысль о создании новой структуры усилителя. Чтобы получить эффект такого слоя в объемном материале, они решили использовать инверсионный слой на поверхности куска кремния с удельной проводимостью противоположного типа. Они использовали контакт типа "кошачий ус" на поверхности и такой же контакт на базе большой площади. Бардин и Браттейн для преодоления пространственного заряда на поверхности полупроводника использовали электролит. Окружающая капля электролита была изолирована от "кошачьего уса". Третий электрод был выполнен в виде проволочной петли, которая находилась в контакте с электролитом.

Первые эксперименты были проведены в ноябре 1947 года (рис. 1.1). Когда на базу подавалось положительное напряжение, большая часть обратного тока носила электронный характер и текла к контакту от инверсионного слоя. Ток уменьшался, когда к электролиту прикладывалось отрицательное напряжение. В соответствии с эффектом поля это было прогнозируемо. Эффект был воспроизводимым, но только на низких частотах. Высокие частоты использовать не удалось из-за низкого быстродействия электролита.

Лучшее быстродействие было получено при использовании инверсионного слоя p-типа на германии n-типа. Стало ясно, что наличие электролита в конструкции не позволит достигнуть высоких результатов, и в связи с этим начались новые поиски. Успех пришел в то время, когда было обнаружено, что при подаче постоянного напряжения смещения на электролит происходит анодное травление германия с образованием оксидной пленки. Эту пленку стали использовать вместо электролита, а на слой оксида положили контактный слой из золотой фольги. Два точечных контакта в германии расположили рядом. На один из

них подали напряжение, полярность которого совпадала с типом проводимости, на другой — обратное (рис. 1.2).



Рис. 1.1. Исследование ведут лауреаты Нобелевской премии 1956 года Уильям Шокли (сидит), Джон Бардин и Уолтер Браттейн (справа). 1947 год

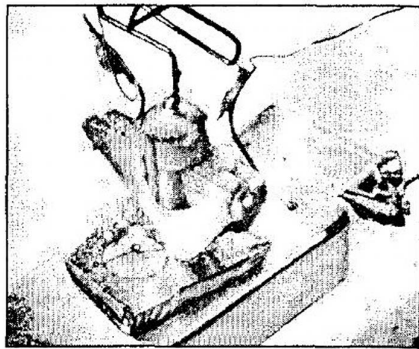


Рис. 1.2. Первый твердотельный усилитель — макет транзистора. Два близкорасположенных контакта выполняли функции эмиттера и коллектора. Они располагались на поверхности германиевого бруска, который служил базой. Вверху находилась пружина, которая прижимала золотую фольгу

Если к металлическому контакту, близкому к "кошачьему ус", приложить напряжение, то инжектируются неосновные носители, и усиленный обратный ток начинает течь к контакту. Если положительное напряжение приложить к контакту, близкому к отрицательно смещенному "кошачьему ус" на германии *n*-типа, то обратный ток увеличивается. Это противоречило ожидаемому от полевого эффекта изменению тока.

Итак, первым твердотельным прибором для усиления электрического тока стал точечный транзистор, в котором два точечных контакта размещаются в непосредственной близости на верхней поверхности небольшого куска германия *n*-типа. При положительном напряжении на эмиттере инжектируются дырки, ток которых направлен к отрицательно смещенному коллекторному контакту. Демонстрация прибора состоялась 23 декабря 1947 года. Экспериментаторы обнаружили, что с помощью тока через один контакт удастся управлять током через другой и усиливать сигналы (до верхней границы звуковых частот) более чем в 100 раз.

Встал вопрос: как назвать свое детище? Вариантов было много. По предложению Дж. Пирса, который в то время был административным руководителем отдела связи фирмы, прибор назвали *транзистором*. Это название отражало сходство полупроводникового прибора с лампами. Действительно, важным параметром лампы является крутизна ее характеристики (*transconductance*), а в полупроводниковом усилителе эффект усиления обеспечивается благодаря переходному сопротивлению (*transresistance*), и в итоге получается *transistor*.

Произошла первая транзисторная революция!

Первые транзисторы не отличались высокой надежностью (рис. 1.2). Виною тому были их конструктивные решения. Отказы шли из-за контактов, выполненных в виде проволочных пружинок ("кошачий ус"). Стоимость транзисторов была велика, а воспроизводимость характеристик плохая. Однако уникальность свойств транзистора стимулировала дальнейшие их исследования.

через год после изобретения транзистора фирма Bell наводнила мир статьями и докладами о транзисторах и их возможностях. Прибор не был засекречен. Видимо, военные не верили в его возможности.

В Советском Союзе транзисторный эффект впервые наблюдали в 1949 году. Авторами первого отечественного точечного транзистора были советские исследователи НИИ-160 (ныне "Исток") Александр Викторович Красилов и его аспирантка Сусанна Гукасовна Мадоян (рис. 1.3). Лабораторный экземпляр работал не более часа, а затем требовал новой настройки — поиска нового точечного контакта с помощью заостренных бронзовых проводочек.

Потом были получены транзисторы в ФИАНе в лаборатории Бенциона Моисеевича Вула, в ЛФТИ в лаборатории Владимира Максимовича Тучкевича, в НИИ-108 в лаборатории Сергея Григорьевича Калашникова, и, наконец, в ИРЭ АН Полным Н. А.

В 1951 году академик, заместитель министра обороны Аксель Иванович Берг созвал совещание специалистов по развитию транзисторостроения. Ленинградский физико-технический институт АН СССР представлял Наследов Д. Н. и Тучкевич В. М., ФИАН АН СССР делегировал Вул Б. М. и Ржанова А. В., от ЦНИИРТИ (НИИ-108) Калашников С. Г. и Пенин Н. А., от Украинской АН — Лашкарев В. Е., от НИИ "Исток" (НИИ-160) — Красилов А. В. Результатом совещания стала постановка ОКР по точечным транзисторам "Точка" в НИИ "Исток" (руководитель Красилов А. В.) и по плоскостным транзисторам НИР "Плоскость". Эта научно-исследовательская работа выполнялась несколькими организациями: ЛФТИ (руководитель Наследов Д. Н.), ФИАН (руководитель Вул Б. М.), НИИ "Исток" (руководитель Красилов А. В.). В конце 1953 года эти работы были сданы Госкомиссии, которая работала во вновь созданном НИИ "Пульсар" (НИИ-35). Сюда же была передислоцирована лаборатория Красилова А. В. В этой лаборатории были изготовлены первые в СССР плоскостные транзисторы П1, П2, П3, которые стали основой для других серийных полупроводниковых приборов. Промышленный выпуск был освоен на ЛОЭП "Светлана" (Я. А. Кацман).

Был создан второй полупроводниковый институт — НИИ-311 в дальнейшем НИИ "Сапфир". Его профилем стали работы по созданию диодов. Одновременно с этим академик Берг А. И. создал новый Институт радиоэлектроники (ИРЭ) в рамках АН СССР, который сам и возглавил. Сюда перешли преимущественно сотрудники ЦНИИ-108, а работы по полупроводникам в нем были свернуты.



Рис. 1.3. Группа отечественных исследователей — пионеров микроэлектроники:

профессор, лауреат Сталинской и Ленинской премий Красилов А. В., профессор, лауреат Ленинской премии Федотов Я. А., старший научный сотрудник Мадоян С. Г., профессор Пенин Н. А., главный научный сотрудник, лауреат Ленинской премии Щиголь Ф. А. ИРЭА

К середине пятидесятых годов общий объем выпуска дискретных полупроводниковых приборов составил 24 миллиона, из которых более десяти процентов составляли транзисторы.

С 1958 года в стране началось интенсивное развитие полупроводниковой промышленности. Под полупроводниковые предприятия на первых порах выделялись не эффективно используемые предприятия: совпартшколы в Новгороде, макаронная фабрика в Брянске, спичечная фабрика в Таллинне, сельхоззавод в Херсоне, швейная фабрика в Воронеже, коммерческий техникум в Риге и т. д. Все эти предприятия быстро реконструировались и оснащались современным оборудованием. За неполное десятилетие была создана отечественная полупроводниковая промышленность, обеспечившая оборонную, ракетно-космическую и народно-хозяйственную мощь державы.

А между тем интерес к транзисторам нарастал лавинообразно. Транзистор стал выдающейся "личностью" в научно-техническом прогрессе. Наступала эпоха транзисторизации...

1.1.3. Вторая транзисторная революция

Итак, "главой" славной династии транзисторов стал германиевый прибор, изобретенный Д. Бардиным, У. Браттейном и У. Шокли, за который в 1956 году они получили Нобелевскую премию. Правда, патент на точечный транзистор получили первые двое в 1950 году, а У. Шокли только в следующем году получил патент на плоскостной транзистор. А всего за первые 20 лет была выдана почти тысяча патентов.

Особое значение приобрели технологические разработки. В 1954 году с помощью *метода зонной чистки* удалось получить слитки полупроводников высокой чистоты. Этот же метод позволил получить равномерное распределение примесей в кристалле. К 1955 году стало возможным получение транзисторов со сплавными и р-, п-переходами, выращенными из расплава. Позже появились и другие разновидности сплавных транзисторов: дрейфовые и сплавные с диффузией.

Разработанный *электрохимический метод* получения тонких базовых слоев позволил создать новые типы дискретных транзисторов: микросплавные, поверхностно-барьерные, сплавно-диффузионные, микрослойные. Частотный диапазон их работы достигал нескольких мегагерц. Началось производство автоматизированного оборудования для производства транзисторов.

Для получения р—п-перехода стали использовать *диффузионный метод*. Сущность его заключалась в равномерном распределении примесных атомов в кристалле при нагреве его в атмосфере, содержащей необходимые примеси. Этот метод положил начало групповой технологии производства приборов. Наиболее популярным среди транзисторов с диффузионной базой стал *мезатранзистор*, отличающийся высокими рабочими частотами.

Использование процесса диффузии позволило резко увеличить точность задания толщины области базы, который является одним из важнейших параметров. Толщина базы должна быть настолько узкой, чтобы практически все электроны, инжектированные эмиттером, попали через базу в коллектор. При тонкой базе можно уменьшить время, за которое инжектированные электроны попадают на коллектор, а следовательно, увеличить частоту переключения транзистора из одного состояния в другое. Транзисторы с тонкой базой получили название *биполярных*, поскольку в них используются два типа носителей — электроны и дырки. В зависимости от расположения легированных областей они именуются п—р—п или р—п—р-типа.

В конце 50-х годов была разработана технология создания планарных транзисторов, конструкция которых имеет плоскую структуру и расположена у поверхности полупроводника

(рис. 1.4). Отличительной особенностью планарной технологии является возможность создания множества приборов на одной подложке. Это открывало путь к групповой технологии производства транзисторов, к автоматизации их производства.

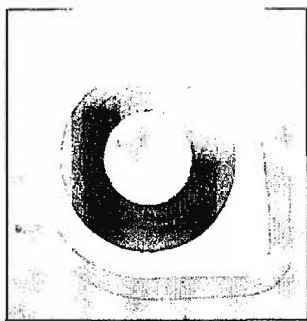


Рис. 1.4. Первый планарный транзистор "бычий глаз", который ознаменовал собой начало второй транзисторной революции. 1959 год

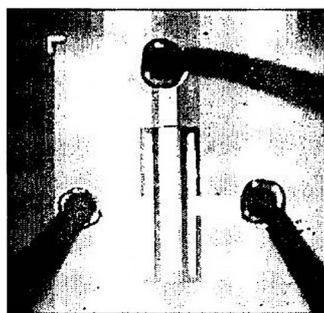


Рис. 1.5. Первый МОП-транзистор с р-каналом. 1964 год

Создание высокочастотных транзисторов вело к увеличению активной площади структуры. Так появились приборы со структурой эмиттера в виде гребешка с хорошо развитым периметром.

История разработки транзисторов знает примеры новых технических решений, которые открывали отдельные направления в полупроводниковой электронике. Таким примером может служить разработка *полевых транзисторов*. Этот прибор одновременно мог выполнять функции резисторов, управляемых напряжением. Первый промышленный полевой транзистор был изготовлен в 1958 году Станиславом Тешнером во Франции и получил название *текнетрон*.

Типичный полевой транзистор был реализован на базе структуры металл — окисел — полупроводник (МОП). В кремниевой подложке, например, р-типа, создаются два островка кремния п-типа. Эти островки снабжены металлическими контактами, один из которых называется *источком*, другой — *стоком*. На поверхности кремния над пространством между истоком и стоком (каналом) наносится слой двуокиси кремния. Поверх этого слоя осаждается тонкая металлическая пленка, образующая третий электрод, называемый *затвором*. Двуокись кремния является хорошим изолятором, поэтому затвор не имеет прямого электрического контакта с полупроводниковой подложкой. Электрическая связь осуществляется через электрическую емкость, другими словами, через электрическое поле, генерируемое зарядом на затворе. Электрическое поле эффективно влияет на движение заряженных носителей в канале полупроводника. В МОП-транзисторе в инверсном канале ток образуют носителями только одного знака. Поэтому этот тип транзисторов иногда называют *униполярным* (рис. 1.5).

Как и биполярный транзистор, полевой МОП-транзистор обладает способностью усиливать электрический ток. Однако в этом случае усиление происходит по напряжению, а не по току, как в биполярном транзисторе.

В настоящее время транзисторы по-прежнему являются основной продукцией полупроводниковой промышленности. Они составляют 60—70% рынка сбыта дискретных приборов. Ныне применяются три типа транзисторных структур: биполярные, полевые и так называемые *БиМОП*, представляющие сочетание первых двух. Биполярные транзисторы выпускаются маломощные (до 1 Вт), средней мощности (1—20 Вт), высокой мощности

(более 20 Вт). Они работают на высоких частотах — свыше 1 ГГц. Маломощные транзисторы весьма миниатюрны и используются в схемах специального назначения, гибридных интегральных схемах и т. п. Мощные транзисторы широко используются в выходных цепях усилителей низкой частоты, генераторах, схемах управления, в качестве релейных элементов, в сервоусилителях, сенсорных переключателях и т. п. Намети́лась тенденция создания высоковольтных, комплементарных, высокочастотных мощных транзисторов. Коэффициент усиления ряда транзисторов достигает значения 5×10^4 и более, рабочие напряжения около 500 В, пиковые значения тока до 200 А; они сохраняют способность работать при температурах корпуса порядка 100 °С.

В мировой электронике идет конкурентная борьба за создание мощных транзисторов дециметрового диапазона для использования в бортовом и наземном оборудовании в диапазоне частот 225—400 МГц. Типичное усиление таких транзисторов составляет 8—9 дБ. Перспективным прибором является биполярный транзистор с проницаемой базой, который сможет работать на частотах 10 ГГц и более.

В последние годы в производстве полупроводниковых приборов отмечается возрастание роли полевых транзисторов из арсенида галлия с затвором на основе барьера Шоттки. Разработаны компактные приборы на частоте до 40 ГГц с коэффициентом усиления от 4 до 14 дБ и коэффициентом шума в пределах 0,5—4,0 дБ. Значительный прогресс достигнут в области мощных полевых СВЧ-транзисторов. Приборы этого типа рассчитаны на частоту 4—25 ГГц с коэффициентом усиления 3—6,5 дБ, выходной мощностью 20—45 дБ и КПД до 40%. Ведутся разработки мощных высоковольтных полевых транзисторов. Разработаны транзисторы с рабочим напряжением до 500 В, работающие при токах 20 А в непрерывном и 50 А в импульсном режимах; типичное время включения составляет 60 нс, выключения — 200 нс. Этот тип транзисторов найдет применение в схемах управления электроприводом, электронных устройствах автомобилей, системах робототехники, управления станками.

1.1.4. Третья транзисторная революция — рождение микроэлектроники

Бурное развитие дискретной полупроводниковой техники, возможность автоматизации производства таких структур привели к идее интеграции. Эта идея, в сущности, не нова. Еще до Второй мировой войны были попытки изготовления интегрального устройства, объединяющего резистор с конденсатором для катодной цепи электровакуумной лампы. Идея не была реализована потому, что выход годных резисторов и конденсаторов был слишком низким. Ситуация в корне изменилась при интеграции полупроводниковых приборов.

И тут уместно вспомнить любопытный факт. В начале 50-х годов, когда транзисторы были еще модной игрушкой в лабораториях ученых, а промышленная техника изготовлялась на электровакуумных лампах, инженер из Великобритании Джеффри Даммер предложил изготавливать электронное оборудование в виде твердого блока из полупроводникового материала. В таком блоке отсутствуют соединительные провода, он включает в себя слои изолирующих, проводящих, усиливающих, выпрямляющих структур, которые в целом надежно выполняют заданные функции. Эта мысль звучала как анекдот даже в аудитории инженеров-электриков. Мало кого интересовали тогда транзисторы вообще, а проблема их надежности вовсе никого не волновала. Это выступление потонуло в море информационных сообщений. Его "раскопали", но не инженеры, а историки науки лишь через 25 лет! Идея интеграции в полупроводниковом производстве пришла с другой стороны — со стороны технологий в электроном материаловедении.

В 1959 году было предложено использовать транзисторные структуры, соединенные проводниками в пределах одной пластины. Такие транзисторы получили название *интегральных*, а кристаллы стали называть *интегральными схемами* (ИС). Заслуга изобретения интегральных схем принадлежит двум исследователям, каждый из которых сделал это независимо друг от друга. Причем оба ничего не знали об идее Даммера, даже не слышали о нем самом. Это были Джек Сент Клер Килби из фирмы Texas Instruments и Роберт Н. Нойс из фирмы Fairchild Semiconductor.

Д. Килби предлагал коммутировать дискретные полупроводниковые элементы тонкой золотой проволокой. С этой целью он изготовил резисторы, положив в основу омические свойства тела полупроводника, а в качестве конденсаторов использовал обратносмещенные $p-n$ -переходы, вырезанные из полупроводника. Из одной из диффузионных областей путем фототравлировки изготавливался мезатранзистор. Так была создана первая твердая схема.

При всем очевидном прогрессе эта схема не позволяла оптимизировать параметры элементов, выход годных составлял около 10%, неэкономичность производства и невозможность доработки схемы поставили вопрос об ее внедрении в производство.

В схеме Р. Нойса фирмы Fairchild Semiconductor была внедрена идея изоляции отдельных элементов с помощью обратносмещенных $p-n$ -переходов, а коммутация элементов осуществлялась через отверстия в окисле с помощью напыляемых металлических токоведущих дорожек (рис. 1.6).

Это изобретение в 2000 году было отмечено Нобелевской премией, которую получил Ж. Килби. К сожалению, Р. Нойс до этого дня не дожил.

С усложнением схем резко возрасла стоимость межсоединений дискретных элементов. Нужно было решить две проблемы. Во-первых, разработать пассивные компоненты (резисторы и конденсаторы), а во-вторых, преодолеть ограничения, связанные с интеграцией — оптимизировать параметры индивидуальных компонентов в схеме, преодолеть синдром невысокого процента выхода годных, получить возможность изменять функции дорогой и трудоемкой схемы.

Первая проблема решалась довольно просто, ключом к ее решению стала технология, с помощью которой удалось получить участки кремния с различной проводимостью и тем самым изготавливать сопротивления в определенных областях транзисторной структуры. В основу конструирования микроэлектронной емкости были положены обратносмещенные $p-n$ -переходы транзисторных структур.

Что касается второй проблемы, то многие годы удавалось преодолевать ограничения, связанные с интеграцией. Удавалось до определенного времени, пока новые разработанные технологические процессы позволяли получать выход годных близкий к 100%.

Прежде всего, на фирме Fairchild удалось пассивировать поверхность кремния его же окислом. Там же был разработан процесс создания базы в коллекторе методом диффузии. Это позволило соединять все три области транзистора путем напыления на окисел токоведущих дорожек.

На фирме Bell Telephone Labs в 1960 году разработали *эпитаксиальный метод* наращивания слоев на монокристаллическом кремнии. Именно этот процесс позволил создавать транзисторные структуры с воспроизводимыми параметрами, повысить выход годных, снизить стоимость. Таким способом удавалось на прочной толстой подложке создать транзисторы с тонкой базой. Открывалась возможность разработки высокочастотных транзисторов большой мощности. На рис. 1.7 показана первая промышленная ИС, выполненная по этой технологии.

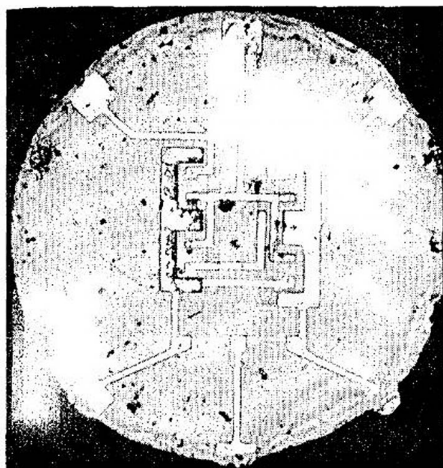


Рис. 1.6. Первый интегральный прибор — триггер, поступивший в продажу. Четыре транзистора $n-p-n$ -типа и два резистора, скоммутированные алюминиевыми токоведущими дорожками. 1961 год

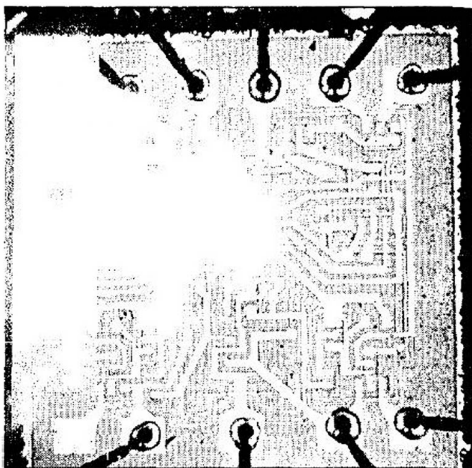


Рис. 1.7. Первая промышленная ИС, выполняющая функцию триггера по схеме диодно-транзисторной логики. Фирма Fairchild. 1964 год

Таким образом, наряду с дискретной полупроводниковой электроникой появилась *интегральная электроника* (ИЭ). Этот термин отражает идею интеграции элементов. Позже за этим направлением в электронике прочно закрепилось название *микроэлектроника*. Этот термин отражает идею микроминиатюризации.

Интегральные транзисторы стали основными изделиями микроэлектроники. Сначала успешно развивались биполярные транзисторы. Наряду с традиционными структурами для дискретной полупроводниковой электроники в семействе интегральных транзисторов появились уникальные. Речь идет о многоэмиттерных и многоколлекторных транзисторах, транзисторах с барьером Шоттки и др. Именно с их помощью удалось создать эффективные схемные решения.

В последние годы интерес проявлен к униполярным транзисторам. Технология их изготовления более простая. К этому типу относятся n -МОП и p -МОП транзисторы, имеющие структуру "*металл—окисел—полупроводник*". Особенно перспективны *комплементарные* или взаимодополняющие транзисторы (КМОП). В этих структурах используются одновременно два транзистора с каналами проводимости p - и n -типов, они отличаются высокой экономичностью и надежностью.

Однако на пути развития традиционных транзисторных структур стоят физические, технологические, экономические и другие ограничения. Примерно те же, о которых волновались еще в 60-е годы.

А что у нас в стране?

В 1962 году в НИИ-35 (НИИ "Пульсар") приступили к разработке первой серии интегральных кремниевых схем ТС-100. Руководителем разработки были Борис Владимирович Малин (начальник отдела) и Анатолий Федорович Трутко (директор института). Цикл планарной технологии включал более чем 300 операций, который был освоен самостоятельно с нулевого уровня. Были созданы опытный цех, научно-технологический отдел, ставший полигоном для обучения специалистов в области планарной технологии. Были сконструированы авто-

матризированные агрегаты пооперационной обработки кремниевых пластин, разработано специальное технологическое оборудование. Работа по серии ТС-100 длилась до 1965 года, а затем еще два года шло освоение заводского производства с военной приемкой.

Для ведущей мировой державы, каким был СССР, масштабы производства ИС были явно недостаточны.

8 августа 1962 года было подписано Постановление ЦК КПСС и СМ СССР о создании Научного центра микроэлектроники в подмосковном городе Крюкове. В соответствии с ним в состав Центра вошли вновь созданные НИИ теоретических основ микроэлектроники, НИИ микросхемотехники, НИИ технологии микроэлектроники, НИИ машиностроения, НИИ специальных материалов и три опытных завода при этих институтах. Началось интенсивное строительство Центра и города со сложной инфраструктурой, обслуживающей Центр. Сегодня этот город носит название Зеленоград, ныне ставший районом Москвы и являющийся в некоем отношении аналогом Кремниевой долины (США).

Научный Центр должен был охватывать все аспекты микроэлектроники, весь цикл "исследование — производство". Постановление было документом де-юре. Де-факто все сложилось по-другому, более масштабно. И это только благодаря усилиям Министра электронной промышленности Александра Ивановича Шокина.

В 1962 году был введен в строй НИИ микроприборов (директор Букреев И. Н.) с опытным заводом "Компонент" и институт по разработке специального технологического оборудования — НИИ точного машиностроения (директор Иванов Е. Х.) с заводом электронного машиностроения "Элион".

В 1963 году был организован НИИ точной технологии (директор Сергеев В. С.), которому через некоторое время был придан завод "Ангстрем", НИИ материаловедения (директор Малинин А. Ю.) с заводом "Элма". Год спустя вошел в строй НИИ молекулярной электроники (и. о. директора Гуреев И. А.) с заводом "Микрон". Одновременно было организовано Центральное бюро применения интегральных схем (ЦБПИМС). Поисковыми исследованиями должен был заниматься НИИ физических проблем (директор Лукин Ф. В.). Было решено готовить кадры на месте, в зоне Научного центра. В 1965 году был образован институт электронной техники (МИЭТ), которому был придан собственный опытный завод.

А пока кадры подрастали, целые коллективы воспитывались в Москве. Так практически коллектив будущего НИИ МЭ воспитывался в отделе Малина Б. В. в НИИ "Пульсар". К этому времени в НИИ "Пульсар" было развито не только полупроводниковое производство, но и реально разрабатывались первые ИС. В 1965 году НИИ МЭ возглавил Камилль Ахметович Валиев, ныне академик, директор физико-технологического института РАН. На основе технологии, близкой к технологии производства планарных транзисторов "Плоскость", коллективом НИИ МЭ под руководством К. А. Валиева уже в 1966 году были разработаны ИС типов "Иртыш", "Микроватт", "Логика", диодно-транзисторные логические схемы. К концу этого года их было выпущено около 100 тысяч штук.

Структура предприятий Научного центра менялась в соответствии с задачами промышленности. Со временем на базе НИИ микроприборов и завода "Компонент" было создано научно-производственное объединение "Элас", тесно связанное с космической отраслью. С момента зарождения квантовой электроники возникли предприятия этого профиля: НПО "Зенит", КБ высокоинтенсивных источников света с заводом "Фотон", филиал НИИ "Фонон".

Параллельно с развитием Научного центра в Зеленограде полупроводниковое и микроэлектронное производство развивалось в подмосковном Томилино и Воронеже на заводах

полупроводниковых приборов, в Ленинграде на НПО "Светлана", в Минске на заводе полупроводниковых приборов (впоследствии НПО "Интеграл"), в Вильнюсе в НИИ "Вента", в Кишиневе на заводе "Мезон", в Тбилиси в НИИ "Мион", в Баку на заводе "Азон", а также в Риге, Новосибирске, Павловском Посаде, Фрязино. Это позволило уже к 1970 году выпустить 3,6 миллионов ИС 69 серии.

В США также вырастали новые фирмы. Так только в Кремниевой долине в 1966—67 годах было создано три новые фирмы, в 1968 — тринадцать, в 1969 — восемь. Росла номенклатура микроэлектронных приборов и интегральных схем на основе использования различных технологических процессов. Появились запоминающие устройства (ЗУ) с произвольной выборкой, перепрограммируемые ЗУ, разработаны ИС с инжекционной и эмиттерно-связанной логикой, приборы с зарядовой связью, комплементарные МОП-приборы. Микроэлектронная технология позволила создать ЗУ на цилиндрических магнитных доменах.

В 70-х годах прошлого века ежегодный объем продаж полупроводниковых приборов и ИС превысил 6 млрд долларов.

Начиналась эра *сверхбольших интегральных схем* (СБИС).

В 1973 году появились первые микропроцессоры, представляющие собой большие интегральные схемы с функциями вычислительных устройств. Начался новый виток развития микроэлектроники. В США фирма Intel захватила мировое лидерство, которое удерживает до сегодняшнего дня.

В Советском Союзе с 1975 года разрабатывались микропроцессоры в Зеленоградском научном центре (НИИ ТТ, НИИ МЭ) в Воронеже (ВЗПП, позже объединение "Электроника"), в Ленинграде (НПО "Светлана"). В 1976 году предприятия МЭП выпустили 100 миллионов интегральных схем, среди которых были микропроцессоры и микропроцессорные комплекты. Американская пресса в те годы писала, что отставание уровня развития советской электроники резко сократилось с 10 до 2—3 лет. К концу 70-х по всей стране выпускалось 700 миллионов ИС.

Это позволило наладить выпуск компьютеров отечественного производства, обеспечить ортовыми ЭВМ все военные объекты от космических аппаратов до подводных лодок.

Отечественная электроника и микроэлектроника была практически уничтожена экономическими реформами, начатыми в 1991 году. Ныне уровень развития отечественной микроэлектроники отстает уже не на год, и не два. Навсегда. Кое-где предприятия отрасли сохранились, и наметилась тенденция к их возрождению в новых экономических и политических условиях. И как писал бывший начальник Главного научно-технического управления МЭП СССР Пролейко В. М.: "Неужели нужно событие, аналогичное Отечественной войне, чтобы руководители страны наконец-то осознали, что значит электроника сегодня? Или чтобы те, кто не хочет это осознать, не могли быть руководителями".

1.5. Грядет ли новая транзисторная революция?

Жизнь в микроэлектронике продолжается. За 40 лет развития ИС плотность размещения транзисторов на кристалле ежегодно увеличивается примерно на 50%. Это первым замечил Гордон Мур, один из основателей фирмы Intel, основного разработчика серийных микропроцессоров. За четыре десятилетия степень интеграции возросла в 10 000 раз!

Сегодня разработчики интегральных схем считают, что экспоненциальный рост степени интеграции будет сопровождаться экспоненциальным ростом потребляемой энергии.

Масштабирование элементов транзистора, выход на субмикронные размеры порождает ряд труднорешаемых проблем. Например, металлические токоведущие дорожки могут потерять свойство проводников и превратиться в емкостные и индуктивные элементы.

И масштабированные транзисторы уже не просто транзисторы. Так по мере уменьшения длины каналов МОП транзисторов уменьшается толщина затворного окисла, а значит, и уменьшается пороговое напряжение. Чем меньше пороговое напряжение транзистора, тем больший ток требуется для переключения транзистора. Растут и токи утечки транзисторных структур. Подсчитано, что КМОП ИС с 0,09 мкм элементами и напряжением питания 1,2 В потребляет ток на два порядка больше, чем МОП микросхема с 0,25 мкм и напряжением питания 2,5 В.

Были предложены новые конструкции транзисторов, использующие баллистические электроны в полупроводниках. В такой конструкции полевого транзистора исток и сток располагаются один над другим на небольшом расстоянии. Эмитируемые из истока электроны благодаря своим волновым свойствам проходят межэлектродное расстояние в кристаллической структуре по баллистическим траекториям без рассеяния. Если сделать расстояние между истоком и стоком или канал проводимости коротким, то быстродействие транзистора станет очень высоким.

Разработаны транзисторы на баллистических электронах, использующие в качестве инжекторов электронов туннельный барьер. Конструктивно такой инжектор выполнен в виде тонкого слоя нелегированного твердого раствора AlGaAs, который располагается между уже легированными областями GaAs. Эти области имеют высокую плотность электронов. Такая конструкция была известна еще под названием *вертикальный транзистор*.

Заметим, что это полупроводниковые приборы на *горячих электронах*. Другими словами, электроны хотя и претерпевают много столкновений, тем не менее обладают высокой скоростью, а следовательно, высокой энергией. Первый такой транзистор был изготовлен из легированного кремния в 1979 году Дж. Шенноном из фирмы Philips. Спустя два года, в 1981 году в Корнельском университете был разработан транзистор на арсениде галлия. Прибор имел два барьера с легированной плоскостью, где происходил перенос горячих электронов.

Масштабы миниатюризации электронных схем, достигнутые сегодня, поражают воображение специалистов даже больше, чем непосвященных в тайны микротехники.

Выдающийся физик современности, лауреат Нобелевской премии 2000 года академик Жорес Иванович Алфров разработал теорию низкоразмерных электронных явлений, а также технологию производства гетероструктур. Эти приборы и устройства широко используются в системах телекоммуникаций, передаче сложных изображений.

А что будет завтра? Каким станет транзистор хотя бы в ближайшие 10 лет?

Прогнозируется, что через 10 лет минимальные топологические нормы составят 0,01 мкм. Исследователи найдут новые технологические решения, позволяющие объединить цифровые и аналоговые методы обработки информации.

Конец начала или начало конца микроэлектроники?

1.1.6. Линии развития, параллельные транзистору

Исследование транзистора, поиск его оптимальных конструкций занимали умы многих исследователей в мире. Транзисторы и другие полупроводниковые приборы стали основой многих исследовательских программ. Изобретательский дух и творческая атмосфера, которые способствовали созданию транзистора, выдвинули электронную промышленность в ряды самых динамичных отраслей. В полупроводниковой электронике, как в фо-

кусе, сконцентрировались достижения физики твердого тела, квантовой механики, кристаллографии, материаловедения и других областей фундаментальных и прикладных наук. Если в 30—40-е годы 90% объема продажи изделий ламповой электроники концентрировалось в основном вокруг радиотехнических устройств, то уже в 50-е годы на радиооборудование приходилось не более 20%. Исследования и разработки новых изделий полупроводниковой электроники все время возрастают, оказывая существенное влияние на прогресс в смежных областях науки и техники. Электронные предприятия ежегодно расширяли номенклатуру своей продукции, рос приток инженеров как в электронную промышленность, так и в исследовательские центры по электронике. Широкий фронт исследований в современной полупроводниковой электронике позволил приложить многим теоретическим изысканиям большую практическую направленность.

В 1930 году советский ученый Яков Ильич Френкель высказал идею, согласно которой при поглощении излучения в кристалле возникают два типа возбуждения: фотоактивный и нефотоактивный. При этом электрон связывался с образованной им дыркой в единую нейтральную систему, которую Френкель Я. И. назвал *экситоном*. В 1952 году было экспериментально доказано существование экситона — квазичастицы, соответствующей электронному возбуждению кристаллов полупроводника или диэлектрика. Выяснилось, что экситон способен мигрировать по кристаллу без переноса электрического заряда и массы.

Было отмечено существование в инерционной поляризующейся среде особого квантового стационарного состояния электрона — *полярона*. На это впервые указали советские физики Лев Дмитриевич Ландау (1933 г.) и Яков Ильич Френкель (1936 г.). Концепция поляронов существенно повлияла на развитие теории полупроводников.

Интересные результаты были получены при изучении воздействия корпускулярного облучения на физические свойства полупроводников, предвосхитившие разработку метода *ионного легирования* и *травления* в микроэлектронике.

Группа исследователей, используя введенный У. Шокли принцип транзистора с ловушкой в коллекторе, в середине 50-х годов предложила конструкцию твердотельного тиратрона путем добавления к транзисторной структуре еще одного р-п-перехода. Такая транзисторная конструкция, полученная на основе р—п—р—п- или п—р—п—р-структур и названная *тиристором*, обладала бистабильными характеристиками и способностью переключаться из одного состояния в другое. В зависимости от способа включения р—п-переходов различали тиристоры диодные (*диинисторы*) и триодные (*триинисторы*). Благодаря двум устойчивым состояниям и низкой мощности рассеяния в этих состояниях тиристоры нашли широкое применение в устройствах для регулирования мощности, в электропреобразовательных высоковольтных устройствах и т. п.

Пожалуй, наиболее интересным случаем, когда фундаментальные исследования привели к появлению полупроводникового прибора, явилось создание *туннельного диода*. Многообразие функций, которые он мог выполнять (генерирование и усиление электромагнитных колебаний, переключение, преобразование частоты и т. д.), по технической значимости поставило создание туннельного диода вровень с открытием транзистора. Работа туннельного диода основана на туннельном эффекте, в соответствии с которым частицы могли с определенной вероятностью проникать через высокопотенциальный барьер. Этот эффект был предсказан в 1939 году русским ученым Георгием Антоновичем Гамовым. Диод на его основе был создан в 1958 году японским физиком Лео Эсаки. Он сумел сформировать чрезвычайно резкий переход между очень сильно легированными р- и п-областями в германии так, чтобы обедненная область в диоде была очень тонкой. Прила-

гая напряженность смещения в прямом направлении, Эсаки обнаружил возрастание суммарного туннельного тока в этом направлении. При увеличении напряжения смещения сверх некоторого значения ток в прямом направлении убывал вследствие уменьшения числа состояний электронов, доступных для туннелирования. Этот эффект эквивалентен возникновению отрицательного сопротивления, которое можно использовать для создания высокочастотных усилителей, генераторов, переключателей. Уже в 1959 году были разработаны туннельные диоды, работавшие на частотах свыше 1 ГГц. В ходе дальнейших исследований Эсаки открыл явление сильного возрастания магнитосопротивления при определенном значении электрического поля (эффект Эсаки), а в 1966 году обнаружил сверхпроводящую энергетическую щель в полупроводниках.

В 1959 году советский ученый Александр Семенович Тагер с сотрудниками открыл явление генерации и усиления СВЧ-колебаний при лавинном пробое полупроводниковых диодов. В этом случае происходит лавинное умножение носителей заряда путем образования пар подвижных носителей при ударной ионизации атомов кристаллической решетки полупроводника подвижными электронами, ускоренными внешним электрическим полем. На основе этого открытия был создан *лавинно-пролетный диод* (ЛПД), представляющий собой полупроводниковый диод с отрицательным сопротивлением в СВЧ-диапазоне. Идея создания ЛПД принадлежит американскому физiku У. Риду (1958 г.), генерацию колебаний впервые осуществил Тагер А. С. с группой сотрудников (1959 г.). В НИИ "Пульсар" под руководством Виктора Михайловича Вальд-Перлова разработано 40 типоминиалов арсенид-галлиевых ЛПД, работающих в диапазоне от 8 до 37 ГГц.

Различают несколько режимов работы ЛПД. Пролетный режим работы основан на использовании лавинного пробоя и пролетного эффекта носителей в обедненной области различных полупроводниковых структур. Этот режим называли также режимом IMPATT (Impact Avalanche Transit Time, ударная ионизация и пролетное время), а диоды, работающие в этом режиме, — *IMPATT-диодами*. Аномальный режим работы ЛПД с захваченной плазмой называли режимом TRAPPAT (Trapped Plasma Avalanche Triggered Transit, захваченная плазма, пробег области лавинного умножения), соответственно этому диоды, работавшие в этом режиме, — *TRAPPAT-диодами*. ЛПД применяли для генерации и усиления в СВЧ-диапазоне на частотах от 10 до 100 ГГц с КПД до 50%.

На базе полупроводниковых параметрических усилительных диодов были созданы усилители сигнала, обладавшие температурой шума в пределах 50—60 К без охлаждения и 25 К и ниже при охлаждении. Они нашли широкое применение в устройствах дальней связи, радиоастрономии, в системах спутниковой связи.

Одним из крупнейших научных достижений физики полупроводников стало создание советским академиком Николаем Геннадиевичем Басовым с сотрудниками полупроводниковых *лазеров* оптического и ближнего инфракрасного диапазонов. Они нашли широкое применение в оптических системах записи, хранения и обработки информации, в системах связи. Весьма перспективным явилось применение полупроводниковых лазеров в волоконно-оптических линиях связи.

В 1963 году американский физик Джон Ганн обнаружил в кристалле арсенида галлия и фосфида индия с электронной проводимостью эффект генерации высокочастотных колебаний тока в сильных электрических полях. Образец полупроводника обладал N-образной вольтамперной характеристикой, имевшей участок с отрицательным дифференциальным сопротивлением. Этот эффект вызван периодическим появлением и перемещением в образце области сильного электрического поля — так называемого *домена Ганна*. На основе этого эффекта СВЧ-генераторы функционировали в диапазоне частот 0,3—2,0 ГГц. Для

этой цели использоваться полупроводники электронного типа GaAs, SnP, CdTe, ZnS, InSb, InAs, а также германий с дырочной проводимостью. Эффект Ганна используется в диодах Ганна для создания генераторов СВЧ-диапазона.

Проблему преобразования световой энергии в электрическую с помощью полупроводников одним из первых поставил советский академик Абрам Федорович Иоффе. основополагающими в области фотоэлектрических свойств полупроводников были работы советских ученых Давыдова Б. И., Курчатова И. В., Лошкарева В. Е., Кушнера Ю. М., Тучкевича В. М., Алферова Ж. И. и др. В 50-е годы был создан германиевый фотоэлемент, работавший в диодном режиме и управлявшийся светом. Разработанные в США и СССР фотодиоды с р—n-переходом позволили существенно повысить КПД фотопреобразователей. Благодаря освоению технологии выращивания монокристаллов кремния были созданы кремниевые ветильные фотоэлементы большой площади со сравнительно большими КПД. Эти приборы нашли широкое применение в наземных и бортовых устройствах непосредственного преобразования солнечной энергии в электрическую.

КПД преобразования кремниевых элементов достиг 10—18%.

Дорогие арсенид-галлиевые приборы выгодно использовать при высокой концентрации света. В таком случае в процессе фотоэлектрического и фототермического преобразований КПД достиг значения 25%. Современной наукой поставлена задача создания установок, способных сделать фотоэлектрические источники конкурентоспособными по отношению к другим видам генераторов.

С 50-х годов стала разрабатываться идея использования р—n-переходов для преобразования ядерной энергии в электрическую. При прохождении через кристалл проводника ионизирующие частицы вдоль своей траектории генерировали электронно-дырочные пары. Эти пары создавали напряжение на р—n-переходе. Созданная на основе этого явления экспериментальная атомная батарея позволяла непосредственно осуществлять преобразование энергии радиоактивного излучения в электрическую энергию. Одновременно были созданы фотопреобразователи на основе сернистого кадмия CdS, чувствительные к рентгеновским лучам и корпускулярному излучению. Они широко использовались в качестве дозиметров рентгеновского излучения и счетчиков α - и β -частиц.

Одним из применений полупроводников явилось детектирование инфракрасного излучения. По своим технико-экономическим характеристикам полупроводниковые детекторы инфракрасного излучения на основе PbS, PbSe, PbTe, InSb существенно превосходили ранее использовавшиеся детекторы теплового излучения. Термовизоры на базе полупроводниковых детекторов инфракрасного излучения, которые начали разрабатываться в 60-х годах, нашли применение в различных областях науки и техники. По мере обработки отдельных узлов улучшились технические характеристики устройств. Применение системы на базе прецизионной оптики и полупроводниковых детекторов излучения со встроенным микропроцессором позволило измерять температуру в диапазоне от 30 до 2000 °C с разрешающей способностью 0,1 °C. Можно получить термальное изображение объекта на экране тепловизора, используя естественное инфракрасное излучение этого объекта. Высокая скорость цифровой памяти позволила хранить и обрабатывать информацию об измеренных значениях температур, выполнять сканирование термически неоднородных поверхностей. В последние годы основными приемниками инфракрасного и субмиллиметрового излучения в диапазоне длин волн от 6 до 500 мкм стали примесные фотоэлектрические приемники на основе кремния, германия, эпитаксиальных пленок арсенида галлия, антимонида индия. Созданы приемники излучения с разогревом носителей тока в полупроводниках. Все они нашли применение для наблюдения источников косми-

ческого излучения, излучения Земли, для изучения распределения яркости Солнца по диску и т. д.

Полупроводники обладают повышенной чувствительностью к воздействию внешних факторов. Температурные, электрические, магнитные, электромагнитные поля, механические деформации зачастую приводят к изменению свойств полупроводниковых приборов за счет изменения их электрических параметров. Эти свойства полупроводников были использованы для создания различных датчиков и приборов на их основе. Разработаны и производятся полупроводниковые резисторы, использующие зависимость электрического сопротивления от внешних факторов: *термисторы*, *фоторезисторы*, *варисторы*, *тензорезисторы*. Промышленные образцы термисторов, меняющих электросопротивление под действием тепловых полей, изготовлены из композиции полупроводников, которые подобраны так, чтобы коэффициент температурной зависимости сопротивления был максимально большим. Термисторы с положительным значением температурного коэффициента получили название *позисторы*, их разработка велась в начале 60-х годов. Термисторы, предназначенные для измерения мощности электромагнитного излучения в видимой, инфракрасной областях спектра, а также в СВЧ-диапазоне, названы *болометрами*.

В ЛФТИ под руководством академика Жореса Ивановича Алферова в семидесятых годах были широко развернуты работы по созданию новых полупроводниковых материалов типа A^3B^5 . Именно эти материалы позволили создать так называемые *гетеропереходы* — полупроводниковый переход между двумя разнородными по химическому составу полупроводниками. В гетеропереходах происходит скачкообразное изменение ширины запрещенной зоны, подвижности носителей, их эффективной массы, энергии электронного сродства.

Нынешние высокие информационные технологии базируются на трех китах:

- ☐ на классической кремниевой технологии;
- ☐ на технологии полупроводниковых гетероструктур;
- ☐ на квантовых полупроводниковых приборах.

Советский физик Жорес Иванович Алферов с сотрудниками Ленинградского физико-технического института и американский физик Герберт Кремер в начале 1963 года начали научное соревнование, целью которого было создание идеальных гетероструктур. На основе этих гетероструктур были созданы лазеры, которые могли работать в непрерывном режиме при комнатной температуре. На их основе были созданы волоконно-оптические системы связи. Эти гетероструктуры легли в основу бортовых и не только солнечных батарей с высоким КПД. *Гетеробиполярные транзисторы* на основе гетеропереходов вошли в состав связной портативной аппаратуры.

Успехи ЛФТИ были тесно связаны с предприятиями электронной промышленности. Это и предопределило успех в таком соревновании. За фундаментальные работы по созданию гетероструктурной электроники академик Алферов Ж. И. получил Нобелевскую премию за 2000 год.

На стыке научных направлений физики полупроводников появились интересные работы. Особенно интересны работы по акустоэлектронике. Здесь был предсказан и экспериментально обнаружен новый тип поверхностных акустических волн — *волны Гуляева — Блюштейна*. Пионером этого направления стал академик Юрий Васильевич Гуляев и его сотрудники. Он провел фундаментальные теоретические и экспериментальные исследования явлений увлечения электронов акустическими волнами, открыл акустоэлектрический и акустомангнетозлектрический эффекты, поперечный акустоэлектрический эффект. Эти разработки на стыке полупроводниковой электроники и акустоэлектроники были широко

использованы при создании акустоэлектронных устройств для систем связи и обработки информации.

Реальные пути стыковки дискретных полупроводниковых приборов с традиционными элементами связи открыла *оптоэлектроника*, представляющая особый раздел науки и техники, посвященный вопросам генерации и приема, а также преобразования и хранения информации на основе сочетания электрических и оптических методов и средств. Оптоэлектроника стала развиваться с 60-х годов на основе достижений квантовой электроники, полупроводниковой электроники, электрооптики голографии, инфракрасной техники и т. д.

На истории развития полупроводниковой и микроэлектроники точку ставить нельзя. Вспреди впечатляющие достижения в нанoeлектронике, одноэлектронике, молекулярной электронике.

Но это уже другая песня.

1.2. Место микроэлектроники в сфере высоких технологий

В ведущих странах мира приоритетное внимание уделяется развитию электроники. В 2000 году объем продаж в этой области составил триллион долларов. Правительства динамично развивающихся стран (США, Франция, Германия) значительные усилия направляют на развитие конкурентоспособной электронной техники.

Другая группа стран, поставивших себе цель занять лидирующие позиции в экономическом росте и росте ВВП (Китай, Малайзия, Испания, Индия), предпринимают колоссальные усилия по созданию собственной электронной промышленности.

С этой целью разрабатываются национальные программы развития электроники. Предусматривается не только прямое государственное финансирование, но также и приоритетная поддержка в виде налоговых льгот, льготных кредитов на закупку технологических линий, государственных гарантий инвесторам, правовой защиты внутреннего рынка от импорта и т. п.

Реальная экономическая ситуация такова, что на рубеже веков роль электроники стала решающей не только в техническом плане, но и в интеллектуальном и, главное, в макроэкономическом аспекте.

Приведем несколько цифр. Экспорт отечественной продукции состоит практически из сырья и оценивается в 36 млрд долларов в год. Из них 14,5 млрд долларов приносит экспорт газа, 14 млрд долларов — экспорт сырой нефти, 4 млрд долларов — металл и 3,5 млрд — военная техника. Прибыль обычно составляет четверть объема экспорта.

А вот другие цифры.

Годовые обороты американских электронных фирм, выпускающих изделия микроэлектроники, составляют порядка 200 млрд долларов. Это в 7 раз больше российского экспорта сырьевых энергоресурсов. Суммарный объем продаж электронных приборов и систем на порядок больше.

Мировое потребление электронной продукции растет и увеличивается на 15% в год, а в странах Азии и Тихоокеанского региона на 19% в год.

В 2003 году объем мирового производства полупроводниковой промышленности достигает рубежа 300 млрд долларов. За три года в три раза!

Технологический уровень современного микроэлектронного производства определяется несколькими основными параметрами:

- ☐ классом чистоты технологических помещений;
- ☐ диаметром обрабатываемых пластин;
- ☐ топологическими параметрами.

Мировой уровень сегодня таков, что класс чистоты технологических помещений близок к 1, диаметр обрабатываемых кремниевых пластин 200—300 мм, а топологические нормы на уровне 0,35 мкм.

В России такого производства нет. И отечественное электронное машиностроение отсутствует.

Можно купить готовый завод стоимостью 1,2—2,0 млрд долларов, позволяющий создавать интегральные схемы с топологическими параметрами 0,35—0,25 мкм, мощностью 20—25 тысяч 8-дюймовых пластин в месяц. Но на такой завод нет денег.

Текущие заказы российских заводов не превышают 20 млн долларов в год — десятая доля экономически выгодного производства. А развитие бытовой электроники... Все можно выгодно купить за рубежом! Так считают те, кто реформирует страну.

Так ли это?

Опыт восточных стран, ныне преуспевших в микроэлектронном производстве, показал, что выпуск на первых порах неконкурентоспособной продукции позволяет развивать собственное производство, отработать технологию и превзойти конкурентов. Современная микроэлектроника находится на переломном этапе технологического развития и дорог каждый день.

Речь идет о тенденции дезинтеграции интегральной электроники (микроэлектроники). Все годы интенсивного развития микроэлектроники в мире доминировали фирмы, опирающиеся на полный цикл производства продукции: разработка, производство и продажа. Такая экономико-технологическая модель позволила корпорациям Intel, IBM, Motorola, Texas Instruments, Micron, NEC, Toshiba, Samsung и другим наладить массовое производство полупроводниковых приборов для товаров массового потребления.

Интегральные схемы специального применения традиционно проектировались так называемыми fabless-фирмами, а изготавливались на кремниевых заводах других компаний. Fabless-фирмы способны разработать сложные интегральные схемы в интересах любого заказчика.

Таким образом, возникла новая экономико-технологическая модель, разработанная интеллектуальная собственность в виде IP-блока, которая продвигается до конечного потребителя за несколько этапов через кремниевое и приборное производство изделий и систем.

Сегодня успешно работает такая экономико-технологическая модель, как fabless-фирма — кремниевый завод. Так 50% ИС под торговым знаком Motorola производится на тайваньских кремниевых заводах.

Фирмой Intel, микропроцессоры которой широко используются во всем мире, создано новое отделение Intel Microelectronics Services. Ее основной задачей является разработка стандартных специализированных ИС с топологическими нормами 0,25 и 0,13 мкм.

В виду того, что собственное производство Intel загружено микропроцессорами и другими стандартными изделиями, производство продукции fabless-фирмы Intel будет проводиться на другом кремниевом заводе. Это либо завод с незагруженными мощностями

крупной фирмы, либо "чистый" завод, построенный специально для реализации интеллектуальной собственности фирм-разработчиков интегральных схем.

Наметилась тенденция увеличения как числа "чистых" кремниевых заводов, так и увеличения числа fabless-фирм.

Для России это шанс встать в ряд технологически и интеллектуально развитых стран, создать национальный банк интеллектуальной собственности IP-блоков, получить значительный экономический эффект от реализации изделий микроэлектроники и электронной компонентной базы и, наконец, достигнуть технологической независимости от иностранных государств.

Без интенсивных усилий по созданию собственного электронного производства Россия:

- ☐ не сможет увеличить потребление изделий электроники, в том числе средств вычислительной техники;
- ☐ не сможет рассчитывать на увеличение экспорта дорогостоящей высокотехнологичной продукции вместо сырья;
- ☐ не сделает экономику страны эффективной;
- ☐ не решит проблемы информатизации страны.

В основах политики РФ в области развития электронной компонентной базы (ЭКБ) в электронных системах, имеющих стратегическое значение для национальной безопасности, предусмотрено использовать только те, которые спроектированы в России и изготовлены на зарубежных технологических линиях только по отечественным фотошаблонам.

В соответствии с программой "Электронная Россия" к 2010 году предусмотрен переход к сквозному технологическому циклу отечественной разработки и производства отечественной ЭКБ на технологическом уровне не хуже 0,1 мкм.

Вот тогда и понадобятся молодые специалисты в области микроэлектроники.

2. Физика полупроводниковых структур

2.1. Полупроводники и их структура

Полупроводники представляют собой широкий класс материалов с электронным механизмом проводимости, в которых концентрация подвижных носителей заряда ниже концентрации атомов, но может меняться под действием температуры, освещения, небольшого количества примесей.

По удельному значению электропроводности полупроводники занимают промежуточное положение между металлами ($10^4 < \sigma \leq 10^6 \text{ Ом}^{-1} \cdot \text{см}^{-1}$) и диэлектриками ($\sigma \geq 10^{-12} \text{ Ом}^{-1} \cdot \text{см}^{-1}$) и лежат в диапазоне $10^3 < \sigma \leq 10^9 \text{ Ом}^{-1} \cdot \text{см}^{-1}$.

Одной из главных отличительных свойств полупроводниковых материалов является возрастание электропроводности с ростом температуры. В широком диапазоне температур электропроводность экспоненциально растет с температурой T по закону

$$\sigma = \sigma_0 \exp\left(\frac{-\epsilon_A}{kT}\right), \quad (2.1)$$

где σ_0 — начальное значение электропроводности, ϵ_A — энергия активации проводимости, которая соответствует энергии связи электронов с атомами; k — постоянная Больцмана.

Полупроводники можно классифицировать по различным признакам, например:

- ☐ по агрегатному состоянию: твердые и жидкие;
- ☐ по структуре: кристаллические и некристаллические;
- ☐ по физическим свойствам: магнитные и сегнетоэлектрические;
- ☐ по химическому составу: элементарные, соединения, органические.

В микроэлектронике в основном используются твердотельные кристаллические структуры, состоящие из элементарных полупроводников (Ge, Si) или полупроводниковых соединений типа Ge-Si, A^3B^5 (GaAs, InSb), A^2B^6 (CdS) и др.

Основным требованием к полупроводниковым материалам, используемым в микроэлектронике, является их бездефектность или малое количество дефектов. Поэтому основным материалом для применений являются монокристаллические полупроводники — твердые тела с регулярной кристаллической структурой.

Кристаллическая структура состоит из множества повторяющихся и примыкающих друг к другу элементарных ячеек определенного размера. В элементарной ячейке частицы занимают строго фиксированные позиции, находясь на определенном расстоянии друг от друга. В результате взаимодействия электронов внешних оболочек атомов в кристалле возникает химическая связь.

По типу химических связей различают четыре основные группы кристаллических структур. В ионных или гетерополярных кристаллах преобладает ионный (электростатический)

характер связи между атомами, возникающий вследствие перехода электронов от одного атома к другому.

В ковалентных или гомеополярных кристаллических структурах валентные электроны соседних атомов обобществляются, образуя двойные или тройные связи между атомами. К такому типу кристаллов относятся алмаз, кремний, карборунд.

В металлических кристаллических структурах, относящихся к третьему типу химических связей, свободные электроны распределяются по всей кристаллической решетке, образуя электронный газ.

В молекулярных кристаллах атомы в молекуле прочно связаны, в то время как сами молекулы между собой связаны слабо. Такая связь характерна для органических соединений.

Кремний, например, имеет кубическую гранецентрированную структуру типа алмаза с постоянной решетки $a = 0,3571 \text{ нм}$ (рис. 2.1, а).

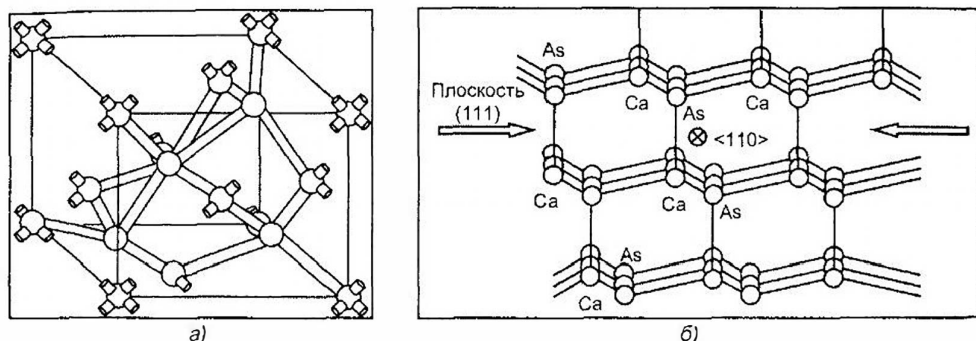


Рис. 2.1. Структура гранецентрированных кристаллических решеток типа алмаза (а) и типа цинковой обманки (б)

Для арсенида галлия характерна классическая структура типа цинковой обманки — минерала сфалерита типа ZnS (рис. 2.1, б). Кристаллическая структура типа цинковой обманки имеет специфические особенности. При разделении кристалла по плоскости (111) в направлении стрелок в верхнем слое располагаются атомы Ga, а в нижнем — атомы As. Это свойство широко используется в процессах при эпитаксиальном выращивании и травлении.

Структура кристаллической решетки имеет дефекты и дислокации.

Различают точечные дефекты в виде пустого узла (вакансия) или междоузельного атома, а также примесные дефекты в виде примеси внедрения или примеси замещения.

Дислокации бывают линейные (краевые) и винтовые (спиральные). Помимо дислокаций и дефектов, в полупроводниковых кристаллах могут иметь место микротрещины, поры, пузырьки и т. д. Все эти неоднородности кристаллов приводят к браку при производстве интегральных схем. Поэтому получению бездефектных полупроводниковых структур уделяется большое внимание.

При наличии дефектов на поверхности кристалла у приповерхностных атомов кристалла нарушаются ковалентные связи из-за отсутствия следующих слоев атомов. Нарушение ковалентных связей приводит к нарушению энергетического равновесия на поверхности. Это может привести к захвату чужеродных атомов из окружающей среды — адсорбции

или к частичному восстановлению оборванных связей и образованию, например, окислов. Такая поверхность не может быть использована в микроэлектронном производстве.

Структура тонкого приповерхностного слоя резко отличается от структуры основного объема кристалла. Граничные слои играют важнейшую роль при создании интегральных схем.

Структура полупроводникового материала может изменяться искусственно по нужному алгоритму. Изменение структуры путем внедрения примесных атомов приводит к целенаправленному изменению проводимости полупроводников. Технологически такое изменение может осуществляться путем высокотемпературной диффузии или ионной имплантации. Целенаправленное локальное изменение проводимости полупроводниковой структуры легло в основу производства интегральных схем.

2.2. Носители заряда в полупроводниках

Электропроводность полупроводников обусловлена двумя типами носителей электрического заряда, которые могут перемещаться под действием градиента концентрации или внешнего электрического поля.

Носителями заряда в полупроводниках являются электроны проводимости и дырки.

Электрон проводимости является наименьшим носителем отрицательного электрического заряда $q = -1,60217733 \times 10^{-19}$ Кл. Масса покоя электрона составляет $m_q = 0,91093897 \times 10^{-30}$ кг.

Масса электрона проводимости может меняться при движении в полупроводниковой кристаллической структуре и поэтому направление ускорения электрона в общем случае не параллельно внешней силе. Вводится понятие эффективной массы электрона

$$m^* = p_0 / v_0,$$

где p_0 и v_0 — абсолютные значения импульса и скорости, соответствующие энергии Ферми.

Дырка представляет собой квазичастицу или незаполненное электронное состояние (вакансию) в валентной зоне полупроводника. Понятие дырки было введено для удобства описания физических свойств полупроводников. Дырке приписывают положительный заряд, по величине равный заряду электрона. Эффективная масса дырки обычно больше, чем у электрона. Подвижность дырок при движении в полупроводниковой структуре меньше, чем у электронов проводимости.

Беспримесный и бездефектный полупроводник с идеальной кристаллической решеткой называется *собственным* полупроводником. Его проводимость называется *собственной* проводимостью. Идеальный собственный полупроводник при температуре, равной абсолютному нулю, обладает свойствами изолятора. По мере повышения температуры в кристалле возникает колебательное движение атомов решетки.

Квант колебаний атомов кристаллической решетки называют *фононом*.

Термин был введен советским физиком И. Е. Таммом по аналогии с квантом электромагнитного поля — фотоном. Фонон является квазичастицей, представляющей собой квант энергии упругих колебаний кристаллической решетки. Энергия фонона E_{ϕ} определяется величиной

$$E_{\phi} = \hbar \omega,$$

где ω — частота колебаний, \hbar — постоянная Планка, равная $h / 2\pi$.

С повышением температуры кристалла количество и энергия фононов возрастает. Фононы способны разорвать ковалентные связи между атомами решетки. Это приводит к одновременному возникновению свободных электронов и незаполненных связей — дырок.

Процесс образования электронно-дырочных пар под действием фононов называется *термогенерацией*.

Возможен и другой механизм создания электронов или дырок под действием квантов энергии. Речь идет о γ -квантах и рентгеновских лучах. Рождение электронно-дырочных пар возникает по всей глубине проникновения таких квантов энергии.

Проводимостью собственного полупроводника можно управлять. С этой целью вводят примеси, которые могут локально изменять тип проводимости полупроводника.

Рассмотрим два случая внедрения примеси.

Если в четырехвалентный атом кремния ввести атом пятивалентного элемента из V группы таблицы Менделеева, то четыре валентных электрона из пяти свяжутся с четырьмя электронами атома кремния. Образуется устойчивая оболочка из восьми электронов, а примесный атом Р превратится в неподвижный положительный ион. Оставшийся свободным электрон добавится к собственным свободным электронам. Такие полупроводники называются *электронными* полупроводниками, или полупроводниками *n-типа* (рис. 2.2, а).

Во втором случае внедрим в структуру четырехвалентного кремния элемент из III группы таблицы Менделеева, например атом бора. В этом случае все три валентных электрона вступят в связь с четырьмя электронами соседних атомов кремния. Дополнительный электрон для образования устойчивой восьмиэлектронной оболочки будет заимствован у ближайшего атома кремния. Таким образом, образуется незаполненная связь, или дырка. Атом примеси превратится в неподвижный ион с отрицательным зарядом (рис. 2.2, б). Дырки примесного происхождения добавятся к собственным дыркам, а полупроводник станет полупроводником *p-типа*, или полупроводником с *дырочной проводимостью*.

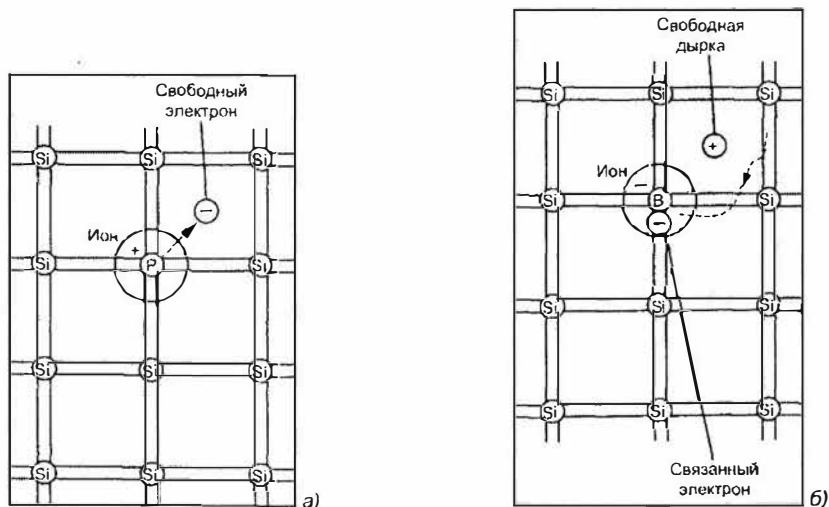


Рис. 2.2. Процесс замещения примесными атомами и формирования области с электронной (а) и дырочной проводимостью (б)

Для получения полупроводников электронного типа из кремния используются обычно элементы пятой группы: фосфор, сурьма, мышьяк. Полупроводники дырочного типа получают внедрением примеси третьей группы: бор, галлий, алюминий.

Примеси в полупроводниках электронного типа называют *донорными*, а в полупроводниках дырочного типа — *акцепторными*.

В GaAs донорами являются элементы шестой группы — сера и теллур, а акцепторами — элементы второй группы — бериллий и цинк. Существуют так называемые *амфотерные* примеси. Они могут быть как донорной, так и акцепторной в зависимости от условий внедрения. Кремний, вводимый в GaAs методом ионной имплантации, является донором, а германий, введенный методом жидкостной эпитаксии, — акцептором.

В примесных полупроводниках концентрация электронов и дырок измеряется в их количестве в объеме кубического сантиметра. Носители преобладающего типа проводимости называются *основными*, а другого типа — *неосновными*.

В полупроводнике п-типа проводимости основными носителями являются электроны, в полупроводнике р-типа проводимости — дырки.

Поведение электронов в монокристаллическом полупроводнике определяется не только корпускулярными свойствами электрона, но и его волновыми свойствами — волновой функцией электрона Ψ .

Изменение во времени состояния квантовых объектов, характеризующихся волновой функцией, описывается уравнением Шредингера. Это уравнение является математическим выражением фундаментального свойства микрочастиц — корпускулярно-волнового дуализма.

Для изолированного атома справедливо уравнение Шредингера

$$\mathbf{H} \cdot \Psi = E \cdot \Psi, \quad (2.2)$$

где $\mathbf{H} = -\hbar^2 / 2m_0 \Delta + V(x, y, z)$ — гамильтониан, $V(x, y, z)$ — потенциальная энергия электрона в кристалле, E — энергия уровня электрона.

Величина $\Psi \cdot \Psi^*$ называется *функцией плотности вероятности*, где Ψ^* — комплексно-сопряженная величина. Физический смысл выражения $\Psi \cdot \Psi^* \cdot dV$ определяет вероятность появления или нахождения электрона в объеме dV .

В этом случае квазиимпульс электрона \vec{p} определяется как $\hbar \vec{k}$, где \hbar — постоянная Планка, \vec{k} — волновой вектор электрона.

В свою очередь энергия электрона E связана с его массой соотношением:

$$E = \hbar^2 \vec{k}^2 / 2m. \quad (2.3)$$

В идеальном монокристалле атомы расположены регулярно, причем расстояние между ними равно шагу решетки a (порядка нескольких ангстрем). Атомы состоят из положительно заряженных ядер и электронов, обладающих эквивалентным по величине отрицательным зарядом.

Так одиночный атом кремния имеет:

- два электрона на атомной орбитали $1s - 2s^2$;
- два электрона на атомной орбитали $2s - 2s^2$;
- шесть электронов на атомной орбитали $2p - 2p^6$;
- два электрона на атомной орбитали $3s - 3s^2$;
- два электрона на атомной орбитали $3p - 3p^2$.

В кристалле наибольшее воздействие соседние ядра атомов оказывают на электроны внешней оболочки $3s^2$ и $3p^2$. Все эти орбитали в кристалле взаимосвязаны и образуют смешанные орбитали, характерные для кристаллов ковалентного типа. Электроны некоторых внешних оболочек вращаются не только вокруг собственного ядра, но и вокруг ядер соседних атомов. При этом на одной смешанной орбитали находятся по два электрона от каждого соседнего атома или восемь электронов.

Орбитали внешних атомных оболочек в кристалле кремния имеют замкнутую форму, которая, образуя ковалентные межатомные связи, охватывает соседние атомы. Электроны, движущиеся по этим орбиталиям, называются *валентными*.

На рис. 2.3 показаны расчетные распределения плотности электрического заряда валентных электронов (изогипсы) для наиболее распространенных в микроэлектронике кристаллов Si и GaAs.

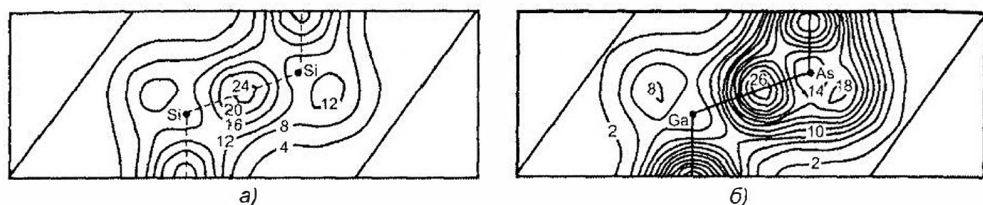


Рис. 2.3. Изогипсы распределения валентных электронов в кремнии (а) и арсениде галлия (б). Цифры указывают число электронов

В кристалле же количество атомов близко к значению 10^{22} см^{-3} (или 10^{28} м^{-3}), причем каждый из них обладает собственной системой энергетических уровней.

Рассмотрим процесс изменения энергетических уровней по мере сближения двух атомов А и В, в каждом из которых электрон находится в основном $1s$ -состоянии.

По мере сближения атомов их волновые функции перекрываются и возникают две комбинации функций: $\Psi_A + \Psi_B$ (рис. 2.4, а).

Состояние, показанное на рис. 2.4, б, характеризуется тем, что электрон "проводит" часть времени в середине между точками А и В. В этой области он находится под влиянием поля притяжения одновременно обоих атомов, что увеличивает энергию связи. Напротив, в положении, отображенном на рис. 2.4, в, плотность вероятности в серединной точке обращается в нуль и добавки к энергии связи не возникает.

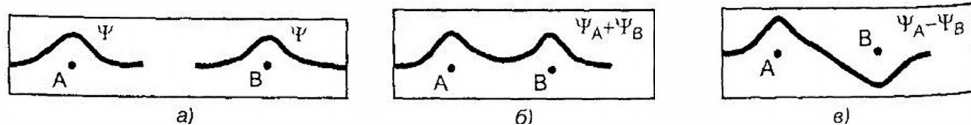


Рис. 2.4. Изменение волновых функций электронов в двух водородных атомах при сближении атомов: а — исходное состояние; б — линейная комбинация волновых функций при сближении; в — линейная комбинация, описывающая возбужденное состояние

Система из многих атомов, каждый из которых имеет свои уровни, образует несколько разрешенных и запрещенных зон.

По мере сближения свободных атомов, кулоновское взаимодействие между атомными ядрами и электронными оболочками приводит к расщеплению энергетических уровней.

Образуется система энергетических зон. Ширина зоны пропорциональна интенсивности взаимодействия и степени перекрытия оболочек соседних атомов.

На рис. 2.5 показаны энергетические уровни в системе из шести водородных атомов в процессе их сближения. Уровни ранее изолированных атомов расщепляются.

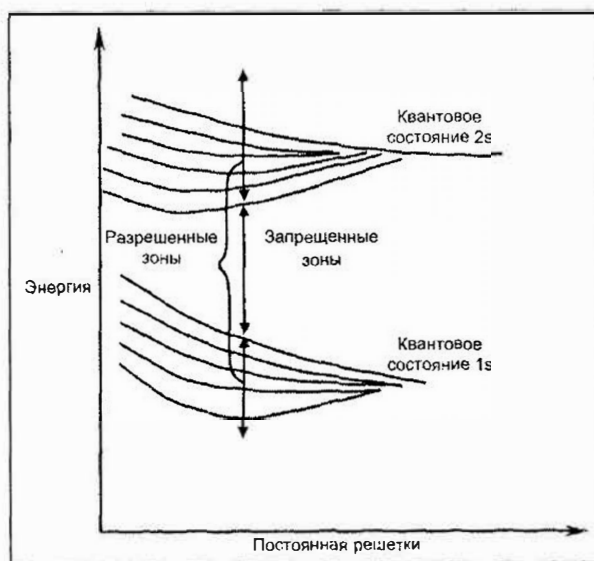


Рис. 2.5. Образование разрешенных и запрещенных зон в системе из шести водородных атомов при их сближении

Количественный анализ полупроводников базируется на зонной теории твердого тела, в соответствии с которой твердое тело характеризуется совокупностью энергетических зон. Верхняя зона, разрешенная для ее заполнения электронами, называется *зоной проводимости* E_c , а нижняя зона — *валентной зоной* E_v . Между ними находится *запрещенная зона* E_g , ширина которой зависит от температуры:

$$E_g = E_{g0} - \epsilon_g \cdot T,$$

где E_{g0} — ширина зоны при $T \approx 0$ К; T — температура; ϵ_g — температурная чувствительность, которая для кремния составляет 3×10^{-4} В/°С. Ширина запрещенной зоны для кремния при комнатной температуре равна $E_g = 1,11$ эВ, а для GaAs — 1,42 эВ.

Увеличение энергии электрона соответствует поднятию электрона на внешние энергетические уровни. Напротив, увеличение энергии дырки соответствует снижению на более низкий энергетический уровень.

На рис. 2.6 приведены зонные диаграммы с донорной примесью замещения (а) и с акцепторной примесью замещения (б).

Полные концентрации для примесного проводника будут соответствовать:

для донорного $n = n_n + n_i$ и $p = p_n$

для акцепторного $n = n_i$ и $p = p_n + p_a$,

где n_i, p_i — концентрация электронов и дырок, обусловленная возбуждением собственного полупроводника, n_n, p_n — концентрация электронов и дырок, образовавшихся вследствие

вне возбуждения донорных и акцепторных примесей. В обычном случае справедливо следующее соотношение:

$$n_n \gg n_i, p_p \gg p_i.$$



Рис. 2.6. Зонная диаграмма с донорной (а) и акцепторной примесью (б)

В полупроводнике при неизменной температуре произведение концентрации электронов и дырок является постоянной величиной.

Существует классическое соотношение:

$$n_p = n_i p_i = n_i^2$$

и поэтому увеличение, например, концентрации электронов приводит к уменьшению концентрации дырок.

Электроны в твердом теле представляют собой идеальный квантовый газ, подчиняющийся статистике Ферми — Дирака для элементарных частиц с полуцелым спином-фермионов. В соответствии с распределением Ферми — Дирака, число частиц в состоянии с энергией E_i при температуре T :

$$\langle n_i \rangle = \frac{1}{\exp\left(\frac{E_i - \chi}{kT}\right) + 1}, \quad (2.4)$$

где i — набор квантовых чисел, характеризующих состояние частицы; k — постоянная Больцмана; χ — химический потенциал. В полупроводниках значение χ соответствует энергии в центре запрещенной зоны — электрическому потенциалу ϕ_F .

Энергию, соответствующую середине запрещенной зоны невырожденного полупроводника, называют уровнем Ферми

$$\phi_F = \frac{1}{2}(E_c + E_v).$$

Уровень Ферми E_F определяют как потенциал, вероятность заполнения которого электроном равна 0,5.

Для идеального газа фермионов, например электронов, уровень Ферми совпадает с химическим потенциалом при $T = 0$ К.

В полупроводнике *n*-типа проводимости концентрация электронов в зоне проводимости больше, чем у собственного полупроводника, и уровень Ферми будет расположен выше середины ближе к донорному уровню.

В полупроводнике *p*-типа проводимости концентрация дырок в валентной зоне будет выше, чем у собственного полупроводника. Уровень Ферми в таком полупроводнике будет расположен ниже середины запрещенной зоны, т. е. ближе к акцепторному уровню.

Одним из фундаментальных положений в физике полупроводников является постулат:

уровень Ферми одинаков во всех частях равновесной системы, какой бы разнородной она не была.

В таком случае справедливы соотношения:

$$E_F = \text{const}, \text{grad}(E_F) = 0. \quad (2.5)$$

2.3. Перенос носителей и управление переносом

В полупроводниках существуют два основных механизма переноса носителей:

- диффузия носителей заряда;
- дрейф носителей заряда под действием внешнего электрического поля.

Диффузия носителей заряда представляет собой направленное перемещение носителей в кристалле в сторону уменьшения их концентрации.

Процесс диффузии ведет к выравниванию неравномерного распределения неравновесных носителей заряда по объему кристалла. Различают *монополярную* диффузию (или диффузию носителей одного заряда) и *биполярно-совместную* диффузию электронов и дырок.

В одномерном случае следует, что если концентрация электронов изменится вдоль координаты x , то возникает электрическое поле $E_{\text{вн}} = \frac{kT}{n} \cdot \frac{dn}{dx}$.

В неоднородно легированных полупроводниках смещение подвижных носителей или их диффузия уравнивается возникновением *встроенного внутреннего электрического поля* $E_{\text{вн}}$.

В неоднородно легированных полупроводниках смещение подвижных носителей или их диффузия уравнивается возникновением *встроенного внутреннего электрического поля* $E_{\text{вн}}$.

Дрейф носителей в кристалле имеет хаотический характер. На рис. 2.7 интерпретированы процессы рассеяния и дрейфа электрона в полупроводниковом кристалле, которые начались в точке O .

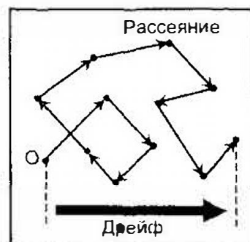


Рис. 2.7. Процессы рассеяния и дрейфа носителей

В черных точках электрон рассеивается и меняет вектор скорости. Рассеивание носит случайный характер.

Дрейф носителей заряда представляет собой упорядоченное движение носителей заряда под действием внешнего электрического поля. Электрический ток, обусловленный дрейфом носителей заряда, называется *дрейфовым*.

Плотность дрейфового тока j определяется соотношением:

$$j = \sigma E,$$

где σ — удельная проводимость, E — электрическое поле.

Поскольку в полупроводнике имеется два типа носителей, то удельная проводимость имеет две компоненты, т. е.

$$\sigma = \sigma_n + \sigma_p,$$

где $\sigma_n = qn\mu_n$ — электронная, а $\sigma_p = qp\mu_p$ — дырочная составляющие. Здесь μ_n и μ_p — подвижности соответствующих носителей заряда, q — заряд носителей, n и p — соответствующие концентрации носителей заряда.

Подвижностью носителей заряда является величина, характеризующая динамические свойства носителей заряда и определяется как отношение средней скорости направленного движения носителей (дрейфовая скорость) к напряженности электрического поля. Подвижность представляет собой многократное повторение следующих фаз процесса: ускорение носителей электрическим полем, их рассеяние на дефектах кристалла или на фононах, изменение вектора движения, повторное ускорение и т. д. Поэтому чем меньше масса носителей, тем более высокую подвижность они имеют. Численно подвижность определяется соотношением:

$$\mu = \frac{q\tau}{m^*},$$

где τ — среднее время пробега между двумя циклами рассеяния, q — заряд электрона, m^* — эффективная масса заряженной частицы.

Подвижность имеет размерность $\text{см}^2/(\text{В} \cdot \text{с})$ и при напряженности поля $E = 1 \text{ В/см}$ значение подвижности численно равно скорости носителя: $\mu = V$.

Таким образом, в полупроводнике движение носителей заряда обусловлено диффузией под воздействием градиента концентрации и дрейфом под воздействием градиента электрического поля.

Полный ток определяется выражением:

$$j = (j_n)_{др} + (j_n)_{диф} + (j_p)_{др} + (j_p)_{диф},$$

где индексы "др" и "диф" относятся к дрейфовым и диффузионным составляющим тока электронов j_n и тока дырок j_p .

Диффузионные составляющие токов можно записать:

$$(j_n)_{диф} = q \cdot \mu_n kT \cdot \frac{dn}{dx} = qD_n \frac{dn}{dx},$$

$$(j_p)_{диф} = -q\mu_p kT \frac{dp}{dx} = -qD_p \frac{dp}{dx},$$

где коэффициенты D_n и D_p называются коэффициентами диффузии электронов и дырок, соответственно, и определяются в соответствии с формулой Эйнштейна:

$$D = \frac{kT}{q} \mu.$$

Коэффициент диффузии имеет размерность $[\text{см}^2/\text{с}]$. При комнатной температуре для германия и арсенида галлия коэффициенты диффузии равны, соответственно: $D_{\text{Ge}} \approx 90 \text{ см}^2/\text{с}$, $D_{\text{GaAs}} \approx 220 \text{ см}^2/\text{с}$. Для кремния коэффициент диффузии составляет всего лишь $38 \text{ см}^2/\text{с}$. Диффузионная длина L связана со временем жизни следующим соотношением:

$$L = \sqrt{D\tau}. \quad (2.6)$$

Дрейфовые составляющие токов можно записать в виде:

$$(j_n)_{\text{др}} = qn\mu_n E = -qn\mu_n \frac{\partial U}{\partial x},$$

$$(j_p)_{\text{др}} = qp\mu_p E = qp\mu_p \frac{\partial U}{\partial x}.$$

В сильных электрических полях происходит разогрев носителей тока. Энергия, получаемая носителями от электрического поля, не успевает рассеяться тепловыми фононами и температура носителей оказывается существенно выше температуры решетки. В этом случае говорят о горячих носителях, например, о горячих электронах.

2.4. Барьеры и контакты

2.4.1. Барьеры на границе кристалла

В соответствии с законом диффузии оценим время, за которое электроны могли бы покинуть кристалл:

$$t_{\text{дифф}} \approx \frac{L^2}{D},$$

где L — характерный размер кристалла, D — коэффициент диффузии. С помощью соотношения Эйнштейна коэффициент диффузии, например, для золота оценивается величиной $0,75 \text{ см}^2/\text{с}$; при подвижности носителей $\mu \approx 30 \text{ см}^2/\text{В}$ время составит $t_{\text{дифф}} \approx 1,3 \text{ с}$. Однако все электроны в этом случае остаются в кристалле. Это объясняется тем, что на границе раздела твердое тело — вакуум существует потенциальный энергетический барьер, препятствующий выходу электронов в окружающую среду. Чтобы преодолеть потенциальный барьер, электрон должен обладать энергией, превышающей значение работы выхода электрона из твердого тела.

Работа выхода представляет собой энергию, которая затрачивается при возбуждении электронов для их вывода из твердого тела в вакуум. Работа выхода электронов из твердого тела для различных веществ колеблется в пределах от 1 до 6 эВ, и существуют разные методики ее определения.

На работу выхода электронов из полупроводников сильно влияют явления на границе раздела "полупроводник—вакуум".

Кристалл полупроводника характеризуется регулярной структурой, которая нарушается ближе к границе раздела. Резко нарушаются и условия связи между валентными электронами, принадлежащими приповерхностным атомам. Обрыв кристаллической решетки способствует появлению дополнительных энергетических уровней. Эти поверхностные состояния получили название *уровней*. Такие уровни характерны для атомарно чистой поверхности, на которой отсутствуют посторонние атомы.

В реальных условиях поверхность всегда покрыта слоем адсорбированных атомов, окислов и т. п. На поверхности всегда существуют структурные дефекты: искажения решетки, вакансии и т. д.

Совокупность или спектр поверхностных состояний реальной поверхности можно менять, изменяя окружающие условия и (или) способы обработки поверхности.

В зависимости от того, какой тип поверхностных состояний имеет место (донорный или акцепторный), соответствующим образом будет заряжаться и поверхность твердого тела. Если на поверхности полупроводника *n*-типа проводимости преобладают акцепторные состояния, то поверхность будет захватывать электроны из объема полупроводника, прилегающего к поверхности, и поверхность в этом случае будет заряжена отрицательно. В приповерхностной области образуется слой, обедненный электронами, и, соответственно, заряженный положительно.

Таким образом, на поверхности полупроводника возникает двойной заряженный слой. Поле этого слоя будет препятствовать выходу электронов из кристалла.

На рис. 2.8 представлена энергетическая диаграмма полупроводника *n*-типа. Поле двойного слоя максимально вблизи поверхности раздела и уменьшается по мере удаления вглубь кристалла.

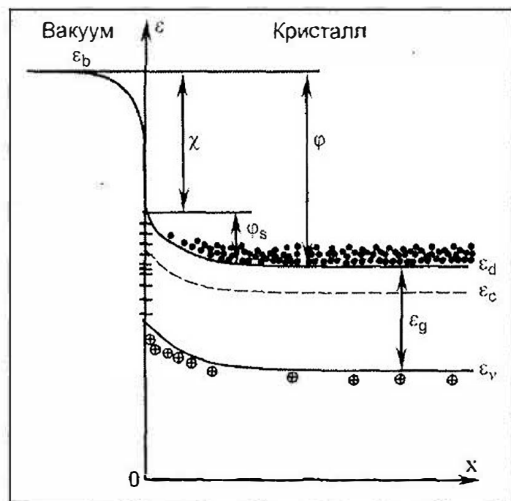


Рис. 2.8. Энергетическая диаграмма границы раздела "полупроводник *n*-типа—вакуум" при отрицательном заряде на поверхностных состояниях. Черточки на границе раздела соответствуют энергетическому положению поверхностных уровней в запрещенной зоне. E_d — объемный донорный уровень

На диаграмме видно, что зоны изгибаются: электроны отталкиваются от границы раздела, в то время как дырки притягиваются. Величину ϕ , называют поверхностным потенциалом, ϕ — работой выхода электронов из кристалла, а χ — электронным сродством.

Рассмотрим случай, когда на поверхности полупроводника *n*-типа будут преобладать донорные поверхностные состояния.

Отдав электроны, такие состояния будут заряжены положительно и будут притягивать электроны из объема кристалла. В то же время дырки будут отталкиваться и уходить вглубь кристалла. На поверхности кристалла возникнет слой, обогащенный электронами. Работа выхода электронов из такого кристалла будет меньше, чем если бы поверхностные состояния отсутствовали.

На рис. 2.9 представлена энергетическая диаграмма границы раздела "полупроводник р-типа—вакуум" с положительным суммарным зарядом поверхностных состояний.

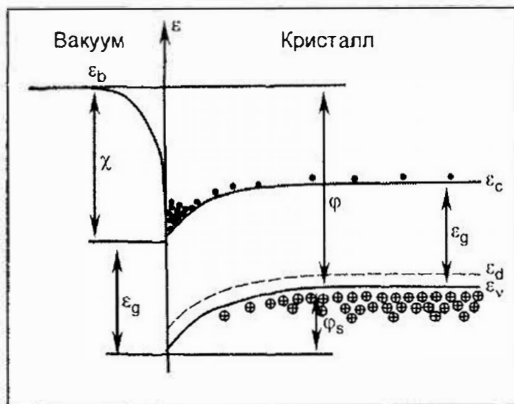


Рис. 2.9. Энергетическая диаграмма границы раздела "полупроводник р-типа—вакуум" с положительным суммарным зарядом поверхностных состояний

В этом случае зоны изгибаются вниз. Наблюдается приток электронов к поверхности и, соответственно, отток дырок вглубь кристалла.

Работа выхода электронов или энергия, затрачиваемая на возбуждение и вывод электрона из полупроводника в вакуум, зависит от типа проводимости и уровня легирования полупроводника, а также от спектра поверхностных состояний. Отрицательный заряд поверхностных уровней увеличивает работу выхода электронов, а положительный — уменьшает ее.

2.4.2. Электронно-дырочные переходы

На границе р- и n-областей создается энергетический барьер, который является основой всех полупроводниковых элементов и компонентов.

Полупроводник р-типа проводимости представляет собой отрицательно заряженные акцепторы, неподвижно закрепленные в кристаллической решетке и положительно заряженные дырки, способные переносить заряды и формировать электрический ток.

Полупроводник n-типа проводимости, напротив, содержит положительно ионизированные доноры, неподвижно закрепленные в решетке и отрицательно заряженные электроны проводимости (рис. 2.10, а).

Если соединить полупроводники р- и n-типов проводимости, то произойдет ряд физических процессов.

Концентрации носителей в полупроводниках различного типа проводимости отличаются на несколько порядков. Различают n^+ —р- или р'—n-переходы, где верхний индекс "+" соответствует слою со значительно большей концентрацией.

Рассмотрим переход p^+ —n в равновесном состоянии, когда внешнее электрическое поле отсутствует. Поскольку концентрация дырок в p^+ -области выше, чем концентрация электронов в n-области, то возникнет процесс диффузии дырок из р-области в n-область и соответствующий ток диффузии $I_{диф.р}$. Одновременно начнется диффузия электронов в р-область и возникнет ток диффузии $I_{диф.н}$ (рис. 2.10, б). Встречная диффузия приведет к появлению в n-области нескомпенсированных положительных зарядов ионов донорной

примеси. Одновременно в р-области возникнет отрицательный заряд ионов акцепторной примеси. Донорные и акцепторные атомы жестко связаны с кристаллической решеткой. В результате у границы р- и п-областей образуется двойной электрический слой пространственного заряда. Поле этого двойного электрического слоя создает потенциальный барьер, препятствующий дальнейшей диффузии дырок в п-область, а электронам в р-область.

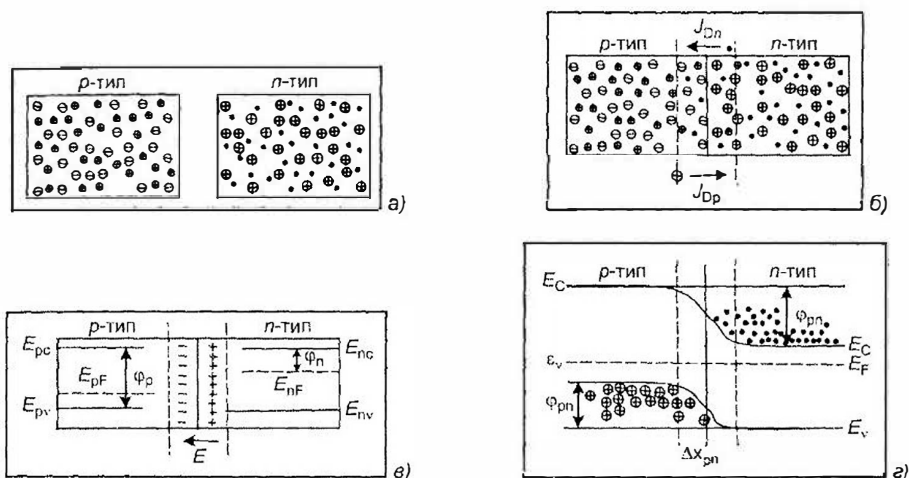


Рис. 2.10. Этапы формирования потенциального барьера на границе р—п-перехода

В то же время в приконтактной зоне происходит интенсивная рекомбинация носителей заряда. Слой обедняется носителями заряда. Когда процесс рекомбинации завершится, то в обедненном слое установится термодинамическое равновесие, и суммарный положительный пространственный заряд будет в точности равен суммарному отрицательному заряду (рис. 2.10, в), другими словами, диффузионные токи станут равными $J_{\text{диф},n} \equiv J_{\text{диф},p}$.

Разность потенциалов, возникающая между р- и п-областями сдвигает электрические уровни в областях на величину, равную разности уровней Ферми в полупроводниках различного типа проводимости (рис. 2.10, г). В такой ситуации электрону из п-области, чтобы попасть в р-область, необходимо преодолеть разность потенциалов (электрический барьер), равную ϕ_{pn} . Такой же барьер нужно преодолеть и дырке, чтобы попасть в п-область.

Значение

$$\phi_{pn} = \phi_p - \phi_n \approx E_g,$$

где E_g — ширина запрещенной зоны.

Таким образом, на границе контакта полупроводников различного типа проводимости образуется р—п-переход. Толщина Δx_{pn} двойного электрического слоя перехода составляет 10^{-4} — 10^{-5} см, а напряженность встроенного электрического поля $E_0 \approx -\text{grad}(\phi_{pn})$.

Если к р—п-переходу приложить внешнее электрическое поле, то система придет в неравновесное состояние. Повышение или понижение потенциального барьера за счет внешнего электрического поля приводит к возникновению тока диффузии и дрейфа.

рассмотрим два случая.

Приложим к р—п-переходу электрическое поле таким образом, чтобы плюс внешнего источника приходился на р-, а минус — на п-области. В этом случае говорят о *прямом включении р—п-перехода*, а напряжение называется *прямым смещением* (рис. 2.11, а). Ток, протекающий через р—п-переход, называется *прямым*. Внешний источник с постоянным напряжением U_{np} создает в р—п-переходе электрическое поле E_{en} противоположного направления относительно внутреннего поля E_0 р—п-перехода.

Результирующее поле в р—п-переходе ослабляется, и потенциальный барьер ϕ снижается на величину U_{np} . В этом случае распределение потенциала в р—п-переходе по координате x показано сплошной линией, а штриховой изображено распределение потенциала в состоянии равновесия.

Под действием поля в р-область попадают дополнительные электроны, а в п-область — соответственно дырки. Этот процесс носит название *инжекции* носителей заряда.

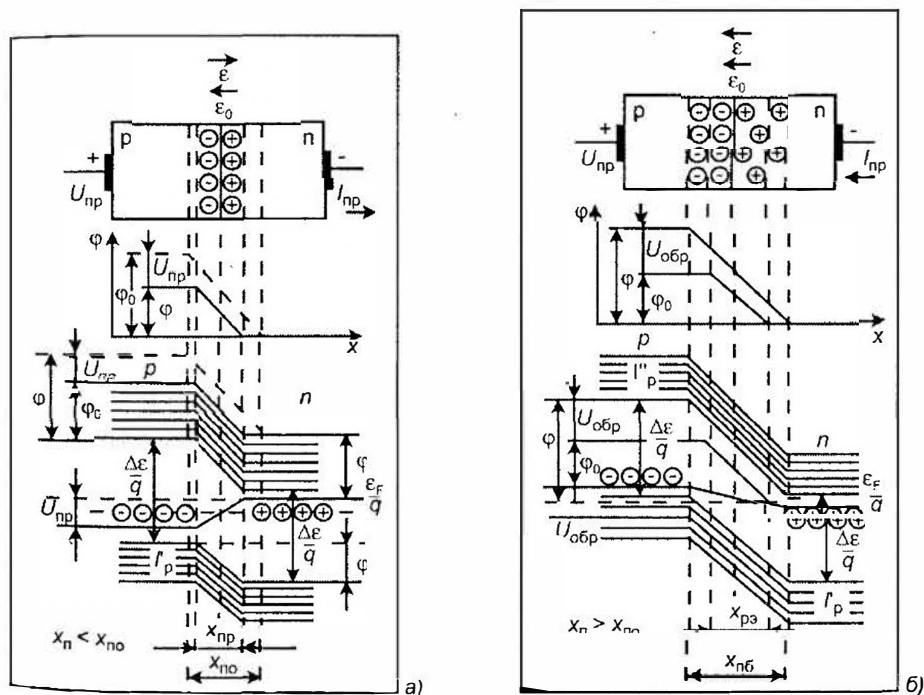


Рис. 2.11. Изменение потенциала и зона диаграмма р—п-перехода, включенного в прямом (а) и обратном (б) направлениях

Одновременно от внешнего источника напряжения через омические контакты в области р- и п-типа попадут равные количества основных носителей. Они нейтрализуют инжектированные заряды. Неосновные носители диффундируют вглубь р- и п-областей. В р—п-переходе при увеличении приложенного напряжения возрастание тока происходит по экспоненциальному закону (рис. 2.12).

Будучи инжектированы в полупроводник с другим типом проводимости, электроны и дырки превращаются в неосновные носители, свойства которых определяются их време-

нем жизни τ и диффузионной длиной $L = \sqrt{D\tau}$. Эффективное время жизни составляет примерно 0,1—5,0 нс. Энергетическая диаграмма р—п-перехода при прямом включении приведена на рис. 2.11, а.

Уровень Ферми в р-области смещен относительно п-области на величину внешнего прямого напряжения $U_{пр}$.

Ширина р—п-перехода при подаче прямого напряжения $U_{пр}$ уменьшается.

Диффузионный ток электронов из п-области в р-область обозначим I'_n , а ток дырок из р-области в п-область — I'_p . Рассмотрим случай, когда к р—п-переходу приложено *обратное напряжение*, т. е. плюс к п-области, а минус к р-области. В этом случае ток через переход называется *обратным током* $I_{обр}$.

Внешний источник создаст в р—п-переходе электрическое поле E , совпадающее по направлению с полем перехода (рис. 2.11, б). В этом случае поле в р—п-переходе усилится, а высота потенциального барьера увеличится на величину $U_{пр}$. Пунктирной линией обозначен потенциал в идеальном переходе $\phi = f(x)$.

Ток определяется диффузией через р—п-переход неосновных носителей, возникающих в результате тепловой генерации в р- и п-областях вблизи перехода.

При небольших обратных смещениях все термически генерированные вблизи р—п-перехода носители попадают в область, где они становятся основными. Дальнейшее увеличение внешнего напряжения приводит к росту тока проводимости за счет генерации носителей в р—п-переходе, размеры которого также увеличиваются.

В процессе *экстракции* неосновных носителей заряда формируется ток, обусловленный дрейфом носителей. Процесс экстракции или вытягивания носителей в контакт с металлом или другим полупроводником приводит к обеднению полупроводника носителями заряда.

Величина дрейфового тока значительно меньше тока диффузии в прямом направлении и выходит на насыщение (рис. 2.12).

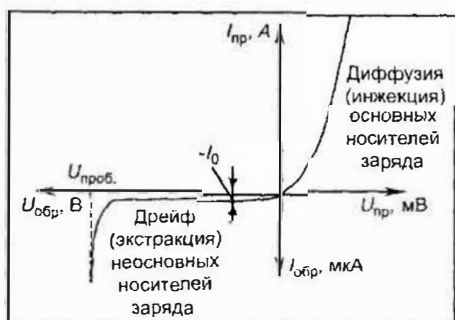


Рис. 2.12. Вольтамперная характеристика р—п-перехода

Зонная диаграмма показывает, что уровень Ферми в р-области будет выше уровня Ферми в п-области на величину $U_{обр}$.

Ток дрейфа неосновных носителей из р-области в п-область обозначим I''_n , а ток дрейфа дырок из п-области в р-область — I''_p .

Полный ток инжекции (диффузии) электронов I''_n определяется как:

$$I''_n = SqL_n n_{p0} / \tau_n \cdot [\exp(U / \phi_T) - 1],$$

где n_{p0} — равновесная концентрация электронов в р-области; S — площадь р—п-перехода; q — заряд; τ_n — время жизни электронов; L_n — диффузионная длина, определяемая как $L_n = \sqrt{D_n \tau_n}$, где D_n — коэффициент диффузии электронов.

Ток инжекции дырок I_p' определяется аналогично:

$$I_p' = Sq L_p p_{n0} / \tau_p \cdot [\exp(U / \Phi_T) - 1],$$

где p_{n0} — равновесная концентрация дырок в п-области, L_p — диффузионная длина дырок, τ_p — время жизни дырок.

Полный ток через р—п-переход составляет:

$$I = I_n + I_p = I_0 [\exp(U / \Phi_T) - 1]. \quad (2.7)$$

Эта величина представляет собой прямую ветвь вольтамперной характеристики идеального р—п-перехода (рис. 2.12). Величина I_0 определяется геометрией р—п-перехода, степенью легирования материала и параметрами полупроводника:

$$I_0 = Sq \cdot [(D_n / L_n) \cdot n_{p0} + (D_p / L_p) \cdot p_{n0}]. \quad (2.8)$$

Уравнение (2.7) описывает зависимость тока через переход от приложенного напряжения и является *вольтамперной характеристикой перехода (ВАХ)*.

При обратном напряжении обратный ток дрейфа I_n'' и I_p'' достигает значения I_0 и остается практически постоянным и не зависит от напряжения до определенного значения равного *напряжению пробоя* $U_{проб}$. При приложении обратного напряжения выше $U_{проб}$ возникает пробой перехода, вызванный лавинным размножением носителей. Пробой может происходить как в объеме, так и по поверхности р—п-перехода. Кроме электрического пробоя в р—п-переходе может произойти и тепловой пробой.

Отличительным свойством р—п-переходов является образование *диффузионной* и *барьерной* емкостей. При приложении прямого напряжения $U_{пр}$ к р—п-переходу будет наблюдаться изменение заряда в обеих областях перехода. С ростом напряжения на р—п-переходе будет увеличиваться взаимная диффузия основных носителей через переход и, соответственно, ток через переход, а также заряд от $Q_{расч}$ до $Q_{реал}$.

В этом случае изменение заряда можно выразить следующим образом:

$$dQ_p = Q_{реал} - Q_{расч} = C_{Dp} \cdot dU_{пр},$$

где C_{Dp} — коэффициент пропорциональности, который и называется *диффузионной емкостью*.

Аналогичный процесс диффузии будет происходить в п-области перехода. Диффузионной емкостью будет величина, определяемая соотношением:

$$dQ_n = Q_{реал} - Q_{расч} = C_{Dn} \cdot dU_{пр}.$$

Общее выражение для диффузионной емкости примет вид:

$$C_D = C_{Dp} + C_{Dn}$$

или

$$dQ_p = dI_p \tau_p$$

$$dQ_n = dI_n \tau_n$$

где τ_p, τ_n — время жизни дырок и электронов, соответственно.

Из уравнения ВАХ р—п-перехода при прямом включении следует:

$$dI_p / dU_{пр} = I_p / \Phi_T \text{ и } C_{Dp} = (I_p \tau_p) / \Phi_T, \quad C_{Dn} = (I_n \tau_n) / \Phi_T.$$

Если диффузионная длина L много меньше толщины области W (р- или п-) или $W \gg L$, то

$$C_D = q(\tau_p I_p + \tau_n I_n) / kT.$$

Если $\tau_p = \tau_n$ и $W \ll L$, то время пролета носителей t определяется как

$$t = W^2 / D_p.$$

Окончательная формула для значения диффузионной емкости примет вид:

$$C_D = \ln \frac{q}{kT} \cdot \frac{W^2}{2D_p}. \quad (2.9)$$

Процесс диффузии через р—п-переход связан с образованием разности потенциалов ϕ_0 , а также с образованием ионов доноров и акцепторов, жестко привязанных к решетке. Образовавшийся потенциальный барьер характеризуется барьерной емкостью, величина которой может быть описана формулой плоского конденсатора:

$$C_n = \epsilon \epsilon_0 S / x_n,$$

где x_n — ширина р—п-перехода; ϵ — относительная диэлектрическая проницаемость; ϵ_0 — электрическая постоянная; S — площадь р—п-перехода.

В связи с зависимостью ширины перехода от внешнего приложенного напряжения U_{np} существует следующая зависимость:

$$C_{\bar{\sigma}} = C_{\bar{\sigma}0} \sqrt{\phi_0 / (\phi_0 + U_{np})},$$

где $C_{\bar{\sigma}0}$ — барьерная емкость р—п-перехода в равновесном состоянии.

При увеличении обратного напряжения $U_{обр}$, приложенного к переходу, его ширина увеличивается. В этом случае дырки в р-области и электроны в п-области под действием поля уходят от границы р—п-перехода. Линии тока внутри переходного слоя замыкаются через токи смещения, образуя непрерывную линию зарядного тока. В такой ситуации р—п-переход ведет себя как плоский конденсатор, емкость которого определяется как

$$C_{\bar{\sigma}} = \epsilon \epsilon_0 S / x_n,$$

где x_n — ширина р—п-перехода в равновесном состоянии.

Соответственно ток через конденсатор составит величину:

$$I_{\bar{\sigma}} = C_{\bar{\sigma}} dU_{об} / dt.$$

Изменение ширины р—п-перехода от приложенного напряжения определяется зависимостью:

$$x_n = x_{n0} (1 - U / \phi_T)^i,$$

где x_{n0} — ширина перехода в равновесном состоянии; U — напряжение, приложенное к р—п-переходу; i — показатель степени ($i = 1/2$ для резкого, $i = 1/3$ для плавного переходов).

Окончательная зависимость барьерной емкости от напряжения примет вид:

$$C_{\bar{\sigma}} = C_{\bar{\sigma}0} (1 - U / \phi_T)^{-i}. \quad (2.10)$$

Из всего вышесказанного можно сделать два вывода:

1. При прямом включении р—п-перехода основным процессом является процесс перемещения диффузионных зарядов, определяемый диффузионной емкостью ($C_D \gg C_{\bar{\sigma}}$).
2. При обратном включении р—п-перехода главную роль играет барьерная емкость, которая отражает перераспределение зарядов в р—п-переходе ($C_{\bar{\sigma}} \gg C_D$).

2.4.3. Гетеропереходы

Гетеропереходом называют переход, образующийся на границе двух полупроводников с различной шириной запрещенной зоны.

Гетеропереход может быть образован как двумя монокристаллическими (аморфными) полупроводниками, так и монокристаллическим и аморфным полупроводниками.

На границе гетероперехода происходит изменение свойств полупроводника, в частности, меняется структура энергетических зон, ширина запрещенной зоны, подвижности и эффективные массы носителей заряда.

Различают анизотипные и изотипные гетеропереходы. *Анизотипные переходы* создаются в результате контакта полупроводников с дырочной и электронной типами проводимости. *Изотипные переходы* возникают в результате контакта полупроводников одного типа проводимости.

Комбинации различных гетеропереходов образуют *гетероструктуры*.

На рис. 2.13 представлены диаграммы электрических зон для идеального (а) и резкого (б) р-п-гетеропереходов.

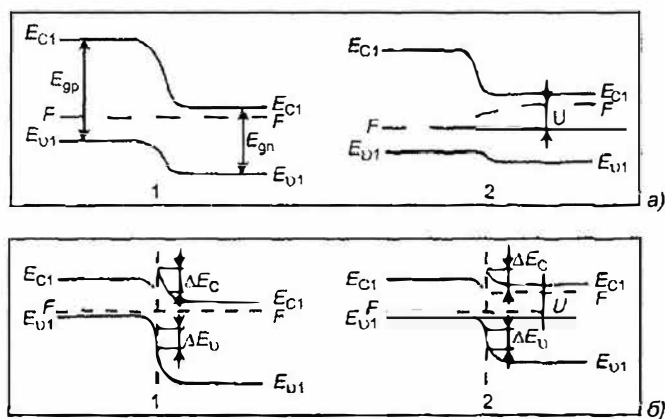


Рис. 2.13. Энергетическая диаграмма для идеального (а) и резкого (б) р-п-гетеропереходов в случае равновесия (слева) и положительного смещения (справа)

Структура гетероперехода, у которого материал р-типа имеет большую ширину запрещенной зоны, чем материал п-типа, т. е. $E_{gp} > E_{gn}$, представлен на рис. 2.13, а (слева). Это идеальный гетеропереход в условиях термодинамического равновесия. Если к нему приложить прямое внешнее напряжение U , то потенциальный барьер, который должны преодолеть дырки при переходе из р- в п-область, уменьшится (рис. 2.13, а, справа). С увеличением напряжения барьер может практически исчезнуть, что приведет к резкому возрастанию дырочного тока.

Совсем иная картина имеет место для электронов, которые желают преодолеть барьер при переходе из п- в р-область. Потенциальный барьер для электронов в этом случае достаточно велик и поэтому электронный ток в прямом направлении мал.

При создании гетеропереходов из полупроводников с различными параметрами (ширина запрещенной зоны, диэлектрическая проницаемость ϵ , работа выхода электронов и элек-

тронное сродство χ) формируется зонная диаграмма, представленная на рис. 2.13, б. Здесь структура, расположенная слева, соответствует состоянию термодинамического равновесия.

Электрическое поле на границе раздела имеет разрыв, который обусловлен различием диэлектрических проницаемостей. Энергетические зоны также имеют разрыв на границе раздела, образуя ступени E_c и E_v . При подаче прямого внешнего электрического поля барьер для электронов будет меньше, чем барьер для дырок (рис. 2.13, б, справа). В этом слое доминирующим будет ток электронов.

Помимо приведенных типов гетеропереходов существует также ряд специальных гетеропереходов, которые представляют значительный практический интерес.

На рис. 2.14, а представлена структура р-п-гетероперехода, которая позволяет эффективно инжектировать дырки в материал п-типа, не требуя намеренно сильного легирования области р-типа. Область р-типа с широкой запрещенной зоной прозрачна для рекомбинационного излучения из области р-типа. В этой же области не наблюдается поглощение света свободными носителями заряда. Такие гетеропереходы используются при создании, например, полупроводниковых лазеров. Штрихом показана область — так называемый *квантовый колодец*, в котором находится двумерный электронный газ.

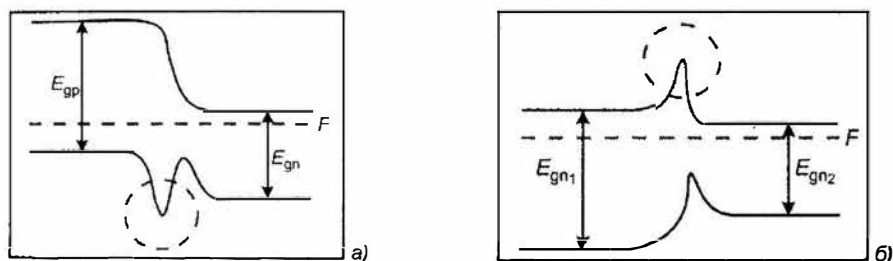


Рис. 2.14. Зонная диаграмма гетероперехода с квантовыми колодцами

Двумерный электронный газ (ДЭГ) представляет собой систему электронов, энергетические уровни которых дискретны и их движение финитно.

Другими словами, в поперечном направлении потенциальная энергия электронов не позволяет им покинуть потенциальную яму, а их соответствующие энергетические уровни дискретны. Таким образом, движение электронов возможно только в плоскости ДЭГ.

Свойства двумерного газа определяются возможностью регулировать и менять в широких пределах плотность электронного газа под действием поперечного электрического поля. Для электронов в области ДЭГ характерна высокая подвижность порядка 9×10^3 см²/В·с, близкая к объемной подвижности электронов в нелегированном GaAs.

Если гетеропереход получен из веществ с различной постоянной решетки, то на границе двух полупроводников могут возникнуть механические дефекты, которые будут играть роль ловушек для дырок и электронов. Со стороны п-типа появится подъем зон, в то время как со стороны р-типа возникнет их понижение (рис. 2.14, б).

В таких гетероструктурах формируются квантовые колодцы для обоих типов носителей заряда.

В табл. 2.1 приведены некоторые параметры полупроводников, образующих гетеропереходы.

Таблица 2.1. Параметры полупроводников, образующих гетеропереходы

Гетеро-переход	Полупроводник	E_g , эВ, при 300 К	Параметр решетки, мм	Коэффициент линейного расширения при 300 К, $10^{-6} \cdot K^{-1}$	Энергия электронного сродства, эВ	Относительная диэлектрическая постоянная, ϵ_r
GaAs-Ge	GaAs	1,43	0,5653	5,8	4,07	11,5
	Ge	0,67	0,5658	5,7	4,13	16,0
AlSb-GaSb	AlSb	1,6	0,6136	3,7	3,65	10,3
	GaSb	0,68	0,6095	6,9	5,06	14,8
AlAs-GaAs	AlAs	2,15	0,5661	5,2	—	—
	GaAs	1,43	0,5653	5,8	4,07	11,5
GaP-Si	GaP	2,25	0,5451	5,3	4,3	3,4
	Si	1,11	0,5431	2,33	4,01	12
ZnSe-Ge	ZnSe	2,67	0,5667	7,0	4,09	9,1
	Ge	2,66	0,5658	5,7	4,13	16,0
ZnSe-GaAs	ZnSe	2,67	0,5669	7,0	4,09	9,1
	GaAs	1,43	0,5653	5,8	4,07	11,5

2.4.4. Контакты

Исторически первыми полупроводниковыми приборами стали диоды на основе контакта "полупроводник—металл". Различают контакты "полупроводник—металл" двух типов: омические контакты или не выпрямляющие и выпрямляющие контакты.

Тип контакта определяется взаимным расположением уровней Ферми в обоих веществах.

Омическим контактом или омическим переходом называется физический контакт, электрическое сопротивление которого мало и не зависит от направления тока в заданном диапазоне значений токов.

В таких контактах отсутствует инжекция неосновных носителей заряда, их удельное сопротивление меньше $10^5 \text{ Ом} \cdot \text{см}^2$. Большинство омических переходов создается на основе $n-n^+$ или p^+-p -переходов. Концентрация легирующей примеси в сильно легированном слое должна быть достаточно высокой, чтобы между металлом и n^+ -полупроводником создать объединенную область. Толщина этой области такова, что переход носителей через потенциальный барьер обеспечивается с помощью механизма туннельного эффекта. Из-за низкой концентрации дырок в выраженном n^+ -слое их инжекция в слабо легированную n -область будет отсутствовать.

Улучшению свойств омического контакта служит шлифовка полупроводника в месте его контакта перед металлизацией. Возникшие в процессе шлифовки дефекты кристаллической решетки работают как центры рекомбинации. В этом случае возникает равновесная концентрация основных и неосновных носителей вблизи поверхности полупроводника.

На границе раздела металла с полупроводником возникает выпрямляющий контакт. Если работа выхода из полупроводника ϕ_p больше работы выхода из металла ϕ_m , то уро-

вень Ферми металла E_{Fm} располагается выше, чем уровень Ферми полупроводника E_{Fp} (рис. 2.15, а).

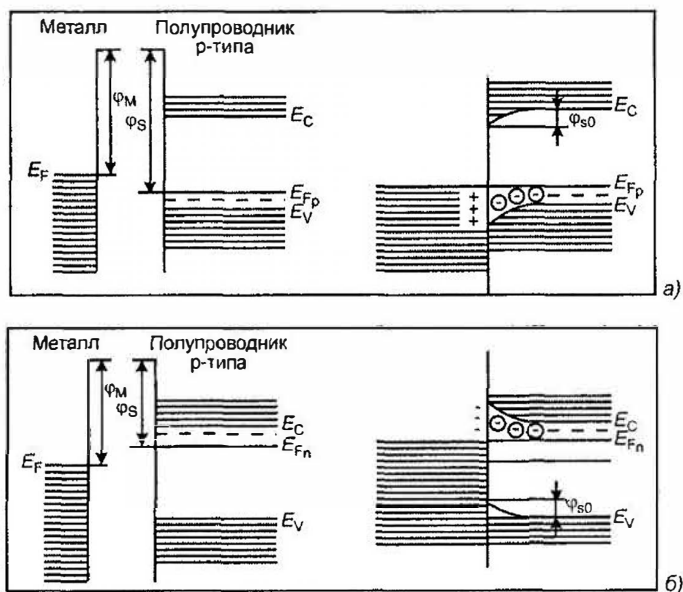


Рис. 2.15. Зонные диаграммы выпрямляющих контактов металла с полупроводником р-типа (а) и п-типа (б)

В момент соприкосновения металла с полупроводником последний заряжается отрицательно, и в результате его энергетические зоны в приконтактной области искривляются вниз. Число электронов в зоне проводимости увеличивается, в то время как число дырок в валентной зоне уменьшается. В результате рекомбинации обнажаются некомпенсированные отрицательные ионы акцепторов, образуется электрическое поле, которое препятствует дальнейшему притоку электронов из металла. В итоге образуется *барьер Шоттки* ϕ_{s0} , представляющий собой потенциальный барьер в приконтактном слое.

Если работа выхода из полупроводника меньше работы выхода из металла, то после соприкосновения электроны полупроводника п-типа переходят в металл (рис. 2.15, б). В приконтактной области полупроводник заряжается положительно, и его энергетические зоны искривляются вверх. Дно зоны проводимости удаляется от уровня Ферми, а потолок валентной зоны, наоборот, приближается к уровню Ферми. Вблизи контактной области концентрация электронов убывает, а концентрация дырок в валентной зоне возрастает по сравнению с этими значениями в глубине полупроводника.

Приконтактный слой полупроводника характеризуется пониженной удельной проводимостью. Этот слой обогащен неосновными носителями заряда и носит название *антизатормозного*. В этом случае также возникает барьер Шоттки ϕ_{s0} .

В зависимости от полярности приложенного внешнего напряжения потенциал и сопротивление приконтактного слоя будут изменяться. Так если положительное напряжение приложено к металлу, а отрицательное — к полупроводнику, то высота потенциального барьера в контакте увеличивается (рис. 2.15, а). Приконтактный слой обедняется основ-

ными носителями — дырками. Сопротивление этого слоя возрастает по сравнению с равновесным. В этом случае говорят, что приложенное напряжение для данного контакта — *обратное*.

Напротив, в системе "металл—полупроводник" *n*-типа с ростом внешнего напряжения потенциал понижается (рис. 2.15, б). Приконтактный слой обогащается основными носителями — электронами, и его сопротивление уменьшается. В этом случае напряжение с такой полярностью будет *прямым*. Такие контакты, как "металл—полупроводник", которые обладают выпрямляющими свойствами, называются *контактами Шоттки*.

Для *n*-Si барьер Шоттки составляет 0,6—0,8 эВ, для *p*-Si — 0,4—0,6 эВ, для *n*-GaAs — 0,8 эВ, а для *p*-GaAs — 0,6 эВ.

2.4.5. Граница раздела

При создании интегральных схем граница раздела "полупроводник—диэлектрик" играет весьма важную роль. Рассмотрим границу раздела "кремний—диоксид кремния" (*Si* — *SiO₂*).

Свойства среды, с которой граничит полупроводник, оказывают определяющее влияние на свойства поверхностного слоя, его кристаллическую структуру, содержание адсорбированных примесей и наличие особых энергетических уровней. Все это влияет на подвижность и время жизни носителей в приповерхностном слое и другие электрофизические параметры.

Главная особенность слоев или пленок диоксида кремния состоит в том, что они всегда содержат примеси донорного типа (натрий, калий, водород), которые имеют тенденцию локализоваться вблизи границы раздела *Si* — *SiO₂*. В результате на границе с кремнием формируется тонкий слой положительно заряженных донорных атомов (рис. 2.16, а). Отдавшие ими электроны переходят в приповерхностный слой кремния. Поверхностная концентрация доноров в двуокиси кремния составляет $\sim (0,5\text{—}2,0) \times 10^{12} \text{ см}^{-2}$. Если пленка *SiO₂* находится на поверхности *n*-типа *Si*, то приповерхностный слой обогащается основными носителями и у границы раздела образуется *n*-канал.

Если кремний обладает проводимостью *p*-типа, то электроны, диффундировавшие из окисла, могут привести к обеднению приповерхностного слоя вследствие рекомбинации с дырками и обнажить отрицательные ионы акцепторов (рис. 2.16, б).

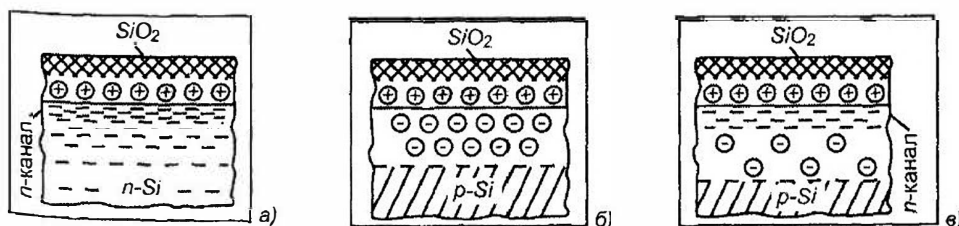


Рис. 2.16. Приповерхностные структуры на границе раздела *Si* — *SiO₂*: а — обогащенный слой; б — обедненный слой; в — обедненный слой с инверсионным каналом

В кремнии *p*-типа помимо обедненного слоя также возможно образование поверхностного *n*-слоя (рис. 2.16, в).

Знание структуры границы раздела *Si* — *SiO₂* позволяет предотвращать нарушения и дефекты в приборах микроэлектроники.

Задачи и упражнения

2.1. Положение уровня Ферми

Дайте элементарное описание зонной структуры энергетических уровней электрона в металле и объясните смысл уровня Ферми. Постройте зависимость концентрации электронов в зоне проводимости металла от энергии при различных температурах. Вычислите, на какой высоте (в эВ) от дна зоны проводимости находится уровень Ферми E_F в натрии, который содержит $2,53 \times 10^{22}$ атомов/см³. Можете предположить, что плотность энергетических уровней в зоне проводимости определяется выражением

$$S(E) = 2^{1/2} m^{3/2} \pi E^{1/2} / h^3.$$

Решение

Энергетические уровни в одиночном изолированном атоме являются строго дискретными. Под влиянием обменных процессов с соседними атомами энергетические уровни трансформируются. Сдвиг внутренних уровней очень незначителен, в то время, как внешние сдвигаются в большей степени. На простейшем примере атома водорода показано, как при построении кристаллической решетки твердого водорода дискретные энергетические уровни свободно изолированного атома водорода превращаются в энергетические зоны. В твердом водороде нет электронов в зоне проводимости и поэтому зазор между зоной проводимости и валентной зоной велик (рис. 2.17).

На рис. 2.18 показаны зависимости концентрации электронов от энергии при различных температурах.

Число энергетических уровней в единице объема определяется выражением

$$S(E) = \frac{2^{7/2} m^{3/2} E^{1/2} \pi}{h^3}. \quad (2.1.1)$$

Умножив это выражение на распределение Ферми — Дирака, найдем действительное число электронов

$$N(E)dE = \frac{2^{7/2} m_e^{3/2} \pi}{h^3} \left\{ \frac{1}{\exp \left[(E - E_F) / kT \right] + 1} \right\} E^{1/2} dE. \quad (2.1.2)$$

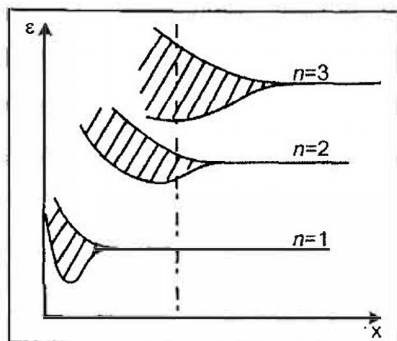


Рис. 2.17. Дискретные энергетические уровни атома водорода в зависимости от расстояния x

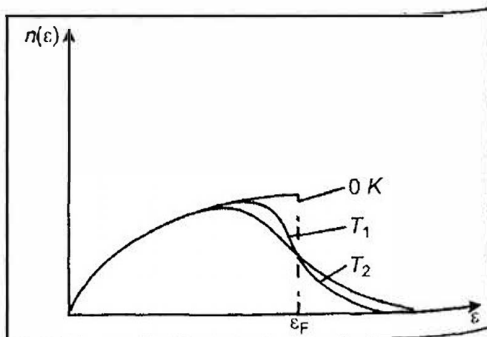


Рис. 2.18. Распределение концентрации электронов по энергиям для разных температур

При $T = 0$ К функция распределения Ферми — Дирака (в фигурных скобках) равна единице. Поэтому число электронов, заполняющих все состояния вплоть до уровня Ферми, определяется интегралом

$$\int_0^{E_F} N(E) dE.$$

Интегрируя уравнение (2.11), находим

$$N(E) = \frac{(2 \cdot 2^{7/2} \cdot m_e^{3/2} \pi)}{3h^3} E_F^{3/2},$$

откуда после преобразования получаем

$$E_F = \frac{h^2}{2m_e} \left[\frac{3N(E)}{8\pi} \right]^{2/3}.$$

Для определенной концентрации электронов имеем

$$E_F = \frac{h^2}{2m_e} \left[\frac{3 \cdot 2.53 \cdot 10^{28}}{8\pi} \right]^{2/3} \text{ Дж.}$$

Подставляя значения констант, получаем, что уровень Ферми E_F или верхний энергетический уровень при температуре 0 К равен 3,14 эВ.

2.2. Уровень Ферми в германии n-типа

Дайте краткое качественное описание процесса протекания электрического тока в полупроводнике n-типа и объясните, как образуются носители заряда. Определите положение уровня Ферми в германии n-типа при температуре 300 К, если на 1×10^6 атомов германия приходится один атом примеси. Обоснуйте сделанные при этом допущения.

Концентрация атомов в германии равна $4,4 \times 10^{28}$ атомов/м³. Константа в выражении, связывающем число электронов в единице объема в зоне проводимости с температурой и энергетическими уровнями, равна $4,83 \times 10^{21} \text{ м}^{-3} \cdot (\text{К})^{-3/2}$. Ширина запрещенной зоны равна 0,72 эВ, а расстояние между дном зоны проводимости и донорным уровнем — 0,01 эВ.

Решение

Электропроводность чистого полупроводника, такого как германий или кремний, может быть увеличена незначительным добавлением соответствующего элемента V группы периодической системы Менделеева. Так, в кремний обычно добавляют фосфор, а в германий — мышьяк или олово. Поскольку концентрация этих добавок обычно порядка одного атома на 10^6 , их называют примесными атомами в чистом полупроводнике, или легирующими элементами.

Примесные атомы имеют пять валентных электронов, тогда как в атомах чистого германия их всего четыре. При комнатной температуре почти все дополнительные электроны являются свободными. Такие примесные атомы называют донорами, и легируемый ими материал имеет проводимость n-типа, поскольку в нем образуется дополнительное количество подвижных отрицательных носителей зарядов.

Энергетическая диаграмма полупроводника n-типа показана на рис. 2.19.

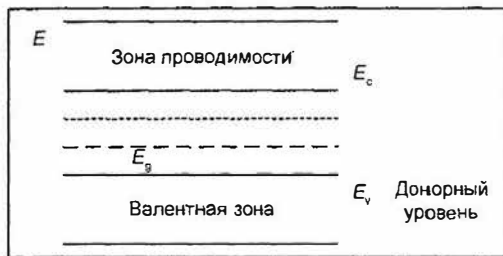


Рис. 2.19. Зонная диаграмма германия n-типа

Каждый донорный атом создает энергетический уровень, расположенный чуть ниже дна зоны проводимости. Согласно заданию этот уровень в германии располагается на расстоянии 0,01 эВ от дна зоны проводимости, а в кремнии — на расстоянии 0,05 эВ. Поэтому с такого уровня электроны легко переходят в зону проводимости. Примесные атомы ионизируются и существуют внутри кристаллов в виде неподвижных положительных зарядов. Вероятность того, что электрон имеет энергию, соответствующую зоне проводимости, намного больше в полупроводнике n-типа, чем в чистом полупроводнике.

Следовательно, уровень Ферми находится между серединой запрещенной зоны и дном зоны проводимости (рис. 2.19). Поскольку концентрация атомов в германии равна $4,4 \times 10^{28} \text{ м}^{-3}$ и на 10^5 атомов германия приходится один атом примеси, то число свободных электронов в нем будет $4,4 \times 10^{22} \text{ м}^{-3}$.

Концентрация свободных электронов в германии определяется следующим образом:

$$n_0 = G T^3 / 2 \cdot \exp[-(E_c - E_F) / kT] = N_c \exp[-(E_c - E_F) / kT]. \quad (2.2.1)$$

Здесь $G = 4,83 \times 10^{21} \text{ м}^{-3} \cdot \text{К}^{-3/2}$ — константа, связывающая число электронов в единице объема в зоне проводимости с температурой и энергией уровней, N_c — эффективная плотность состояний в зоне проводимости, определяемая выражением

$$N_c = 2(2\pi m^* kT / h^2)^{3/2} = 2 \{ [(2\pi m^* k)^{3/2} T^{3/2}] / h^3 \}.$$

Согласно условиям задачи $n_n = 4,4 \times 10^{22} \text{ м}^{-3}$.

Поскольку разница между донорным уровнем и дном зоны проводимости равна 0,01 эВ, а ширина запрещенной зоны $E_g = 0,72$ эВ, можно предположить, что все атомы примеси ионизованы, и поэтому можно пренебречь тепловой генерацией носителей. Тогда можно вычислить величину N_c при заданной температуре $T = 300 \text{ К}$. Подставляя в выражение для N_c значение $T^{3/2} = 300^{3/2} = 5196$, находим, что $N_c = 4,83 \cdot 10^{21} \cdot 5196 = 25 \times 10^{24} \text{ м}^{-3}$.

Перепишем выражение (2.2.1) в виде

$$\exp[(E_c - E_F) / kT] = N_c / n_n$$

или

$$(E_c - E_F) / kT = \ln(N_c / n_n).$$

Отсюда получаем

$$E_c - E_F = E_g - E_F = kT \cdot \ln(N_c / n_n) \text{ Дж.}$$

Таким образом,

$$E_F = 0,72 - 300 \cdot 1,38 \cdot 10^{-23} \cdot (1 / 1,610 - 19) \cdot \ln(25 \cdot 10^{21} / 4,4 \cdot 10^{22}) = 0,165 \text{ эВ.}$$

Это означает, что уровень Ферми находится примерно на 0,17 эВ ниже дна зоны проводимости.

2.3. Концентрация электронов и дырок в примесных полупроводниках

Выведите точное выражение для концентрации дырок p_n в полупроводнике п-типа через концентрацию доноров N_d и собственных носителей тока n_i . Выведите также выражение для концентрации электронов n_p в материале р-типа через концентрации акцепторов N_a и собственных носителей тока n_i . Предположите, что все примеси ионизованы.

Найдите отношение N_d / n_i , если $n_n = 1,005 \cdot N_d$. Вычислите p_n , если концентрация собственных носителей n_i для германия при температуре 300 К равна $4 \times 10^{19} \text{ м}^{-3}$. Найдите приближенные выражения для p_n и n_p , если $N_d \gg n_i$ в материале п-типа и $N_a \gg n_i$ в материале р-типа.

Решение

Для полупроводника п-типа справедливо следующее равенство:

$$n_n = p_n + N_d \quad (2.3.1)$$

Здесь n_n — концентрация электронов в материале п-типа, м^{-3} ; p_n — концентрация дырок в этом материале, м^{-3} ; N_d — концентрация доноров (атомов примеси), м^{-3} .

Используем классическое соотношение

$$n_n p_n = n_i^2 \quad (2.3.2)$$

где n_i — число пар "электрон—дырка", м^{-3} . Отсюда находим концентрацию электронов

$$n_n = n_i^2 / p_n \quad (2.3.3)$$

Подставляя (2.3.3) в (2.3.1) получаем

$$n_i^2 / p_n = p_n + N_d \quad (2.3.4)$$

Так приходим к уравнению

$$p_n^2 + N_d p_n - n_i^2 = 0 \quad (2.3.5)$$

Решение этого уравнения имеет вид

$$p_n = N_d / 2 \{ [1 + 4 \cdot (n_i / N_d)^2]^{1/2} - 1 \}. \quad (2.3.6)$$

Аналогично для полупроводника р-типа имеем

$$p_p = n_p + N_a \quad (2.3.7)$$

где p_p — концентрация дырок, м^{-3} ; n_p — концентрация электронов, м^{-3} ; N_a — концентрация акцепторной примеси, м^{-3} .

Снова используя соотношение

$$p_p n_p = n_i^2 \quad (2.3.8)$$

определяем p_p . Подставляя полученное выражение вместо p_p в (2.20), после соответствующего преобразования получаем уравнение

$$n_p^2 N_a n_p - n_i^2 = 0, \quad (2.3.9)$$

решение которого имеет вид

$$n_p = N_a / 2 \{ [1 + 4 \cdot (n_i / N_a)^2]^{1/2} - 1 \}. \quad (2.3.10)$$

Таким образом, полученные решения (2.3.6) и (2.3.10) представляют собой точные соотношения для концентрации электронов и дырок в примесных полупроводниках.

Преобразуя выражение (2.3.1) и подставляя в него, согласно условиям задачи, $n_n = 1,005 \cdot N_d$, получаем

$$p_n = n_n - N_d = 1,005 \cdot N_d - N_d = 0,005 \cdot N_d.$$

Подставляя значение $p_n / N_d = 0,005$ в соотношение (2.3.6), имеем

$$0,005 = 1 / 2 \{ -1 + [1 + 4(n_i / N_d)^2]^{1/2} \}$$

или

$$1,01 = [1 + 4(n_i / N_d)^2]^{1/2}.$$

Отсюда находим отношение $N_d / n_i = 14,1$.

$$N_d = 14,1 \cdot 4 \cdot 10^{19} = 5,64 \times 10^{20} \text{ м}^{-3}.$$

Следовательно, концентрация дырок равна

$$p_n = 0,005 \cdot N_d = 0,005 \cdot 5,64 \cdot 10^{20} = 2,82 \times 10^{18} \text{ м}^{-3}.$$

Если $N_d \gg n_i$ и к соотношению (2.3.6) применить биномиальную формулу

$$(1+x)^m = 1 + mx + m(m-1)x^2/2 + \dots,$$

то для первых членов получим

$$p_n \approx 1 / 2 \{ -N_d + N_d[1 + 2(n_i / N_d)^2] \} = n_i / N_d \quad (N_d \approx n_n).$$

При условии, что $N_d \gg n_i$, для полупроводника р-типа имеем

$$n_p \approx n_i / N_d \quad (p_p \approx N_d).$$

2.4. Подвижность электронов

Образец германия имеет концентрацию доноров $N_d = 2 \times 10^{20} \text{ м}^{-3}$. Эффективная масса электрона m^* равна $1,57 \cdot m_0$, а донор можно считать рассеивающим центром с $r = 5 \times 10^{-2} \text{ мкм}$. Чему равны средняя длина свободного пробега и среднее время между столкновениями при 300 К? Определить подвижность электронов.

Решение

Средняя длина свободного пробега $\lambda = 1 / (N \pi r^2)$, где r — радиус сферического рассеивающего центра, а N — концентрация носителей заряда. В данном случае

$$\lambda = 1 / (2 \cdot 10^{20} \cdot \pi \cdot 25 \cdot 10^{16}) = 0,64 \times 10^{-6} \text{ м}.$$

Среднее время между столкновениями $\tau = \lambda / v$, где v — средняя скорость электронов. Известно также, что

$$m^* v^2 / 2 = 3 / 2 kT.$$

Тогда

$$\tau = 2(m^* / 3kT)^{1/2} = 0,64 \cdot 10^{-6} \cdot (1,57 \cdot 9,106 \cdot 10^{-31} / 3 \cdot 1,38 \cdot 10^{-23} \cdot 300)^{1/2} = 0,69 \times 10^{-11} \text{ с}.$$

Подвижность

$$\mu = q\tau / m^* = 1,602 \cdot 10^{-19} \cdot 0,69 \cdot 10^{-11} / 1,57 \cdot 9,106 \cdot 10^{-31} = 0,77 \text{ м}^2 / (\text{В} \cdot \text{с}).$$

2.5. Диффузионная длина

Вычислить диффузионную длину электронов в германии р-типа и дырок в германии п-типа, если время жизни неосновных носителей заряда $\tau_n = \tau_p = 10^{-4} \text{ с}$, коэффициенты диффузии для германия р-типа $D_n = 47 \times 10^{-4} \text{ м}^2 / \text{с}$ и для германия п-типа $D_p = 47 \times 10^{-4} \text{ м}^2 / \text{с}$.

решение

Из выражения

$$D_n \tau_n = L_n^2$$

находим диффузионную длину электронов

$$L_n = \sqrt{D_n \tau_n} = \sqrt{99 \cdot 10^{-4} \cdot 10^{-4}} = 0,99 \text{ мм.}$$

Диффузионная длина дырок

$$L_p = \sqrt{D_p \tau_p} = \sqrt{47 \cdot 10^{-4} \cdot 10^{-4}} = 0,69 \text{ мм.}$$

2.6. Дрейфовые и диффузионные токи

Покажите, в чем заключается различие между дрейфовым и диффузионным токами в полупроводнике. Какую роль играет каждый из этих токов в работе полупроводникового прибора.

Докажите, что коэффициент диффузии D для любого типа носителей заряда определяется выражением

$$D = kT\mu/q,$$

где k — постоянная Больцмана, T — абсолютная температура, μ — подвижность носителей заряда, q — заряд электрона.

Решение

Если полупроводниковый образец находится во внешнем электрическом поле E , то через него течет дрейфовый ток плотностью

$$J = nq\mu E.$$

При инжекции носителей заряда с одной из сторон образца в этой области образуется высокая, а с противоположной стороны — низкая избыточная концентрация носителей. Поэтому носители диффундируют из области с высокой концентрацией.

Для плотности электронного и дырочного тока имеем

$$J_n = qD_n \left(\frac{\Delta n}{dx} \right),$$

$$J_p = -qD_p \left(\frac{\Delta p}{dx} \right).$$

Следовательно, диффузионные электронный и дырочный токи записываются в виде

$$I_n = qD_n S \left(\frac{\Delta n}{dx} \right),$$

$$I_p = -qD_p S \left(\frac{\Delta p}{dx} \right),$$

где D_n и D_p — коэффициенты диффузии электронов и дырок, $\text{м}^2/\text{с}$; S — площадь потоков.

В общем случае электроны и дырки могут перемещаться посредством дрейфа и диффузии. При этом плотность электронного тока имеет вид

$$J_n = nq\mu_n E + qD_n \left(\frac{\Delta n}{dx} \right), \quad (2.6.1)$$

а плотность дырочного тока

$$J_p = pq\mu_p E - qD_p \left(\frac{\Delta p}{dx} \right). \quad (2.6.2)$$

Уравнения (2.6.1) и (2.6.2) называют уравнениями потока, или диффузионно-дрейфовыми уравнениями. В обоих уравнениях члены, описывающие дрейфовый ток, положительны, т. к. противоположно заряженные частицы в электрическом поле, двигаясь в разные стороны, создают ток одного и того же направления.

Члены, описывающие диффузионный ток, имеют противоположные знаки, поскольку градиент концентрации заставляет оба вида частиц диффундировать в одном направлении. А в силу того, что заряды частиц имеют разные знаки, токи, создаваемые их диффузией, текут в противоположных направлениях.

Рассмотрим неравномерно легированный образец. В условиях теплового равновесия, т. е. в условиях отсутствия освещения, градиентов температуры и внешнего электрического поля, имеем

$$J_p = 0.$$

Используя это условие в уравнении (2.6.2), получаем

$$\Delta p q \mu_p E = q D_p \left(\frac{d\Delta p}{dx} \right), \quad (2.6.3)$$

а затем, подставляя $E = -dV/dx$ и разделяя переменные, получаем уравнение

$$\frac{d\Delta p}{\Delta p} = \left(-\frac{\mu_p}{D_p} \right) dV. \quad (2.6.4)$$

Интегрирование этого уравнения и соответствующее преобразование дают

$$\Delta p = A \exp \left(-\frac{\mu_p V}{D_p} \right), \quad (2.6.5)$$

где A — константа.

Аналогично находим

$$\Delta n = B \exp \left(-\frac{\mu_n V}{D_n} \right). \quad (2.6.6)$$

Сравнивая (2.6.5) и (2.6.6) с выражениями

$$\Delta p = n_i \exp \left(-\frac{qV}{kT} \right), \quad (2.6.7)$$

$$\Delta n = n_i \exp \left(-\frac{qV}{kT} \right), \quad (2.6.8)$$

получаем соотношение Эйнштейна

$$\frac{\mu_p}{D_p} = \frac{\mu_n}{D_n} = \frac{q}{kT},$$

или в общем виде

$$D = \frac{kT\mu}{q}.$$

Другое доказательство. Для газа в равновесном состоянии имеем $p \approx nkT$. Градиент давления

$$dp/dx = kT(dn/dx)$$

представляет собой ту силу, которая заставляет газ расширяться в доступный для него объем.

В полупроводнике, в котором на расстоянии x от инжектирующего контакта существует неравномерно распределенная избыточная концентрация электронов Δn , возникает внутреннее поле E . Сила, действующая на электрон, равна qE , а сила, действующая на Δn электронов, есть $\Delta nqE = \Delta nqu/\mu$.

В условиях теплового равновесия

$$\Delta nqE\mu_n = qD_n \left(\frac{d\Delta n}{dx} \right),$$

т. е. дрейфовая составляющая тока равна диффузионной. Таким образом, для силы, действующей на Δn электронов, или для градиента давления имеем

$$\Delta neE = \left(\frac{qD_n}{\mu_n} \right) \left(\frac{d\Delta n}{dx} \right) = \left(\frac{dp}{dx} \right) = kT \left(\frac{d\Delta n}{dx} \right).$$

Отсюда находим, что

$$\frac{qD_n}{\mu_n} = kT.$$

Аналогично находим соотношение для дырок $qD_p/\mu_p = kT$. Таким образом, в общем виде для коэффициента диффузии носителей заряда имеем соотношение:

$$D = \frac{kT\mu}{q}.$$

2.7. Потенциальный барьер на р—п-переходе

Объясните кратко, почему при нулевом смещении на границе р—п-перехода существует потенциальный барьер, и покажите, что его высота определяется выражением

$$\phi = \frac{kT}{e} \ln \frac{N_a N_d}{n_i^2},$$

где N_a и N_d — концентрации акцепторов и доноров в дырочной и электронной областях, соответственно; n_i — концентрация собственных пар "электрон—дырка". Подробно расскажите, как приложенное внешнее напряжение действует на потенциальный барьер.

Решение

В реальных р—п-переходах либо $N_a > N_d$, либо $N_a < N_d$, где N_a и N_d — концентрации акцепторов и доноров в дырочном и электронном материалах. В любом из этих случаев заряд по обе стороны границы перехода распределен по-разному и, следовательно, на переходе существует потенциальный барьер.

В наиболее распространенном случае $N_a > N_n$ в дырочной области концентрация дырок p_p высока, а концентрация электронов n_p очень низка. В электронной области высокая концентрация электронов n_n и очень низкая концентрация дырок p_n .

При нулевом смещении дырочный и электронный ток каждый состоит из двух компонент.

Дырочный ток имеет следующие две компоненты:

- дрейфовый ток, создаваемый дырками — неосновными носителями в электронной области — при их переходе из электронной в дырочную область под действием электрического поля;
- диффузионный ток, создаваемый дырками — основными носителями дырочной области — при диффузии их из этой области через потенциальный барьер.

Поэтому плотности электронного и дырочного тока определяются выражениями

$$J_p = pq\mu_p E - qD_p \frac{dp}{dx}, \quad (2.7.1)$$

$$J_n = nq\mu_n E + eD_n \frac{dn}{dx}. \quad (2.7.2)$$

Запишем условие нулевого смещения

$$J_p = J_n = 0.$$

Применим его к (2.7.1), откуда получим

$$pq\mu_p E = qD_p \frac{dp}{dx}$$

или, разделяя переменные,

$$\left(\frac{\mu_p}{D_p} \right) E dx = \frac{dp}{p}. \quad (2.7.3)$$

На рис. 2.20 приведено распределение электростатического потенциала на р-п-переходе в условиях равновесия.

Согласно соотношению Эйнштейна для диффузии,

$$\frac{\mu_p}{D_p} = \frac{\mu_n}{D_n} = \frac{q}{kT}.$$

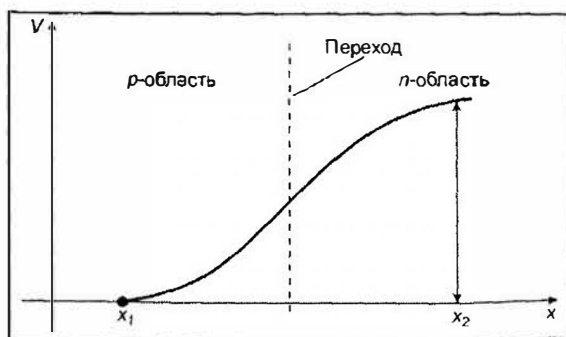


Рис. 2.20. Распределение потенциала на р-п-переходе

Подставляя это соотношение в уравнение (2.7.3), получаем

$$\frac{q}{kT} E dx = \frac{dp}{p}. \quad (2.7.4)$$

Интегрируя это уравнение в пределах от V_1 до V_2 и подставляя $E = -dV/dx$, имеем

$$-\frac{q}{kT} (V_2 - V_1) = \ln \frac{p_2}{p_1}. \quad (2.7.5)$$

Выражение (2.7.5), исходя из данных, приведенных на рис. 2.20, можно записать в следующем виде:

$$-\frac{q}{kT} \varphi = \ln \frac{p_n}{p_p}, \quad (2.7.6)$$

где $\varphi = V_2 - V_1$ — диффузионный потенциал, p_n — концентрация дырок в электронной области в точке x_2 , p_p — концентрация дырок в дырочной области в точке x_1 . Но

$$p_p \approx N_a$$

$$p_n \approx n_i^2 / N_d$$

поэтому выражение (2.7.6) можно записать как

$$-\frac{e}{kT} \varphi = \ln \frac{n_i^2}{N_a N_d} = -\ln \frac{N_a N_d}{n_i^2}.$$

Следовательно, высота потенциального барьера φ равна

$$\varphi = \frac{kT}{e} \ln \frac{N_a N_d}{n_i^2}.$$

Для электронов получаем уравнение, аналогичное (2.7.3):

$$\frac{\mu_n}{D_n} E dx = -\frac{dn}{n}.$$

Подставляя в него соотношение Эйнштейна, имеем

$$\frac{q}{kT} E dx = -\frac{dn}{n}. \quad (2.7.7)$$

Интегрируя это уравнение в пределах от V_1 до V_2 , получаем

$$-\frac{q}{kT} (V_2 - V_1) = -\ln \frac{n_2}{n_1}. \quad (2.7.8)$$

В точке x_1 концентрация $n_1 = n_p$ и в точке x_2 $n_2 = n_n$.

Но

$$n_n \approx N_d$$

$$n_p \approx n_i^2 / N_a.$$

Следовательно, выражение (2.7.8) можно записать следующим образом:

$$\frac{e}{kT} \varphi = \ln \frac{N_d N_a}{n_i^2},$$

откуда получаем выражение для высоты потенциального барьера:

$$\phi = \frac{kT}{e} \ln \frac{N_a N_d}{n_i^2}.$$

При прямом смещении V электронная область заряжена отрицательно относительно дырочной. При этом потенциальный барьер уменьшается до величины

$$\phi - V$$

и обедненный слой сужается. Таким образом, через р—п-переход может протекать большой ток.

При обратном напряжении смещения V электронная область находится под положительным потенциалом относительно дырочной области. Высота потенциального барьера увеличивается и становится равной

$$\phi + V,$$

т. е. обедненный слой расширяется. Через р—п-переход может идти только очень небольшой ток. Следовательно, в последнем случае переход является выпрямляющим.

2.8. Уравнение полупроводникового диода

Объясните качественно работу р—п-перехода, используемого в выпрямителе. Покажите, что если рассматривать движение носителей заряда через потенциальный барьер на переходе, то вольтамперная характеристика идеального р—п-перехода имеет вид

$$I = I_0 [\exp(qV / kT) - 1],$$

где I_0 — обратный ток насыщения; V — приложенное напряжение; T — температура перехода, К; k — постоянная Больцмана.

Постройте вольтамперную характеристику, соответствующую написанному выше уравнению. На том же графике приведите типичную вольтамперную характеристику реального перехода и отметьте причины расхождения между этими кривыми.

Решение

В отсутствие напряжения смещения и тока (условие равновесия) в р—п-переходе концентрация акцепторов N_a в р-области обычно больше концентрации доноров N_d . В р-области имеется высокая концентрация дырок p_p и низкая концентрация электронов n_p , а в п-области — высокая концентрация электронов n_n и низкая концентрация дырок p_n . Электроны и дырки рекомбинируют в области металлургического перехода, приводя к возникновению обедненного слоя толщиной W_d , в котором нет свободных носителей заряда. Числа ионизованных атомов примеси по обе стороны перехода равны между собой. Пусть W_n и W_p — толщины обедненного слоя в п- и р-областях. Если р-область является сильнолегированной, то обедненный слой простирается в глубь п-области, т. е. $W_n > W_p$.

При прямом смещении V п-область находится под отрицательным потенциалом относительно р-области и потенциальный барьер снижается до величины $q(\phi - V)$. Таким образом, может диффундировать большее число дырок из р-области в п-область и большее число электронов из п-области в р-область. Напряжение смещения не оказывает влияния на движение неосновных носителей p_p и n_n (т. е. на дрейфовые токи). Толщина обедненного слоя уменьшается, и через переход идет большой ток.

При обратном напряжении смещения V (п-область находится под положительным потенциалом относительно р-области) потенциальный барьер возрастает до величины $q(\phi + V)$. Толщина обедненного слоя увеличивается, и через переход может идти небольшой ток. Таким образом, р—п-переход проявляет выпрямляющие свойства.

При нулевом напряжении смещения концентрации неосновных носителей определяются выражениями

$$p_e = p_p \cdot \exp(-q\phi / kT) \quad (2.8.1)$$

$$n_p = n_n \cdot \exp(-q\phi / kT). \quad (2.8.2)$$

Пусть p_e — общее число дырок в п-области при прямом смещении V , а p_n — число дырок в условиях равновесия. Для p_e можно записать следующее выражение:

$$p_e = p_p \cdot \exp[-q(\phi - V) / kT]. \quad (2.8.3)$$

Избыточная плотность дырок в п-области $\Delta p(0) = p_e - p_n$. Следовательно, используя выражения (2.8.1) и (2.8.3), получаем

$$\begin{aligned} \Delta p(0) &= p_e - p_n = p_p \cdot \exp[-q(\phi - V) / kT] - p_p \cdot \exp(-q\phi / kT) = \\ &= p_p \cdot \exp(-q\phi / kT) [\exp(q\phi / kT) - 1]. \end{aligned}$$

И тогда

$$\Delta p(0) = p_n \cdot \exp[(q\phi / kT) - 1]. \quad (2.8.4)$$

Наличие этой избыточной концентрации дырок в области перехода приводит к диффузии дырок в глубь п-области. Плотность тока через переход, обусловленного этими дырками, равна

$$J_p(0) = qD_p(d\Delta p / dx) \text{ при } x = 0.$$

В р-области соответственно возникает избыточная концентрация электронов $\Delta n(0)$:

$$\Delta n(0) = n_p \cdot [\exp(q\phi / kT) - 1]. \quad (2.8.5)$$

В процессе диффузии концентрация носителей изменяется по закону:

$$\Delta p(x) = \Delta p(0) \exp(-x / L_p). \quad (2.8.6)$$

Дифференцируя это выражение по x , находим

$$d\Delta p(x) / dx = -\Delta p(0) / L_p \cdot \exp(-x / L_p),$$

откуда при $x = 0$ получаем

$$d\Delta p(x) / dx = -\Delta p(0) / L_p.$$

Соответствующая плотность дырочного тока определяется выражением

$$J_p(0) = -qD_p(-\Delta p(0) / L_p) = qD_p[\Delta p(0) / L_p]. \quad (2.8.7)$$

Преобразуем выражение (2.8.7) к виду

$$\Delta p(0) = J_p(0)L_p / qD_p. \quad (2.8.8)$$

Подставив это выражение в (2.8.4), имеем

$$J_p(0) = qD_p p_n / L_p [\exp(qV / kT) - 1].$$

Аналогично получаем выражение для плотности электронного тока

$$J_n(0) = qD_n n_p / L_n [\exp(qV / kT) - 1].$$

Общая плотность тока

$$J = J_p(0) + J_n(0).$$

Подставляя в это уравнение соответствующие выражения для плотностей дырочного и электронного тока, получаем

$$J = q(D_p p_n / L_p + q D_n n_p / L_n) \cdot [\exp(qV / kT) - 1].$$

Полный ток равен $J \cdot S$, где S — площадь границы перехода, т. е.

$$I = qS(D_p p_n / L_p + q D_n n_p / L_n) \cdot [\exp(qV / kT) - 1]. \quad (2.8.9)$$

В случае прямого смещения потенциал V положителен, а в случае обратного смещения — отрицателен. При обычной комнатной температуре $q / kT = 40 \text{ В}^{-1}$. Поэтому при увеличении обратного напряжения смещения (V отрицательно) $\exp(-qV / kT)$ в выражении (2.48) стремится к нулю. Следовательно, для отрицательных значений V обратный ток насыщения I_0 определяется из (2.8.9) следующим образом:

$$I_0 = -qS \cdot (D_p p_n / L_p + q D_n n_p / L_n). \quad (2.8.10)$$

Подставляя (2.8.10) в (2.8.9), получаем искомое выражение для вольтамперной характеристики

$$I = I_0 \cdot [\exp(qV / kT) - 1].$$

На рис. 2.21 построены вольтамперные характеристики идеального и реального переходов. В реальных переходах из-за загрязнения поверхности обратное сопротивление может быть порядка 100 МОм. Как видно из рисунка, в результате этого вольтамперная характеристика отклоняется от идеальной. Если загрязнение удастся значительно снизить, то сопротивление возрастает вплоть до 1000 МОм.

При прямом смещении наличие сопротивления материала диода вне пределов области объемного заряда (обедненного слоя) приводит к тому, что в реальных диодах для получения того же значения тока необходимо приложить большее напряжение.

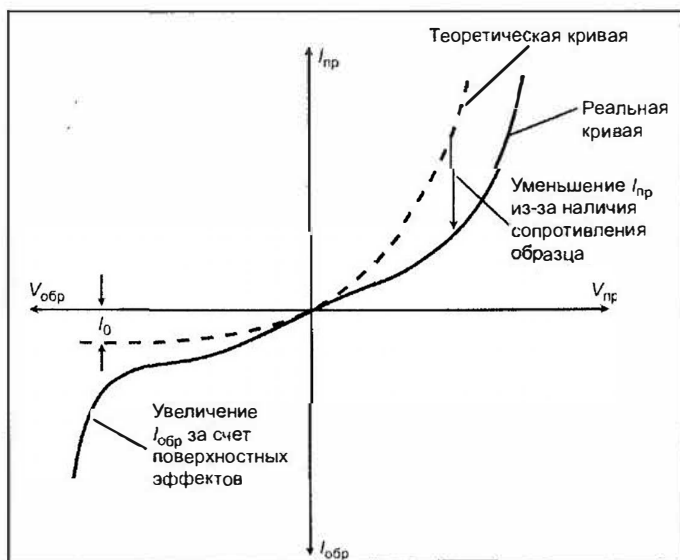


Рис. 2.21. Вольтамперная характеристика полупроводникового диода

2.9. Барьерная емкость

Барьерная емкость диода равна 200 пФ при обратном напряжении 2 В. Какое требуется обратное напряжение, чтобы уменьшить емкость до 50 пФ, если контактная разность потенциалов $\phi_k = 0,82$ В?

Решение

Барьерная емкость резкого р—п-перехода $[Ф/м^2]$ определяется по формуле

$$C_{\bar{\sigma}} = [eqN_aN_d / 2(N_a + N_d)]^{1/2} \cdot V^{-1/2},$$

где V — напряжение на р—п-переходе; N_a и N_d — концентрации примесей на каждой из сторон р—п-перехода. Следовательно, для данного диода

$$C_{\bar{\sigma}} = K / (V_{обр} + \phi_k)^{1/2},$$

где K — некоторая постоянная; $V_{обр}$ — обратное напряжение; ϕ_k — контактная разность потенциалов.

При $V_{обр} = 2$ В имеем $C_{\bar{\sigma}} = 200$ пФ.

Тогда

$$K = 200 \cdot 10^{-12} (2 + 0,82)^{1/2} = 3,35 \times 10^{-10} \text{ пФ} \cdot \text{В}^{1/2}.$$

Находим теперь обратное напряжение, при котором $C_{\bar{\sigma}} = 50$ пФ:

$$50 \cdot 10^{-12} = 3,35 \cdot 10^{-10} / (V_{обр} + 0,82)^{1/2},$$

откуда имеем $V_{обр} = 44,1$ В.

Контрольные вопросы

1. Что такое полупроводник?
2. Чем определяется электропроводность полупроводников?
3. Что такое процесс термогенерации электронно-дырочных пар?
4. Чем и с какой целью легируются полупроводники?
5. Как формируются разрешенные и запрещенные уровни в полупроводнике?
6. Изложите основные положения зонной теории полупроводников.
7. Что такое уровень Ферми в полупроводниковых структурах?
8. Как определяется коэффициент диффузии?
9. Как охарактеризовать процесс диффузии носителей заряда в полупроводниках?
10. Как охарактеризовать процесс дрейфа носителей заряда в полупроводниках?
11. Что такое встроенное электрическое поле в полупроводниках?
12. Чем характеризуется подвижность носителей в полупроводниках?
13. Какими параметрами характеризуется р—п-переход?
14. Что такое прямое и обратное включение р—п-перехода?
15. Опишите полный ток через р—п-переход.
16. Определите понятие диффузионной емкости.
17. Определите понятие барьерной емкости.
18. Что такое гетеропереход? Назовите основные типы гетеропереходов.

19. Что такое квантовый колодец?
20. Какие типы контактов вы знаете?
21. Что такое барьер Шоттки?

Рекомендуемая литература

1. Бонч-Бруевич В. Л., Калашников С. Г. Физика полупроводников. Учебное пособие для вузов. — М.: Наука, 1990.
2. Ефимов И. Е., Козырь И. Я., Горбунов Ю. Я. Микроэлектроника. Физические и технологические основы, надежность: Учебное пособие. — 2-е изд. — М.: Высшая школа, 1986.
3. Зи С. Физика полупроводниковых приборов: В 2-х книгах. — 2-е изд. — М.: Мир, 1984.
4. Линч П., Николайдес А. Задачи по физической электронике. Под. ред. проф. Г. В. Скроцкого. — М.: Мир, 1975.
5. Морозова И. Г. Физика электронных приборов. Учебник для вузов. — М.: Атомиздат, 1980.
6. Степаненко И. П. Основы микроэлектроники: Учебное пособие для вузов. — 2-е изд. — М.: Лаборатория базовых знаний, 2000.
7. Терсхов В. А. Задачник по электронным приборам. — 2-е изд. — М.: Энергоатомиздат, 1983.
8. Шалимова К. В. Физика полупроводников. Учебник для вузов. — М.: Энергоатомиздат, 1985.

3. Интегральные транзисторные структуры

3.1. Классификация транзисторных структур

Транзистор (от английского *transfer* — переносить и латинского *resistor* — сопротивляюсь) — электронный прибор на основе полупроводникового кристалла, имеющий три электрода и предназначенный для усиления, генерирования и преобразования электрических колебаний. Обычно в транзисторе существуют два взаимодействующих р—п-перехода.

Транзисторы по физическому принципу работы делятся на три основных класса:

- ☐ униполярные (полевые);
- ☐ биполярные;
- ☐ перспективные транзисторные структуры.

В *униполярных* транзисторах физические процессы протекания электрического тока в полупроводнике обусловлены носителями заряда одного знака — электронами или дырками. Основным физическим процессом перемещения носителей является дрейф в электрическом поле.

В *биполярных* транзисторах физические процессы обусловлены переносом носителей заряда обоих знаков. В основе работы биполярных транзисторов лежат процессы инжекции неосновных носителей, диффузии и дрейфа основных и неосновных носителей тока.

В настоящее время известно большое число различных типов транзисторов. На основе интегральных транзисторов формируются различные пассивные и активные элементы: диоды, резисторы, транзисторы. Поэтому уместно рассмотреть различные транзисторные структуры.

К *перспективным* транзисторным структурам отнесем транзисторные структуры, использующие различные физические явления в полупроводниках, например:

- ☐ транзисторы типа Terahertz;
- ☐ транзисторы с проницаемой базой;
- ☐ баллистические транзисторы;
- ☐ МОП-транзисторы с управляемой проводимостью канала;
- ☐ МОП-транзисторы с двойным затвором (вертикальные МОП-транзисторы).

В рамках данного курса микроэлектроники будут рассматриваться интегральные транзисторные структуры, применяемые в интегральных схемах различных типов.

Дискретные полупроводниковые приборы, включая транзисторы и тиристоры, в этом курсе рассматриваться не будут.

3.2. Интегральные униполярные (полевые) транзисторы

Полевым или униполярным транзистором называется транзистор, в котором управление происходит под действием электрического поля перпендикулярного току.

Проводящий слой, по которому протекает ток, называется *каналом*. Различают *p*- и *n*-канальные транзисторы. Каналы могут быть приповерхностными и объемными, горизонтальными и вертикальными.

В свою очередь приповерхностные каналы делятся на обогащенные или обедненные носителями, либо инверсионные слои. Их формирует внешнее электрическое поле. Обедненные каналы представляют собой участки однородного полупроводника, отделенные от поверхности обедненным слоем.

На рис. 3.1 приведены схемы каналов в униполярных транзисторах. Транзисторы с приповерхностным каналом имеют структуру *металл — диэлектрик — полупроводник* (МДП). Такие транзисторы принято называть *МДП-транзисторами*. Если диэлектриком является диоксид кремния SiO_2 , то используется название *МОП-транзисторы*. Транзисторы с объемным каналом получили название *полевых транзисторов*.

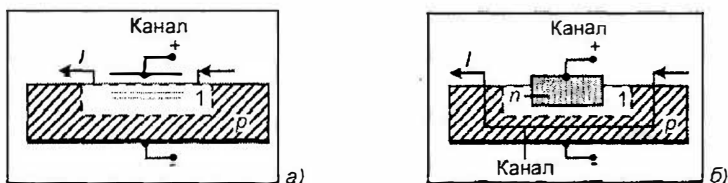


Рис. 3.1. Каналы в униполярных транзисторах: а — приповерхностный *n*-канал; б — объемный *p*-канал; 1 — обедненный слой

3.2.1. МДП-транзисторы с индуцированным каналом

Классическая структура интегрального МДП-транзистора приведена на рис. 3.2. МДП-транзистор имеет три электрода: *исток* (И), *сток* (С) и *затвор* (З). Исток и сток формируются методом диффузии или методом ионной имплантации. Управляющим электродом является затвор — металлический электрод, перекрывающий канал между истоком и стоком. Иногда исток напрямую замыкают на электрод подложки (*Ил*). Если на электроды подан нулевой потенциал, то вблизи n^- -областей истока и стока имеются области объемного заряда, возникающие за счет разности работ выхода электронов из полупроводника с различными типами электропроводности. Между истоком и стоком при $U_{\text{ин}} = 0$ существует большое сопротивление, эквивалентное сопротивлению двух встречно включенных диодов при нулевом смещении. Поэтому при подаче напряжения $U_{\text{ин}}$ во внешней цепи ток будет мал. Если на затвор подать отрицательное напряжение, то приповерхностный слой обогатится дырками, что не изменит тока во внешней цепи.

Если на затвор подать положительное напряжение $U_{\text{ин}} > 0$, то под действием электрического поля все основные носители (дырки) отождутся полем в глубь полупроводника (эффект поля).

Эффектом поля в полупроводниках называется изменение концентрации свободных носителей заряда в приповерхностном слое под действием внешнего электрического поля.

перпендикулярного каналу. Сначала образуется обедненный слой (объемный заряд акцепторов), куда устремляются неосновные носители — электроны. У самой поверхности, границы раздела "полупроводник—диэлектрик", электроны образуют *инверсионный слой* — проводящий канал. Такой тип канала называется *индуцированным*. Ток стока резко возрастает, и в дальнейшем зависит от напряжения $U_{\text{зи}}$. Толщина индуцированного канала обычно лежит в пределах 1—2 нм. Напряжение на затворе, при котором образуется канал и транзистор начинает работать, называется *пороговым* и обозначается U_0 .

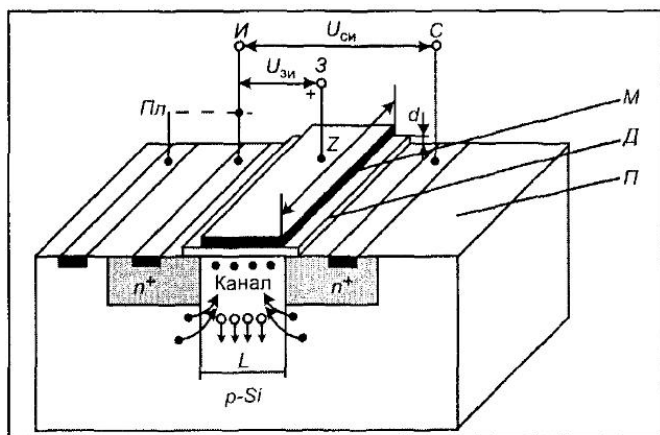


Рис. 3.2. Структура интегрального МДП-транзистора: И — исток; С — сток; З — затвор; Пл — подложка; М — металл; Д — диэлектрик; П — полупроводник. Точками обозначены электроны, а дырки — кружочками

На рис. 3.3, а приведена стокowo-затворная вольтамперная характеристика МДП-транзистора с индуцированным каналом.

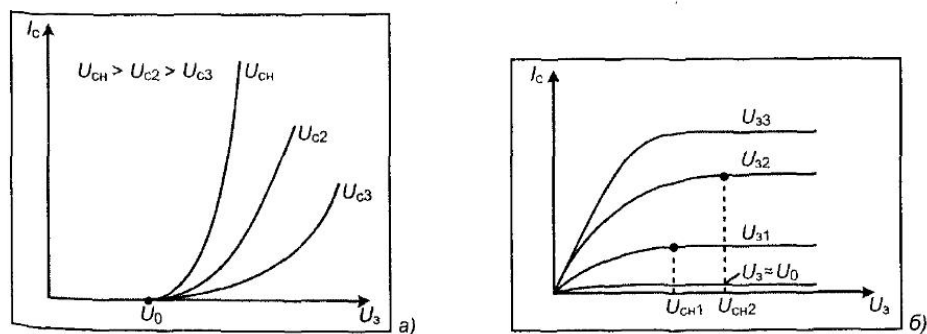


Рис. 3.3. Статические характеристики МДП-транзистора:

- а — стокowo-затворная вольтамперная характеристика МДП-транзистора с индуцированным каналом;
б — семейство передаточных стокowych характеристик МДП-транзистора с индуцированным каналом

Пороговое напряжение U_0 определяется удельной емкостью затвор — канал, зонной диаграммой металл — диэлектрик — полупроводник. Практические значения полного порогового напряжения лежат в пределах 0,5—1,5 В.

Стоково-затворные характеристики транзистора зависят от режима его работы. При напряжении $U_{сн} > 0$ ток протекает по каналу, создавая распределение потенциала по длине канала от истока к стоку. Разность потенциалов между затвором и поверхностью в направлении стока уменьшается, одновременно уменьшается напряженность поля в диэлектрике и удельный заряд электронов в канале. Сечение канала в точке $x = L$ у стока сужается.

При некотором напряжении на стоке разность потенциалов между затвором и поверхностью в точке $x = L$ делается равным пороговому напряжению. Это критическое напряжение на стоке называется *напряжением насыщения* $U_{сн}$. Именно в этот момент образуется *горловина канала* — слой объемного заряда, соприкасающийся с поверхностью. Ранее это слой был отделен от поверхности границы раздела "полупроводник—диэлектрик" поверхностным каналом. При образовании горловины канала ток в рабочей цепи не зависит от напряжения на стоке и происходит процесс насыщения.

На рис. 3.3, б приведено семейство передаточных стоковых характеристик МДП-транзистора с индуцированным каналом, где в качестве параметра используется напряжение на затворе.

В области пологих участков ВАХ вводятся следующие малосигнальные параметры транзистора:

- крутизна — $S = \left. \frac{dI_c}{dU_{зп}} \right|_{U_{сн}} = \text{const}$;
- сопротивление — $r_{\bullet} = \left. \frac{dU_{сн}}{dI_c} \right|_{U_{зп}} = \text{const}$;
- коэффициент усиления — $R = \left. \frac{dU_{сн}}{dU_{зп}} \right|_{I_c} = \text{const}$.

Все три параметра связаны между собой соотношением

$$R = S \cdot r_{\bullet}.$$

3.2.2. МДП-транзисторы со встроенным каналом

В МДП-транзисторе технологическим путем можно создать канал, соединяющий исток со стоком. Такой транзистор получил название — *транзистор со встроенным каналом*.

Структура транзистора со встроенным каналом приведена на рис. 3.4, а.

На стоково-затворной характеристике (рис. 3.4, б) видно, что при нулевом напряжении на затворе по его каналу течет ток, и транзистор способен усиливать сигнал. При подаче на затвор отрицательного напряжения ток в канале уменьшается вследствие действия отрицательного поля затвора и при некотором напряжении $U_{отс}$. Это объясняется тем, что при отрицательном напряжении на затворе канал обедняется носителями, и, следовательно, ток стока уменьшается. При увеличении напряжения канал обогащается неосновными носителями, и ток увеличивается. Передаточная характеристика МДП-транзистора также имеет участок насыщения (рис. 3.4, в).

На рис. 3.5 обобщены рассматриваемые типы МДП-транзисторных структур и приведены их условные обозначения.

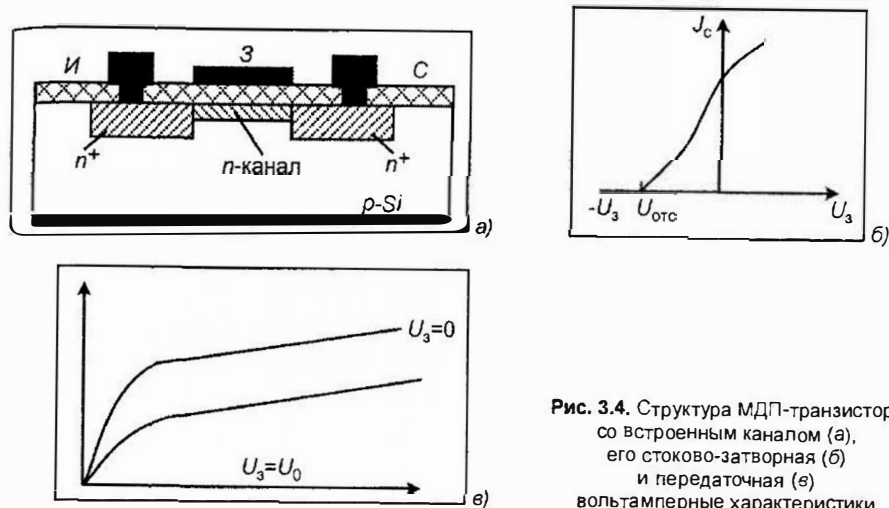


Рис. 3.4. Структура МДП-транзистора со встроенным каналом (а), его стоково-затворная (б) и передаточная (в) вольтамперные характеристики

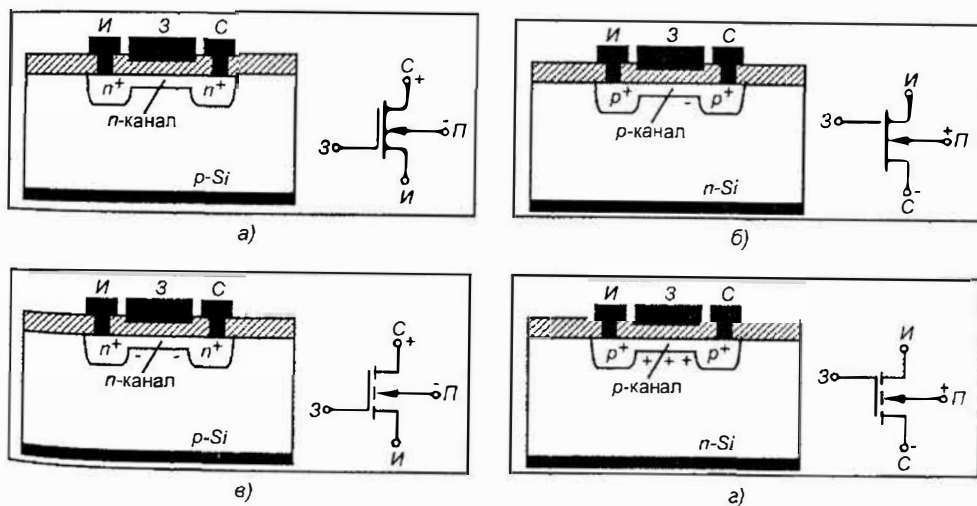


Рис. 3.5. Структуры и условные обозначения МДП-транзисторов: со встроенным каналом п-типа (а); р-типа (б); с индуцированным каналом п-типа (в); р-типа (д)

3.2.3. Комплементарные структуры

Сочетание транзисторных структур с каналами п- и р-типов представляют *комплементарную структуру*.

В таких структурах применяются транзисторы с индуцированными каналами. Отличительной особенностью комплементарных структур является противоположная полярность питающих и управляющих напряжений (рис. 3.6).

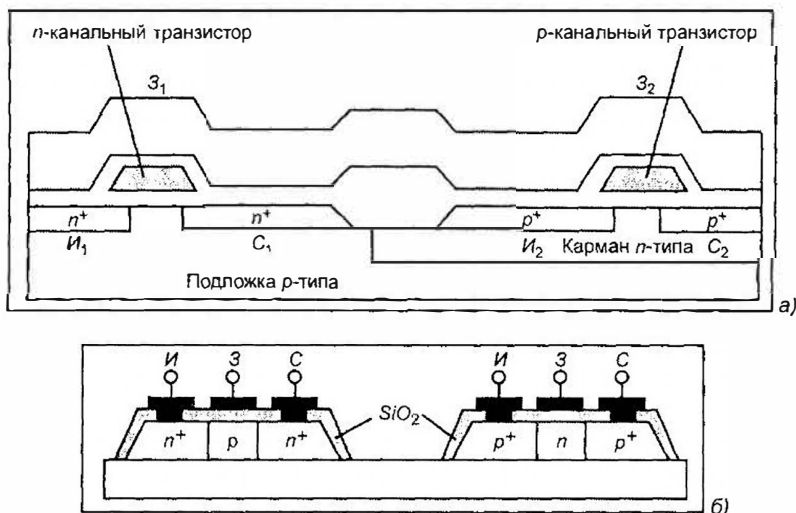


Рис. 3.6. Комплементарные транзисторные структуры, изготовленные в кристалле полупроводника (а), на диэлектрике (б)

Такая комбинация транзисторов позволяет сочетать высокое быстродействие и предельно малое потребление энергии от источника питания. Один из транзисторов в зависимости от типа подложки создают в изолирующем кармане. Создаются также охранные области, позволяющие устранить утечки тока и паразитные связи между МДП-транзисторами. По надежности и популярности предпочтение отдается КМОП-транзисторам, которые реализованы на диэлектрической подложке, например, на сапфире. Это позволяет получать транзисторы без токов утечки, с отсутствием паразитных емкостных связей между областями транзистора и подложки. Такая технология позволяет повысить быстродействие транзисторов, создать на их основе радиационно стойкие интегральные схемы.

3.2.4. Транзистор с управляющим р—п-переходом

Транзистор с управляющим р—п-переходом существенно отличается от транзистора с изолированным затвором и составляет отдельный тип транзисторных структур.

Под металлическим электродом затвора, расположенным непосредственно на полупроводнике, сформирован р-слой. Между затвором и любым из двух других электродов существует р—п-переход. Между истоком и стоком может протекать ток по каналу толщиной d , величина которого зависит от напряжения, приложенного к затвору.

Между истоком и стоком прикладывается напряжение $U_{кс}$, смещающее р—п-переход в запиорном направлении (минус на затворе) (рис. 3.7).

В таком случае под затвором возникает обедненный слой, имеющий высокое напряжение. Чем выше напряжение на затворе $U_{кс}$, тем больше толщина обедненного слоя d , в котором ток не течет. Таким образом, увеличение напряжения на затворе соответствует сужению канала, что вызывает уменьшение тока в канале. При дальнейшем увеличении напряжения $U_{кс}$ возможно полное перекрытие токопроводящего канала. Это напряжение называется *напряжением отсечки*.

Аналогично работает и *полевой транзистор с затвором в виде барьера Шоттки* (ПТШ). Эти приборы стали активно разрабатываться с момента выявления полупроводниковых

свойств арсенида галлия. Его преимущество по сравнению с кремнием заключается, прежде всего, в более высокой подвижности электронов в слабых электрических полях и скоростью насыщения в сильных полях. Для арсенида галлия характерна большая ширина запрещенной зоны и, следовательно, более высокое сопротивление нелегированного материала. Это обстоятельство позволяет создавать полуизолирующие подложки для интегральных схем.

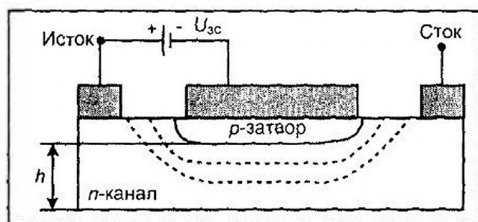


Рис. 3.7. Схема транзистора с управляющим p-n-переходом

Сравнительные характеристики электрофизических параметров кремния и арсенида галлия приведены в табл. 3.1. Хорошо видно, что по ряду параметров арсенид галлия превосходит кремний. По таким параметрам, как плотность поверхностных состояний и подвижность дырок, он проигрывает. Они имеют высокое и низкое значения, соответственно, что затрудняет разработку как высококачественных МДП-транзисторов, так и биполярных транзисторов. Наиболее оптимальным активным элементом на основе арсенида галлия является полевой транзистор с затвором в виде барьера Шоттки (ПТШ) или, как его еще называют, с управляющим переходом "металл—полупроводник" — МЕП-транзистор. МЕП-транзистор формируется на подложке из нелегированного арсенида галлия, имеющую слабо выраженную проводимость p-типа.

Таблица 3.1

Электрофизические параметры	Si	GaAs
Подвижность, $\text{см}^2/(\text{В} \times \text{с})$		
электронов	$(0,8—1,3) \times 10^3$	$(4—5) \times 10^3$
дырок	350	250
(при концентрации доноров 10^{17} см^{-3})		
Скорость насыщения в сильном электрическом поле, см/с	$0,8 \times 10^7$	2×10^7
Максимальное удельное сопротивление нелегированного материала, $\text{Ом} \times \text{см}$	10^3	$10^7—10^9$
Время жизни неосновных носителей, с	10^{-3}	10^{-8}
Плотность поверхностных состояний в МДП-структуре, см^{-2}	10^{10}	$10^{12}—10^{13}$
Экономика производства	+	~

Методом ионного легирования кремнием, серой или селеном формируют сильнолегированные области истока (И) и стока (С) n'-типа, а затем напыляют сплав Ti-W. Остальная поверхность покрывается диэлектриком, например SiO_2 . Металлический электрод затвора образует с каналом барьер Шоттки, типичная равновесная высота которого равна 0,8 В. Проводящий канал расположен между истоком и стоком и ограничен сверху обедненной

областью толщиной $L_{\text{об}}$. При изменении положительного напряжения на затворе $U_{\text{зи}}$ толщина проводящего канала изменяется следующим образом

$$d_k(U_{\text{зи}}) = d_0 - L_{\text{ро}}(U_{\text{зи}}).$$

Одновременно меняется проводимость и ток стока I_c .

Пороговое напряжение соответствует моменту, когда граница обедненного слоя достигает подложки и канал перекрывается (рис. 3.8, б, кривая 1, 2). Если пороговое напряжение отрицательно, то при $U_{\text{зи}} = 0$ канал является проводящим, и транзистор называют *нормально открытым* (рис. 3.8, б, кривая 1). Эта ситуация аналогична поведению МДП-транзистора со встроенным каналом.

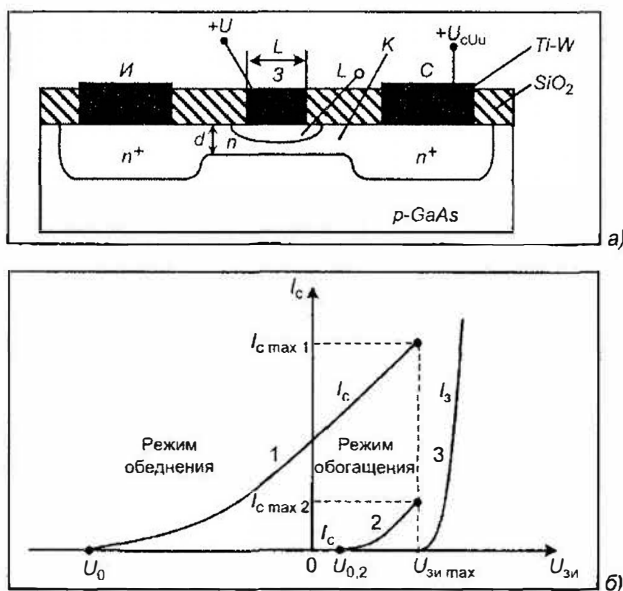


Рис. 3.8. Структура арсенид-галлиевого ПТШ-транзистора (а) и его ВАХ (б): $L_{\text{об}}$ — обедненная область; К — канал

Если пороговое напряжение $U_0 > 0$, то канал перекрыт обедненным слоем и транзистор называется *нормально закрытым*, что аналогично МДП-транзистору с индуцированным каналом. При больших положительных напряжениях на затворе в его цепи может возникнуть паразитный ток I_c (рис. 3.8, б, кривая 3). Это обусловлено открытием перехода "металл—полупроводник". Поэтому ток стока ограничивается значением $I_{c \text{ max}}$.

3.2.5. Полевой транзистор на гетероструктурах

Гетероструктуры представляют собой комбинации различных гетеропереходов. Как уже говорилось, гетеропереход представляет собой полупроводниковый переход между двумя разнородными по химическому составу или фазовому состоянию полупроводниками. Гетеропереходы формируются в объеме одного полупроводника.

Анизотипные переходы предусматривают контакт полупроводников с электронной и донорными типами проводимости. *Изотипные* переходы возникают при контакте полупроводников с одним типом проводимости.

На границе гетероперехода происходит скачкообразное изменение таких свойств, как:

- ширина запрещенной зоны;
- подвижность носителей заряда;
- эффективная масса носителей;
- энергия сродства к электрону и др.

Энергетическая диаграмма гетеропереходов характеризуется скачками энергии в зонах проводимости и валентной зоне. Высота потенциальных барьеров в них различна. В частности, в таких структурах возможно получение односторонней инжекции носителей заряда.

На рис. 3.9 приведена зонная диаграмма гетероперехода между арсенидом галлия и арсенидом галлия алюминия $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$. Величина x характеризует содержание алюминия, и с ростом x увеличивается ширина запрещенной зоны данного твердого раствора. Для типичного значения $x = 0,3$ ширина запрещенной зоны твердого раствора $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ равна 1,8 эВ.

У границы раздела двух полупроводников в зоне проводимости образуется квантовый колодец или зона двумерного электронного газа (ДЭГ).

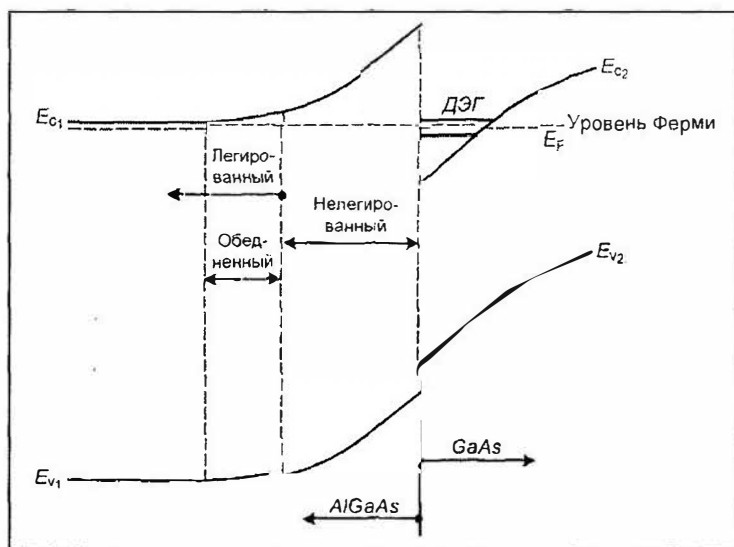


Рис. 3.9. Зонная диаграмма $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$ гетероперехода со слоем двумерного электронного газа

Двумерный электронный газ или 2D-газ представляет собой систему электронов, энергетические состояния которых соответствуют свободному движению только вдоль определенной плоскости.

Важным свойством двумерного электронного газа является то, что возможно регулирование в широких пределах плотности электронов под воздействием поперечного электрического поля. Электроны в ДЭГ имеют повышенную эффективную концентрацию и подвижность.

На основе таких гетероструктур изготавливаются *гетеротранзисторы* (рис. 3.10). Его конструкция многослойна. Между металлическим затвором и легированным слоем на основе $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ формируется управляющий переход "металл—полупроводник". Обедненная область этого перехода располагается в слоях арсенида галлия алюминия. Различают нормально открытый (рис. 3.10, а) и нормально закрытый транзистор (рис. 3.10, б). При $U_{из} < 0$ в слое нелегированного арсенида галлия на границе с гетеропереходом в области ДЭГ формируется канал нормально открытого транзистора. На рис. 3.10, а эта область ограничена штриховой линией.

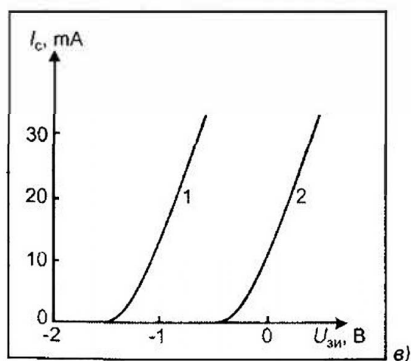
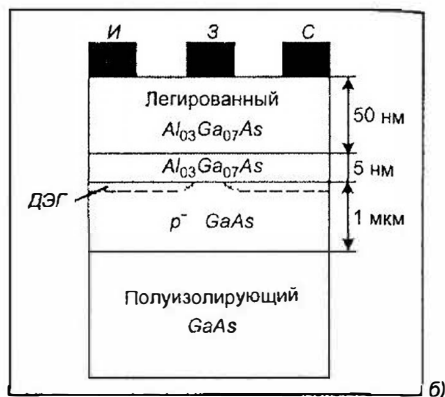
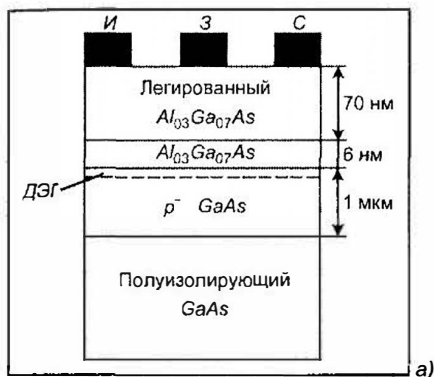


Рис. 3.10. Транзистор на гетероструктуре:
а — нормально открытый транзистор;
б — нормально закрытый транзистор;
в — ВАХ нормально открытого (1) и нормально закрытого (2) транзисторов

Под действием управляющего напряжения $U_{из}$ изменяется толщина обедненной области перехода "металл—полупроводник", концентрация электронов в области ДЭГ и ток стока. Электроны в ДЭГ поступают из истока. При отрицательном напряжении "затвор—исток", величина которого равна пороговому значению, обедненная область расширяется настолько, что перекрывает поток электронов, и ток стока становится равным нулю. В параллельно закрытому транзисторе при $U_{из} = 0$ проводящий канал отсутствует вследствие того, что область ДЭГ двумерного электронного газа перекрывается объединенной областью управляющего перехода. При подаче напряжения $U_{из} > 0$, величина которого равна пороговому значению, обедненная область управляющего перехода сужается настолько, что ее нижняя граница попадает в область ДЭГ.

На рис. 3.10, в приведены стокково-затворные характеристики нормального открытого (1) и нормального закрытого (2) транзисторов. Большое значение крутизны для нормально

закрытого транзистора обусловлено меньшей толщиной легированного донорами GaAlAs. Этот тип транзисторов перспективен для использования в СВЧ-микросхемах.

3.2.6. V-МДП-транзисторы

рассмотренные выше структуры имеют планарную конструкцию, ток носителей в которых переносится в горизонтальном направлении.

Транзисторы типа V-МДП относятся к типу транзисторов с вертикальным токопереносом — от расположенного в подложке истока к верхнему стоку.

Этот тип транзисторов изготавливается путем селективного травления в исходной структуре кремния V-образных канавок (рис. 3.11).

V-МДП-структура является весьма компактной. Истоковая область играет роль шины земли и не требует дополнительной площади кристалла для заземления.

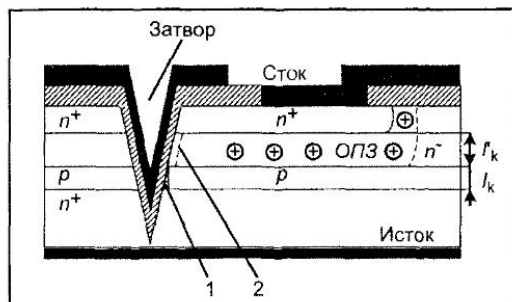


Рис. 3.11. Структура V-МДП-транзистора:
1 — область канала;
2 — область дрейфа электронов

Проводящий канал образуется в p-слое и его длина определяется микронными размерами, а ширина всем периметром V-образного углубления. Большая ширина канала позволяет получить транзисторы с большим током и большим усилением. Область объемного пространственного заряда (ОПЗ) позволяет увеличить пробивное напряжение транзистора и снизить значение паразитной емкости "затвор—сток". Трехмерность V-МДП-транзисторов является большим плюсом при создании объемных интегральных схем с высокой плотностью упаковки на кристалле.

3.3. Интегральный биполярный транзистор

3.3.1. Эпитаксиально-планарный транзистор

Биполярный транзистор — электронный прибор с тремя чередующимися полупроводниковыми областями электронного (n) или дырочного (p) типа проводимости, в котором протекание тока обусловлено носителями заряда обоих знаков.

В основе работы биполярного транзистора лежат физические явления диффузии вследствие градиента концентрации и дрейфа носителей вследствие градиента электрического потенциала. Полный ток носителей состоит из диффузионной и дрейфовой составляющих. Различают биполярные транзисторы p—n—p- и n—p—n-типов.

Классическая конструкция биполярного транзистора, используемая в интегральных схемах, выполняется по эпитаксиально-планарной технологии. Это означает, что транзистор

выполняется в эпитаксиальном слое толщиной $W_{эл}$, а выводы от эмиттерной, базовой и коллекторных областей расположены в одной плоскости на поверхности подложки. Такая технология позволяет производить транзисторные структуры с высоким процентом выхода годных, а планарность выводов позволяет создать автоматизированную систему коммутации транзистора с другими элементами микросхемы — пленочными металлическими проводниками.

Одним из важных требований, предъявляемым к технологии изготовления транзисторных структур, является минимизация занимаемой площади, что позволяет повысить плотность упаковки элементов интегральной схемы и способствует повышению степени интеграции на пластине.

На рис. 3.12 приведена классическая структура эпитаксиально-планарного транзистора с изоляцией р—п-переходом (а) и его топология (б).

Транзистор выполнен на высокоомной подложке р-типа. Методами диффузии примесей через маску получены скрытый слой (СС), п-, р- и n^+ -области. Глубина залегания эмиттерного перехода составляет 1,5—2,0 мкм. В данной конструкции транзистора используется изоляция р—п-переходами: сбоку изолирующими областями, а снизу — скрытым слоем. Недостатком изоляции р—п-переходами является наличие барьерной емкости, которая снижает граничную частоту и увеличивает задержку переключения сигнала.

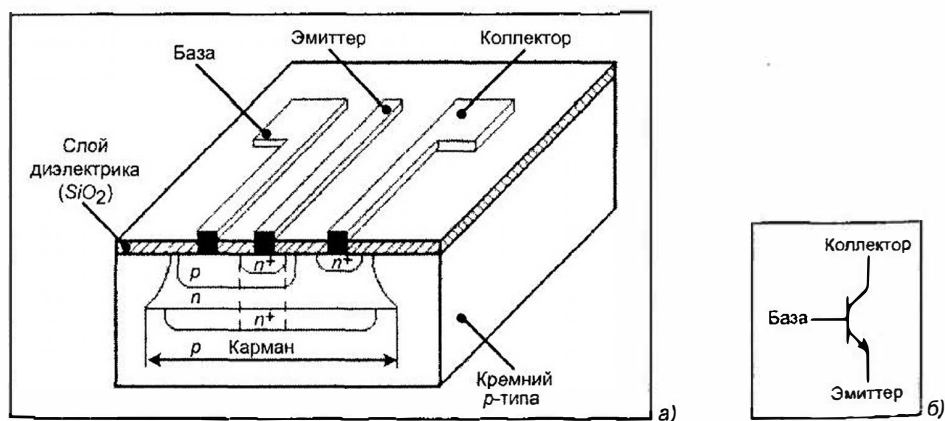


Рис. 3.12. Структура эпитаксиально-планарного n^+ —р—п-транзистора (а) и его условное обозначение (б)

Под эмиттерной областью расположена активная область транзистора, представляющая собой n^+ —р—п-структуру. Физические процессы диффузии и дрейфа носителей в этой области и определяют эффективность работы транзистора.

Конструктивно биполярный транзистор представляет собой совокупность двух взаимодействующих р—п-переходов, включенных навстречу друг другу.

Эпитаксиально-планарный транзистор р—п—р-типа имеет те же области, но с противоположной транзистору п—р—п-типа проводимостью. На условном обозначении стрелка эмиттера направлена к базовому электроду.

3.3.2. Физические основы работы биполярного транзистора

Рассмотрим физические процессы в структуре типа p^+-n-p . Аналогичные процессы происходят и в структуре типа n^+-p-n . В активном режиме работы транзистора эмиттерный переход включен в прямом направлении, а коллекторный — в обратном (рис. 3.13). Эмиттерная область легирована достаточно сильно, так, чтобы выполнялось условие $p_3 \gg n_6$, где p_3 — концентрация дырок в эмиттере, n_6 — концентрация электронов в базе. В этом случае основные носители эмиттерной области — дырки, вследствие градиента концентрации диффундируют в базу, где они являются неосновными носителями. Этот процесс перехода дырок в базовую область называется *инжекцией*.

Количественно процесс инжекции характеризуется *коэффициентом инжекции*:

$$\gamma = \frac{I_{p3}}{I_{p3} + I_{n3}},$$

где I_{p3} и I_{n3} — ток дырок и электронов через эмиттерный $p-n$ -переход, соответственно.

Эффективность эмиттера тем выше, чем больше γ , и при $\gamma = 1$ ток электронов $I_{n3} = 0$. Другими словами, максимальная эффективность эмиттера достигается при отсутствии тока электронов из базы в эмиттер. Обычно коэффициент инжекции составляет $\gamma \approx 0,9995$.

Оказавшись в области базы, дырки перемещаются под действием градиента концентрации к коллекторному переходу. На своем пути они встречают электроны, являющиеся основными носителями в области базы, и там рекомбинируют. Толщину базовой области W подбирают так, чтобы потери на процессе рекомбинации были минимальны. Обычно $W \ll L_6$, где L_6 — диффузионная длина неосновных носителей в базе, которая может меняться в пределах 0,3—1,5 мм.

Процесс рекомбинации дырок определяется коэффициентом

$$\eta = \frac{I_{pk}}{I_{p3}},$$

где I_{pk} — ток дырок в коллекторный переход, I_{p3} — ток дырок через эмиттерный переход.

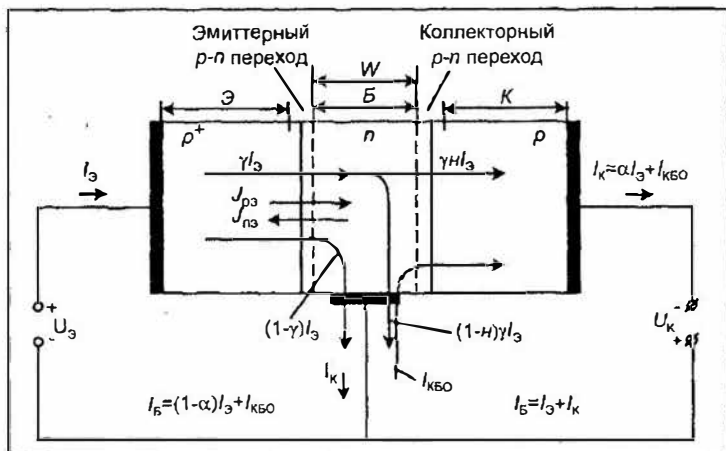


Рис. 3.13. Схема токопередачи в биполярном p^+-n-p -транзисторе

В типичных биполярных транзисторах значение $n \approx 0,95—0,99$, что означает, что лишь незначительная часть дырок рекомбинирует в области базы. Достигнув области коллектора, дырки дрейфуют под действием градиента внешнего отрицательного электрического поля и втягиваются в коллектор. Этот процесс называется *экстракцией* (извлечением) зарядов. Сквозь коллекторный переход протекает ток, который равен

$$I_{pk} = \gamma n I_s.$$

Полный ток в цепи коллектора равен

$$I_k = \alpha I_s + I_{kbo},$$

где α — коэффициент передачи тока эмиттера к коллектору, $\alpha = \gamma n$; I_{kbo} — обратный ток коллектора, сильно зависящий от температуры.

Ток в цепи базы обусловлен изменением заряда базы и определяется концентрацией основных носителей базы

$$I_b = I_s - I_k.$$

В базе транзистора происходят следующие процессы:

- ☐ диффузия электронов из базы в эмиттер навстречу потоку дырок;
- ☐ рекомбинация с дырками дрейф электронов и дырок через обратно включенный коллекторный переход.

Таким образом:

$$I_b = (1 - \gamma) I_s + (1 - n) \gamma I_s - I_{kbo} = (1 - \alpha) I_s - I_{kbo}.$$

Если в цепь эмиттера подать входные сигналы, а в цепь коллектора включить нагрузку, то транзистор будет работать как усилитель сигналов. Такой режим работы транзистора называется *активным*.

Помимо активного усилительного режима работы транзистора существует и другие. Так если оба перехода включить в прямом направлении, то будет происходить встречная инжекция неосновных носителей в базу. Сопротивление транзистора минимизируется вследствие насыщения базы носителями. Такой режим работы транзистора называется *режимом насыщения*.

Если оба перехода включить в обратном направлении, то пройдет процесс экстракции неосновных носителей в базу, и сопротивление базовой области станет большим. Это *режим отсечки тока*, и он соответствует закрытому состоянию транзистора.

Таким образом, транзистор может работать в режимах насыщения, усиления и отсечки.

3.3.3. Малосигнальные параметры

Под *малым сигналом* будет подразумеваться сигнал, который при его увеличении на 50% изменяет заданный параметр в пределах заданной точности. При этом используются небольшие участки ВАХ активной области транзистора. Большой сигнал (импульсный сигнал) использует значительные участки ВАХ, включая ее нелинейные части.

Различают различные способы включения транзистора как четырехполюсника: с общей базой (ОБ), с общим эмиттером (ОЭ) и общим коллектором (ОК) (рис. 3.14).

Как и для любого четырехполюсника, для биполярного транзистора существуют два уравнения, которые связывают между собой четыре физические величины посредством четырех коэффициентов, играющих роль независимых параметров.

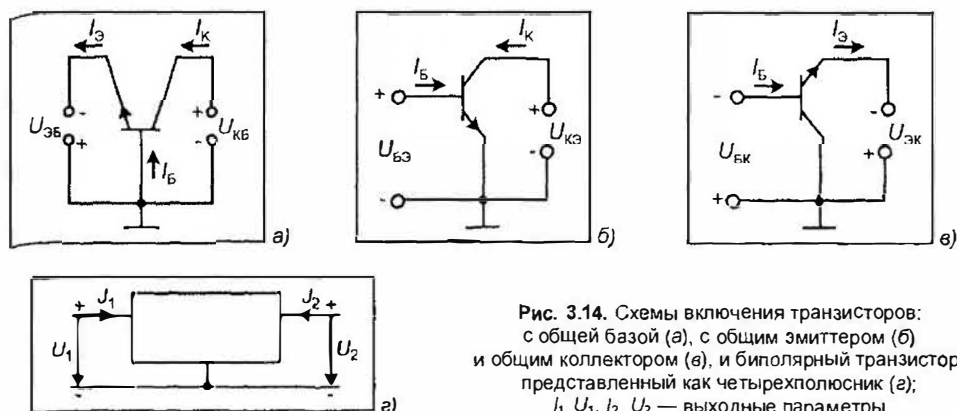


Рис. 3.14. Схемы включения транзисторов: с общей базой (а), с общим эмиттером (б) и общим коллектором (в), и биполярный транзистор, представленный как четырехполюсник (г); I_1, U_1, I_2, U_2 — выходные параметры

Введем систему гибридных параметров транзистора h_{ij} , физические размерности которых неодинаковы.

$$\begin{aligned} u_1 &= h_{11}i_1 + h_{12}u_2 \\ i_2 &= h_{21}i_1 + h_{22}u_2 \end{aligned} \Leftrightarrow \begin{pmatrix} u_1 \\ i_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} i_1 \\ u_2 \end{pmatrix},$$

где:

$h_{11} = \left. \frac{u_1}{i_1} \right|_{u_2=0}$ — выходное сопротивление в режиме короткого замыкания на выходе по переменному току;

$h_{12} = \left. \frac{u_1}{u_2} \right|_{i_1=0}$ — коэффициент обратной связи в режиме холостого хода на входе по переменному току, показывающий, какая доля выходного напряжения поступит на вход транзистора за счет обратной связи;

$h_{21} = \left. \frac{i_2}{i_1} \right|_{u_2=0}$ — коэффициент передачи тока в режиме короткого замыкания на выходе по переменному току (безразмерная величина);

$h_{22} = \left. \frac{i_2}{u_2} \right|_{i_1=0}$ — выходная проводимость в режиме холостого хода на входе по переменному току.

Все эти параметры можно измерить приборами и вычислить по статическим характеристикам транзистора. На рис. 3.15 приведены вольтамперные входные и выходные статические характеристики транзистора, включенного по схеме ОБ и ОЭ.

Значения всех параметров транзистора зависят от режимов работы и температуры. При математическом моделировании работы транзисторов существует несколько моделей. В биполярных транзисторах входной управляющей величиной является либо ток базы, либо ток эмиттера. Выходными управляемыми величинами являются эмиттерный или коллекторный токи.

Приведем без выкладок значения некоторых h -параметров

$$h_{11\text{об}} = \frac{U_j}{I_j} \Big|_{U_{\text{к}}=0} \approx r_j + r_{\text{б}}(1-\alpha) \quad (\text{десятки Ом — кОм});$$

$$h_{21\sigma} = \frac{I_k}{I_{\sigma}} \Big|_{U_{\kappa=0}} = \alpha - \frac{r_k}{r_k + r_{\sigma}} \quad (\approx 0,95 \div 0,99);$$

$$h_{12\sigma} = \frac{U_{\sigma}}{U_k} \Big|_{I_{\sigma}=0} \approx \frac{r_{\sigma}}{r_k} + \mu_{\text{жк}} \quad (10^{-4} \div 10^{-3});$$

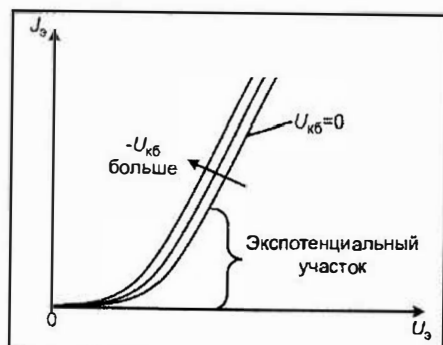
$$h_{22\sigma} = \frac{I_k}{U_k} \Big|_{I_{\sigma}=0} \approx \frac{1}{r_k} \quad (\approx 10^{-7} \div 10^{-6} \text{ Ом}^{-1});$$

$$h_{1\sigma} \approx r_{\sigma} + (1 + \beta)r_{\sigma}, \quad (r_{\sigma} \approx 50200 \text{ Ом});$$

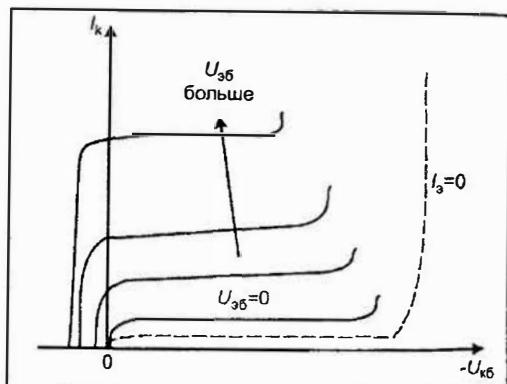
$$h_{12\sigma} \approx \frac{r_{\sigma}}{r_k(1 - \alpha)};$$

$$h_{21\sigma} \approx \beta, \quad (\beta \approx 100);$$

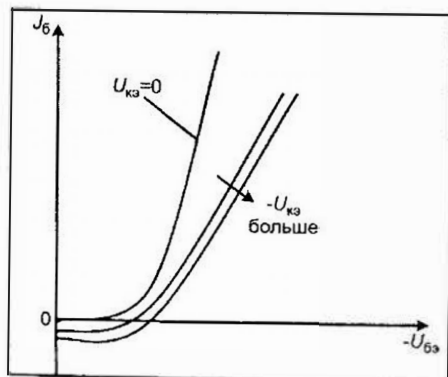
$$h_{22\sigma} \approx \frac{1}{r_k(1 - \alpha)}.$$



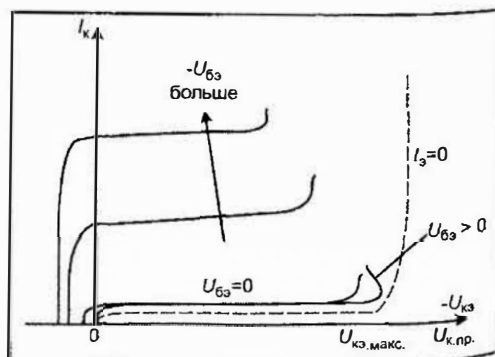
а)



б)



в)



г)

Рис. 3.15. ВАХ транзисторов с общей базой (а, б) и общим эмиттером (в, г)

В выходных характеристиках в качестве параметра выбирается выходное напряжение.

Существуют следующие зависимости:

$$I_K = \alpha I_E,$$

$$I_K = B I_B,$$

$$B = \frac{\alpha}{1 - \alpha},$$

где $\alpha = I_K/I_E \approx \gamma n$; $\alpha = 0,99 \dots 0,995$, $B = 100 \dots 200$.

В свою очередь γ можно определить как

$$\gamma = \left(1 + \frac{W D_2 N_b}{L D_b N_e} \right)^{-1},$$

где W — ширина базы, L_2 — диффузионная длина неосновных носителей в эмиттере, D_2 и D_b — коэффициенты диффузии носителей в базовой и эмиттерной областях, легированных до концентраций N_b и N_e .

Тогда $\gamma = 0,98 \dots 0,99$.

Коэффициент переноса n можно представить как

$$n = 1 - \frac{W^2}{2 D_b \tau},$$

где τ — время жизни неосновных носителей в базе.

Величина n очень близка к единице.

3.3.4. Эпитаксиально-планарные п—р—п-транзисторы

Биполярный эпитаксиально-планарный транзистор п⁺—р—п-типа является одним из основных элементов интегральных схем (рис. 3.12). По техническим параметрам он превосходит транзистор р—п—р-типа.

Минимальные значения горизонтальных размеров транзистора определяются, прежде всего, топологическими нормами литографического процесса и глубиной боковой диффузии примеси под окисел. Топология транзистора может быть асимметричной и симметричной, при которой базовые и коллекторные электроды симметрично облегают эмиттер.

При проектировании транзисторов большой мощности следует обеспечить максимальное отношение периметра эмиттера к его площади. С этой целью создаются "гребенчатые" конструкции, позволяющие создать узкие эмиттеры с большим периметром.

Характеристики транзистора зависят от частоты сигнала, структуры транзистора и наличия в ней паразитных элементов.

Транзисторы п—р—п-типа имеют более высокую предельную частоту, чем р—п—р-транзисторы.

Типичные параметры интегральных п—р—п-транзисторов для интегральных схем приведены в табл. 3.2.

Таблица 3.2

Параметр	Номинал	Допуск	Температурный коэффициент $^{\circ}\text{C}^{-1}$
Коэффициент усиления B	100—200	± 30	$+5 \times 10^{-3}$
Предельная частота f_m , МГц	200—500	± 20	—
Коллекторная емкость C_{κ} , пФ	0.3—0.5	± 10	—
Пробивное напряжение $U_{\text{бв}}$, В	40—50	± 30	—
Пробивное напряжение $U_{\text{кбв}}$, В	7—8	± 5	$+(2-6) \times 10^{-3}$

3.3.5. Многоэмиттерные и многоколлекторные транзисторы

Многоэмиттерные транзисторы (МЭТ) широко используются в микросхемах. Число эмиттеров в них достигает от 3 до 8, и МЭТ можно рассматривать как совокупность транзисторов с общей базой и коллекторами (рис. 3.16, а).

Для подавления паразитных транзисторов p^+-p-p^+ -типа расстояния между соседними эмиттерами должны быть больше диффузионной длины носителей в базовом слое (≈ 10 мкм).

Многоколлекторный биполярный транзистор (МКТ) представлен на рис. 3.16, в. Он представляет собой МЭТ, работающий в инверсном режиме. Общим эмиттером служит общий эпитаксиальный слой, а коллектором является сильно легированные небольшие области p^+ .

При конструировании МКТ основное внимание уделяется обеспечению высокого коэффициента передачи тока от общего эмиттера к каждому из p^+ -коллекторов. Поэтому скрытый слой максимально приближают к базовому слою, а p^+ -области располагают близко друг к другу. Транзисторы этого типа широко применяются в интегральных схемах ИТЛ, ТТЛ и ТТЛШ-логики.

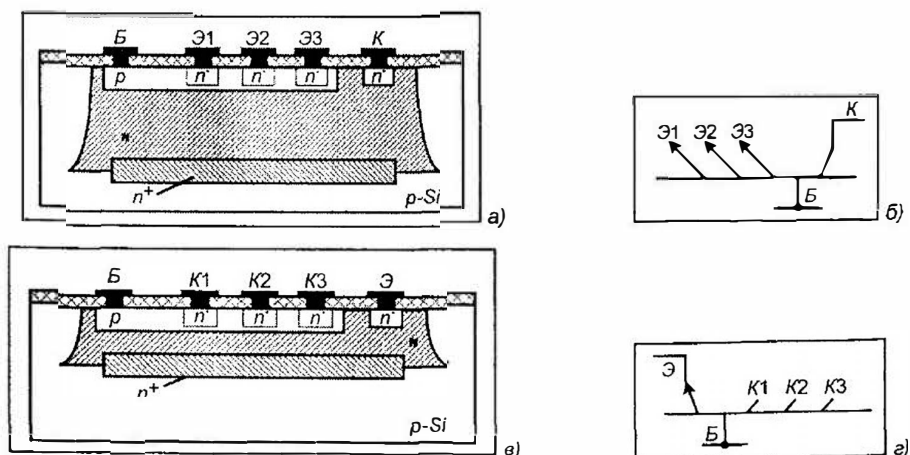


Рис. 3.16. Многоэмиттерный транзистор (а) и его условное обозначение (б), многоколлекторный транзистор (в) и его условное обозначение (г)

3.3.6. Транзисторные структуры интегрально-инжекционной логики

Появление весьма эффективной интегрально-инжекционной логики (I^2L) способствовало разработке биполярного транзистора со встроенным инжектором (рис. 3.17, а). Транзисторы логики состоят из горизонтального транзистора T_1 $p-n-p$ -типа, выполняющего функции генератора тока, и вертикального транзистора тока T_2 $n-p-n$ -типа, выполняющего функции инвертора.

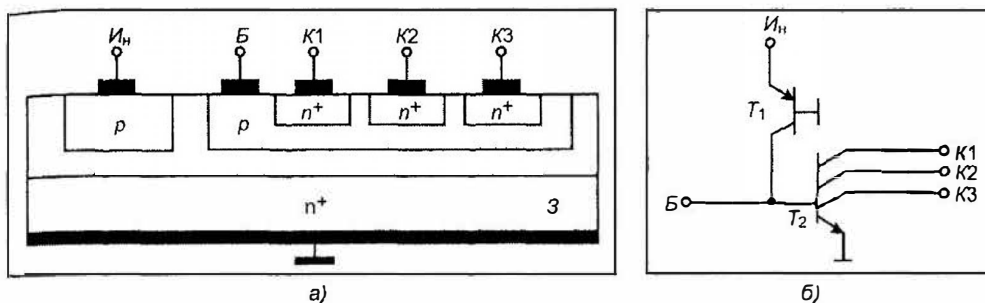


Рис. 3.17. Структура биполярного транзистора для I^2L (а) и схема его включения (б)

Оба транзистора реализованы в одном кристалле так, что базовая область транзистора T_1 с эмиттерной областью транзистора T_2 , а коллекторная область T_1 совмещается с базовой областью T_2 . Эмиттерная область T_1 называется инжектором носителей и подключается к источнику питания. Транзистор T_1 может быть многоэмиттерным, а транзистор T_2 всегда является многоколлекторным, электроды которого являются логическими выходами.

Транзисторы такой конструкции обладают малыми паразитными емкостями, отсутствием процесса накопления зарядов в базовой области, небольшим перепадом уровней сигналов. Это обуславливает высокое быстродействие транзисторной структуры. Интегральные схемы на транзисторах I^2L типа отличаются высокой степенью интеграции и низким уровнем потребляемой мощности.

3.3.7. Транзистор с барьером Шоттки

Обычный биполярный транзистор при работе в активном режиме имеет тенденцию насыщаться носителями в районе базовой области. Избавиться от такого эффекта позволяет включение в структуру транзистора диода. Это позволяет шунтировать базовую и коллекторную область транзистора и предотвратить процесс насыщения носителями базовой области.

На рис. 3.18, а приведена схема включения диода Шоттки, который представляет собой несколько расширенный электрод базы, перекрывающий частично область коллектора.

Такие транзисторы отличаются высоким быстродействием вследствие создания постоянной ненасыщенной структуры. Работа транзисторов Шоттки в ненасыщенном режиме приводит к увеличению падения напряжения на переходе "база—эмиттер". Это позволяет уменьшить ток потребления в статическом режиме и, соответственно, потребляемую мощность.

Транзисторы с диодом Шоттки широко используются при конструировании интегральных логических схем (ТТЛШ) с высоким быстродействием.

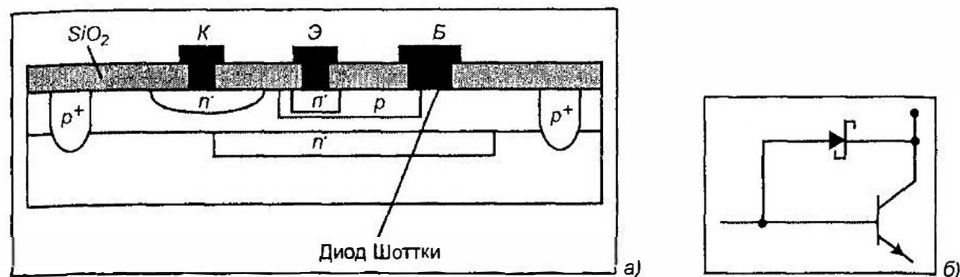


Рис. 3.18. Структура транзистора с диодом Шоттки (а) и его условное обозначение (б)

3.4. Перспективные транзисторные структуры

Основная тенденция развития микроэлектроники заключается в непрерывном росте степени интеграции элементов на кристалле.

В соответствии с этой тенденцией уменьшаются линейные размеры транзисторных структур и, соответственно, увеличивается плотность размещения элементов на кристалле. Естественно, что такая тенденция не может продолжаться, и прежде всего из-за достижения физических пределов отдельных транзисторных структур.

Физические процессы диффузии, дрейфа, статистические законы, законы термодинамики работают в определенных граничных и начальных условиях.

Уменьшение геометрических параметров транзисторных структур, переход от микро- к наноразмерам приводит к переходу в другую область физики явлений, к принципиально другому процессу обработки информационных сигналов.

Рассмотрим пример. Уменьшение длины затвора и толщины затворного оксида транзистора является одним из условий повышения быстродействия интегральных схем вообще и микропроцессоров в частности.

Современные транзисторные структуры имеют толщину подзатворного оксида $\sim 0,8$ нм, что оценивается толщиной в три атомных слоя. При этом наблюдается рост тока утечки, вызванный процессами туннелирования через тонкий слой оксида. Даже в отключенном транзисторе происходят процессы утечки зарядов.

С уменьшением толщины электродов истока, стока и затвора возрастает их омическое сопротивление и, как следствие, необходимо большее значение напряжения для переключения прибора, увеличения потребляемой мощности.

Сравнительный анализ показывает, что по плотности потребляемой мощности самые распространенные микропроцессоры Pentium превосходят разогретую нагревательную плитку (рис. 3.19). Следовательно, необходима разработка системы эффективного теплоотвода.

Для выхода из сложившегося тупика предложено несколько вариантов. Первый вариант связан с увеличением числа логических ячеек, работающих при разных пороговых напряжениях, соответствующих их структуре. Токи утечки при этом могут быть снижены в транзисторных структурах, к быстродействию которых не предъявляются жесткие требования. Все это связано с новым подходом к процессу проектирования интегральных схем. Другой вариант связан с созданием принципиально новых транзисторных структур.

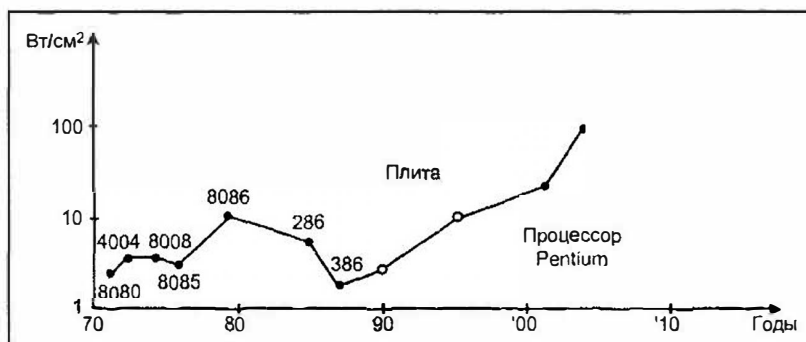


Рис. 3.19. Рост плотности потребляемой мощности микропроцессорами фирмы Intel

3.4.1. Транзисторы со сверхтонким основанием

Транзисторы со сверхтонкой пленкой-основанием выполняются на пленке кремния толщиной 30—40 нм, нанесенной сверху оксидного слоя. При этом это основание может быть частично или полностью обеднено носителями. Если тонкая пленка полностью обеднена подвижными носителями при всех значениях напряжения смещения, то в области канала заряда нет. В этом случае электрическое поле в инверсионном слое прибора меньше, чем в обычных приборах с сильнолегированной областью канала, выполненных на объемном материале. Такие транзисторы обладают большой крутизной ВАХ. Такая структура лежит в основе Terahertz-транзисторов фирмы Intel, частота переключения которых составляет 1000 ГГц или 1 ТГц (рис. 3.20).

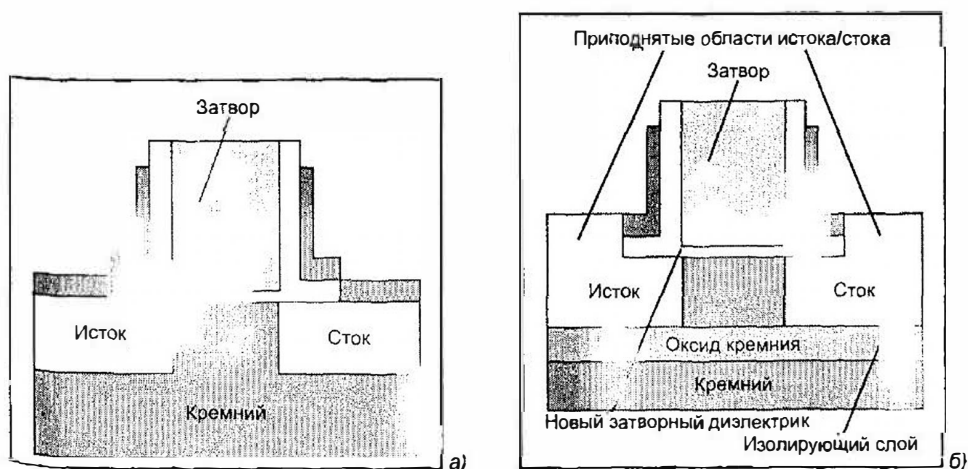


Рис. 3.20. Структура МОП транзистора (а) и Terahertz -транзистора (б)

Транзистор изготавливается на слое из кремния толщиной 30 нм, использование оксида кремния вместо традиционного диоксида кремния позволяет снизить ток утечки через затворный диэлектрик на четыре порядка.

Terahertz-транзисторы превосходят стандартные КМОП-структуры по быстродействию на 25% и по потребляемой мощности на 30%. При напряжении 1,3 В рабочий ток транзистора равен 650 мкА/мкм, а ток утечки всего 9 нА/мкм.

Как предполагается, такие транзисторы станут базовыми разработками микропроцессоров фирмы Intel с минимальными топологическими нормами 20 нм, быстродействием 20 ГГц и рабочим напряжением 1 В. В одном чипе микропроцессора разместится 10^6 Terahertz-транзисторов.

3.4.2. МОП-транзисторы с управляемой проводимостью канала

Транзистор с управляемой проводимостью канала характеризуется высокой подвижностью носителей в области канала.

В основе конструкции этих транзисторов лежит традиционная конструкция МОП-транзистора. Особенность заключается в технологии применения слоя напряженного кремния, осаждаемого на подложку соединения SiGe. Из рис. 3.21, а хорошо видно, что постоянная кристаллической решетки обычного кремния и твердого раствора SiGa различны.

При эпитаксиальном выращивании кремния на SiGe-подложке атомы кремния стремятся подстроиться под структуру подложки, и поэтому происходит растягивание решетки кремния. В этом случае слой кремния становится напряженным. Оказывается, что скорость дрейфа электронов в таком материале на 70% выше, чем в обычном кремнии. Это позволяет, не масштабируя геометрические параметры транзистора, увеличить его быстродействие на 35%.

Такая технология позволяет создавать транзисторные структуры как в объеме, так и в пленке. Интегральные схемы на таких транзисторах будут обладать высокой плотностью размещения элементов, уменьшенным значением рассеивания мощности.

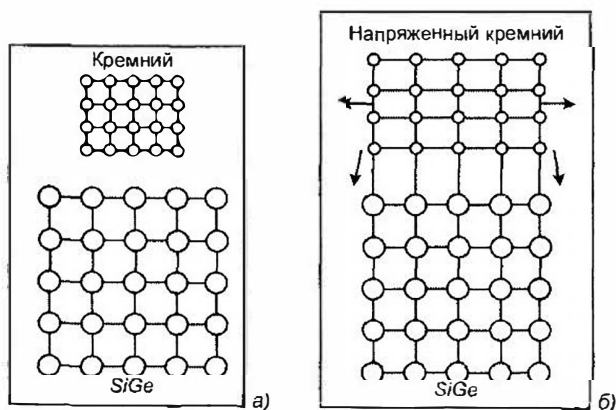


Рис. 3.21. Формирование напряженного слоя кремния на SiGe-подложке

3.4.3. МОП-транзисторы с двойным затвором

Среди прочих существуют также конструкции *МОП-транзисторов с двумя затворами*. Фирма IBM разработала транзистор с двумя затворами принципиально новой конструкции. Конструкция FinFET-транзистора представляет собой кремниевое тело (столбик,

вставка-fin), которое обернуто затвором. Конструкция затвора такова, что формируются два самосовмещенных канала с двух сторон кремниевое тела (рис. 3.22, а). Выступающая передняя область тела представляет собой исток транзистора, а задняя область — сток.

Ток течет в плоскости, параллельной плоскости тела. Активная ширина прибора W равна высоте тела-столбика и может быть увеличена за счет включения нескольких столбиков.

Структура FinFET аналогична традиционной МОП-структуре, хотя и является квазипланарной. Отличительной особенностью является активная область, которая в данном случае формируется вставками. Высота тела вставки составляет 180 нм, толщина затворного диэлектрика — 2,2 нм, поликремниевое затворное — 75 нм.

Фирмой IBM разработаны симметричные и асимметричные n - и p -канальные транзисторы с длиной канала ~ 30 нм. Типичный ток в n -канальном транзисторе составляет 1300 мкА/мкм, в p -канальном — 850 мкА/мкм. Усилиями разработчиков Калифорнийского университета в Беркли (США), фирмы Intel созданы FinFET-транзисторы с длиной канала 20 нм, в котором размеры кремниевой вставки определяются промежутками между поликремниевыми затворами, а области истока и стока формируются процессами литографии.

Характеристики FinFET-транзисторов позволяют надеяться на их использование в интегральных схемах с увеличенной плотностью упаковки и уменьшенной рассеиваемой мощностью.

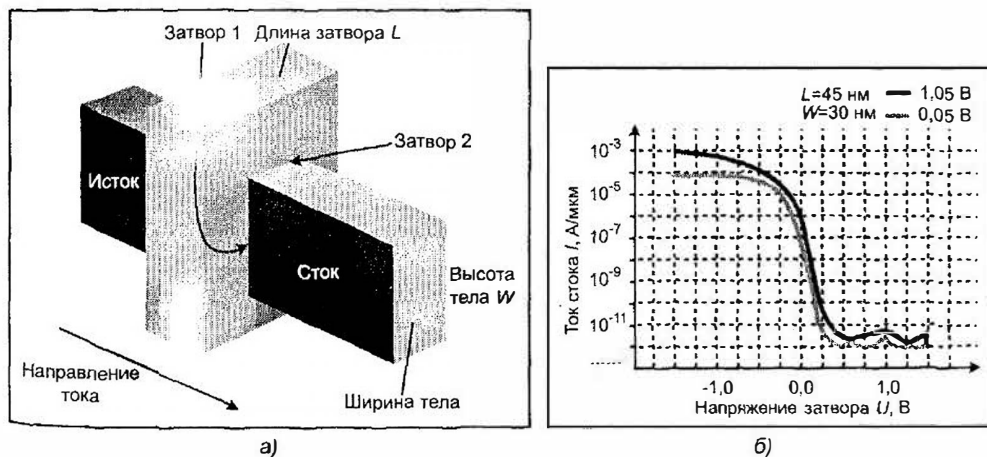


Рис. 3.22. Квазипланарная структура FinFET-транзистора (а) и его стоково-затворная характеристика (б)

3.4.4. Вертикальные МОП-транзисторы

Приборы с вертикальным затвором (verticals replacement-gate-VRG) позволяют с высокой точностью формировать длину канала.

В таких транзисторах токи утечки очень малы. В качестве затворного диэлектрика с высокой диэлектрической проницаемостью и используется оксид гафния (HfO_2).

Разработанные вертикальные полевые транзисторы на основе GaAs имеют ряд преимуществ по сравнению с кремниевыми. Эти преимущества связаны с более высокой под-

вижностью и дрейфовой скоростью носителей, а также с возможностью использования гетеропереходов для инъекции электронов в активную область транзистора. На рис. 3.23 приведены структуры вертикального полевого транзистора на основе GaAs. Транзисторы этого типа имеют короткий канал, высокое быстродействие, малую потребляющую мощность. Этот тип транзисторов может стать основой разработок GaAs схем с высокой плотностью упаковок.

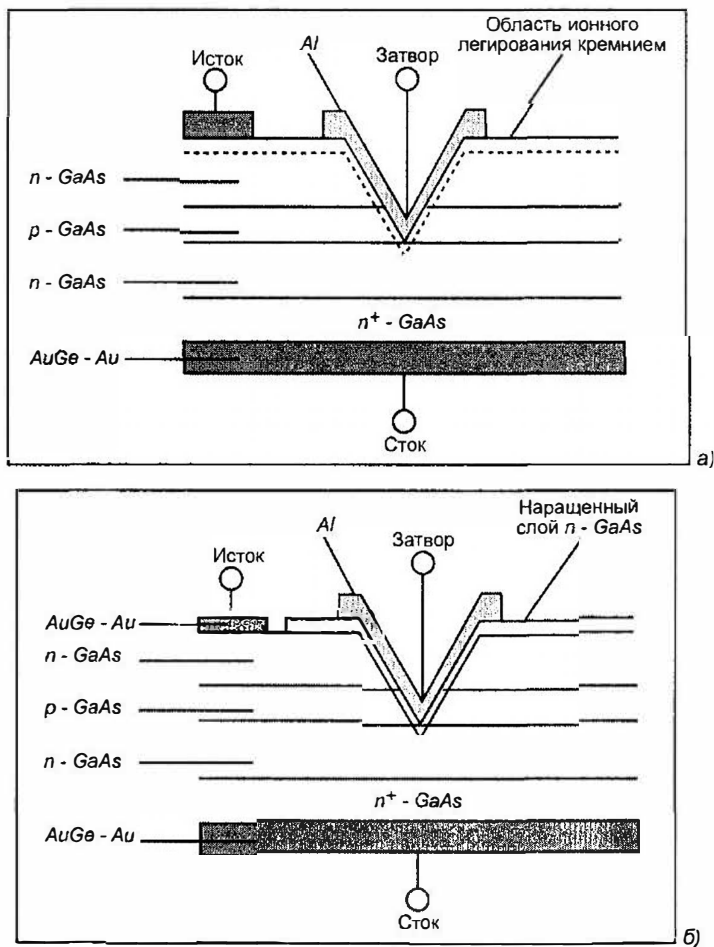


Рис. 3.23. Структура вертикального GaAs полевого транзистора с ионно-легированным затвором Шоттки на GaAs (а) и затвором Шоттки с наращенным слоем n-GaAs (б)

3.4.5. Арсенид-галлиевые транзисторы

Развитие *арсенид-галлиевых транзисторов* идет по пути увеличения скорости электронов, уменьшения их пролетного времени, переходом к баллистическому режиму работы. Следствием этих факторов является сокращение размеров и уменьшение потребляемой

мощности. Полевые транзисторы на гетероструктурах с селективным легированием используют свойства двумерного электронного газа. Электроны движутся от истока к стоку в тонком двумерном инверсионном слое на границе между широкозонным AlGaAs и узкозонным нелегированным GaAs (рис. 3.24).

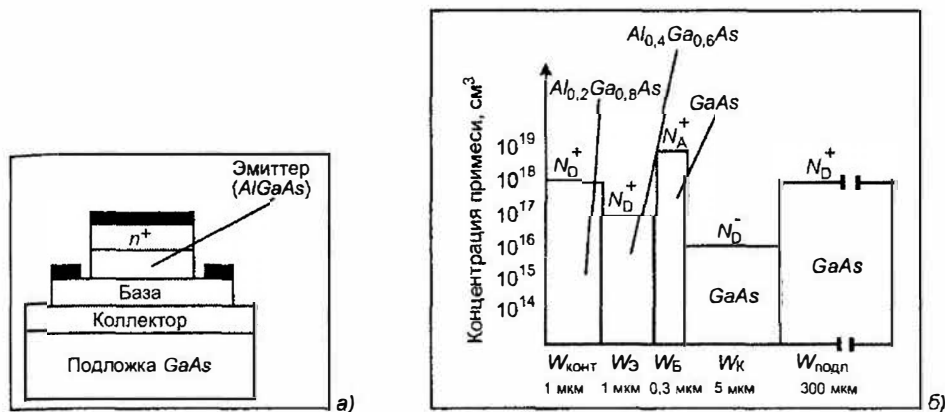


Рис. 3.24. Структура биполярного гетеротранзистора (а) и соответствующий профиль легирования (б)

Тонкий слой нелегированного AlGaAs разделяет слой двумерных электронов в GaAs от доноров в легированном слое AlGaAs под затвором. Уменьшение примесного рассеяния приводит к увеличению подвижности носителей, уменьшает сопротивление канала и увеличивает крутизну вольтамперной характеристики. Быстродействие таких транзисторов близко к рекордному значению и определяется временем задержки сигнала, составляющим при комнатной температуре порядка 10 пс.

Транзистор с проникающей базой формируется путем размещения тонкой вольфрамовой сетки внутри арсенид-галлиевого эпитаксиального слоя (рис. 3.25). Такой транзистор относится к вертикальным структурам.

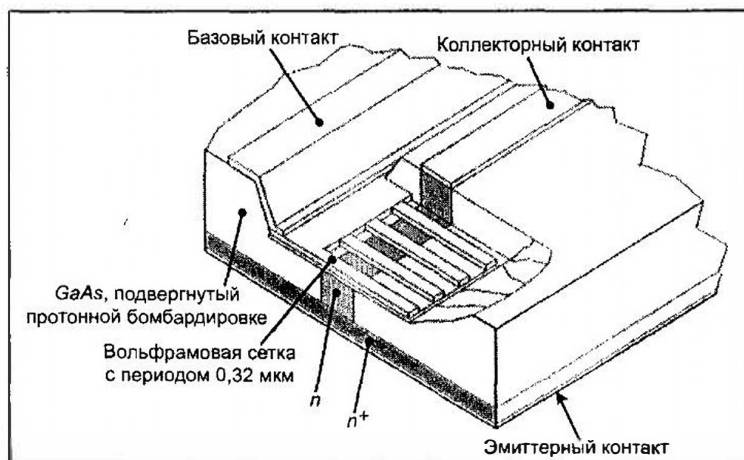


Рис. 3.25. Схема транзистора с проникающей базой

Металл образует барьер Шоттки GaAs, встроенный потенциал которого полностью обедняет GaAs в местах между полосками сетки. Положительный потенциал на базовом электроде уменьшает ширину обедненного слоя, и между эмиттером и коллектором возникают проводящие каналы. Электроны движутся от эмиттера к коллектору через окно в сетке транзистора. Также транзисторы могут работать на частоте до 200 ГГц.

Вертикальной структурой является также баллистический транзистор (рис. 3.26). Поле барьера Шоттки затворного электрода можно изменять проводимость канала. Особенность конструкции является расположение затвора на поверхности, а не в объеме. Этот прием упрощает технологию изготовления.

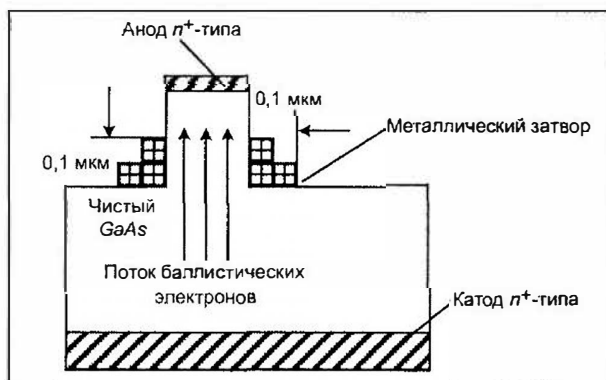


Рис. 3.26. Структура вертикального баллистического транзистора

Задачи и упражнения

3.1. Определение параметров полевого транзистора

Удельная проводимость канала n-типа полевого транзистора $\sigma = 20,9 \text{ Ом/м}$ и ширина канала $\omega = 6 \text{ мкм}$ при напряжении "затвор—исток", равным нулю.

1. Найти напряжение отсечки $U_{\text{отс}}$ считая, что подвижность электронов $\mu_n = 0,13 \text{ м}^2/(\text{В}\cdot\text{с})$, а относительная диэлектрическая проницаемость кремния $\epsilon = 12$.
2. При напряжении затвора, равном нулю, сопротивление "сток—исток" равно 50 Ом. При каком напряжении затвора сопротивление "сток—исток" станет равным 200 Ом?

Решение

1. Удельная проводимость канала n-типа

$$\sigma = \mu_n q N_d$$

где N_d — концентрация примесных атомов; μ_n — подвижность электронов.

Следовательно,

$$N_d = \sigma / \mu_n q = 20,9 / (0,13 \cdot 1,6 \cdot 10^{-19}) = 10^{21} \text{ м}^{-3}.$$

Напряжение отсечки

$$U_{отс} = qN_d a^2 / 2\epsilon\epsilon_0,$$

где a — половина ширины канала при напряжении "затвор—исток" $U_{зи} = 0$:

$$U_{отс} = 1,6 \cdot 10^{-19} \cdot 10^{21} \cdot (3,0 \cdot 10^{-6})^2 / (2 \cdot 12 \cdot 8,85 \cdot 10^{-12}) = 6,8 \text{ В.}$$

2. Сопротивление участка "сток—исток" транзистора при низком напряжении можно представить в виде резистора, сопротивление которого определяется как

$$R_{си} = \frac{R_0}{1 - (|U_{зи}| / U_{отс})^{1/2}},$$

где R_0 — сопротивление "сток—исток" при $U_{зи} = 0$. При $R_{си} = 200$ Ом имеем:

$$200 = \frac{50}{1 - (|U_{зи}| / 6,8)^{1/2}},$$

откуда $|U_{зи}| = 3,83$ В (отрицательно относительно истока).

3.2. Определение крутизны МОП-транзистора

В МДП-транзисторе с каналом п-типа ширина затвора $0,8$ мм, длина канала $l = 5$ мкм, толщина (оксидного) слоя диэлектрической изоляции $d = 150$ нм, подвижность электронов в канале $\mu_n = 0,02$ м²/(В·с), относительная диэлектрическая проницаемость оксидной пленки $\epsilon = 3,7$, напряжение "сток—исток" в пологой части характеристики (при насыщении) $U_{си} = 8$ В. Определить крутизну прибора в области насыщения.

Решение

Крутизну прибора в области насыщения вычислим по следующей формуле:

$$S = \mu_n C_s U_{си} l^2,$$

где μ_n — подвижность электронов в канале, C_s — емкость затвора; l — длина канала; $U_{си}$ — напряжение "сток—исток" при насыщении.

Определим емкость затвора

$$C_s = \epsilon\epsilon_0 l / d = 3,7 \cdot 8,85 \cdot 10^{-12} \cdot 0,8 \cdot 10^{-3} \cdot 5 \cdot 10^{-6} / (150 \cdot 10^{-9}) = 0,87 \text{ пФ.}$$

Следовательно, крутизна

$$S = 0,02 \cdot 0,87 \cdot 10^{-12} \cdot 8 / (5 \cdot 10^{-6})^2 = 5,57 \text{ мА/В.}$$

3.3. Определение параметров канала

Полевой МОП-транзистор с каналом р-типа работает в режиме обогащения и имеет следующие параметры: ширина и длина затвора $Z = W = 50$ мкм; длина канала $L = 5$ мкм; толщина слоя окисла $x_{ок} = 0,1$ мкм; пороговое напряжение $U_{пор} = -1$ В; $\epsilon_{ок} = 4$; $\mu_p = 190$ см²·В⁻¹·с⁻¹.

1. Вычислите ток I_c , сопротивление канала $r_c = 1 / g_c$ и крутизну S , если прибор работает в линейном режиме при напряжениях $U_g = 3$ В и $U_c = -0,1$ В.
2. Определите значения $I_{нас}$ и $S_{нас}$, считая, что транзистор работает при напряжениях $U_g = -4$ В и $U_c = -5$ В.

Решение

1. Ток стока определяется по формуле

$$I_c = \frac{Z}{L} \mu_p C_0 \left[(U_s - U_{\text{нпр}}) U_c - \frac{1}{2} U_c^2 \right] =$$

$$= 10 \cdot 190 \cdot 3,54 \cdot 10^{-8} \cdot [(-371) \cdot (-0,1) - 0,5 \cdot 0,01] = 13,12 \text{ мкА}$$

$$g_c = \frac{Z}{L} \mu_p C_0 |U_s - U_{\text{нпр}} - U_c| =$$

$$10 \cdot 190 \cdot 3,54 \cdot 10^{-8} \cdot |-3 - (-0,1)| = 1,28 \times 10^{-4} \text{ см};$$

$$r_c = 1 / g_c = 7,825 \text{ кОм};$$

$$S = \frac{Z}{L} \mu_p C_0 |U_c| = 10 \cdot 190 \cdot 3,54 \cdot 10^{-8} \cdot 0,1 = 6,73 \text{ мкА/В.}$$

2. Ток насыщения

$$I_{c\text{нас}} = \frac{Z}{L} \mu_p C_0 \frac{(U_s - U_{\text{нпр}})^2}{2} = 0,303 \text{ мА},$$

$$S_{\text{нас}} = \frac{Z}{L} \mu_p C_0 |U_s - U_{\text{нпр}}| = 0,202 \text{ мА/В.}$$

3.4. Пороговое напряжение транзистора

Вычислите пороговое напряжение МОП-транзистора с каналом р-типа и ориентацией поверхности кристалла (111). Транзистор имеет алюминиевый затвор, толщина оксидного слоя 120 нм, концентрация легирующей примеси в подложке $N_{d2} = 3 \times 10^{15} \text{ см}^{-3}$. Первоначальная концентрация $N_{d1} = 10^{15} \text{ см}^{-3}$.

Из-за роста концентрации изменяется уровень E_F , т. к.

$$\left. \begin{aligned} E_{F1} - E_i &= kT \ln \frac{N_{d1}}{n_i} \\ E_{F2} - E_i &= kT \ln \frac{N_{d2}}{n_i} \end{aligned} \right\} \Rightarrow qU_F = E_{F2} - E_{F1}.$$

Это приводит к тому, что изменяется величина $\Phi_{\text{ин}} = -0,3 + U_F$. При ориентации (111) поверхностная плотность заряда, локализованного на границе раздела "полупроводник—оксид", $Q_{\text{ss}} = 5 \times 10^{-11} \text{ Кл/см}^2$, $q = 8 \times 10^{-8} \text{ Кл/см}^2$. Считается, что $n_i = 1,4 \times 10^{10} \text{ см}^{-3}$.

Решение

$$qU_F = E_{F2} - E_{F1} = kT \ln(N_{d2}/N_{d1}) = 0,026 \ln(3 \cdot 10^{15}/10^{15}) = 0,03 \text{ эВ},$$

$$U_F = 0,03 \text{ В.}$$

При $x_{\text{ок}} = 120 \text{ нм}$ удельная емкость:

$$C_0 = \epsilon_{\text{окс}} / x_{\text{ок}} = 2,84 \times 10^{-8} \text{ Ф/см}^2;$$

$$U_n = U_S = -2U_F = -2U_F \ln(N_{d2}/n_i) = -2 \cdot 0,026 \ln[3 \cdot 10^{15}/(1,4 \cdot 10^{10})] = -0,64 \text{ В};$$

$$Q_H = (2\epsilon_{ин} q N_{д2} |2U_F|)^{1/2} = (2 \cdot 12 \cdot 8,85 \cdot 10^{-14} \cdot 1,6 \cdot 10^{-19} \cdot 3 \cdot 10^{15} \cdot 2 \cdot 0,64)^{1/2} = 2,55 \cdot 10^{-8} \text{ Кл/см}^2;$$

$$U_{пор} = \Phi_{МП} - \frac{Q_H}{\epsilon_{ок}} + U_S - \frac{Q_B}{C_0} = -0,27 \frac{5 \cdot 10^{11} \cdot 1,6 \cdot 10^{-19}}{2,84 \cdot 10^{-8}} - 0,64 \frac{3,61 \cdot 10^{-8}}{2,84 \cdot 10^{-8}} = -5 \text{ В.}$$

3.5. Определение тока через базу р-п-р-транзистора

Объясните механизм протекания тока через базу р-п-р-транзистора, работающего в условиях нормального смещения. Для иллюстрации постройте зонную диаграмму.

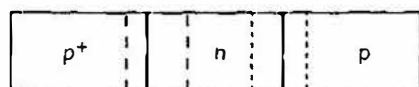
Пусть транзистор имеет следующие параметры: эффективность эмиттера 99%, коэффициент переноса 99,5%, коэффициент ударной ионизации 100%. Вычислите ток коллектора, если ток базы равен 20 мкА, а ток утечки "коллектор-база" при разомкнутой цепи эмиттера составляет 1 мкА.

Решение

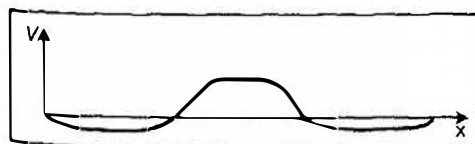
Как в р-п-р, так и в п-р-п-транзисторах при нормальном смещении переход эмиттер — база смещен в прямом направлении, а переход "коллектор-база" — в обратном направлении.

Однако в р-п-р-транзисторе электропроводность эмиттера р-типа обычно выше, чем электропроводность базы п-типа, и на обоих переходах преобладает дырочный ток. Таким образом, ток в р-п-р-транзисторе обусловлен главным образом дырками.

На рис. 3.27 показаны распределения напряжения V и энергии ϵ в р-п-р-переходе при нулевом смещении. Толщина обедненного слоя в р-области меньше, чем в п-области, уровень Ферми ϵ_F не изменяется во всех трех областях, и результирующий ток равен нулю.

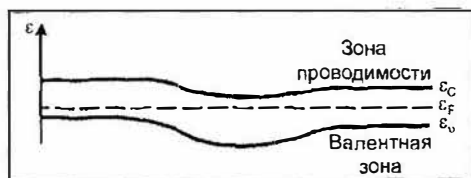


а)



б)

Рис. 3.27. Профиль транзистора (а) и распределение напряжения (б) и энергии зон по его длине (в)



в)

На рис. 3.28 приведены распределения напряжения V и энергии ϵ при нормальном рабочем смещении. Видно, что толщины обедненных слоев, а следовательно, и распределения напряжения и энергии изменяются.

У границы левого р-п-перехода устанавливается избыточная концентрация дырок, в результате чего происходит их инжекция с границы обедненного слоя "эмиттер-база". Инжектированные дырки диффундируют через область базы, в которой незначительное их число рекомбинирует с электронами. Следовательно, существует поток электронов в

базовую область (из внешней цепи), непрерывно восстанавливающий число электронов, теряемых при рекомбинации. Остальные дырки, не участвующие в рекомбинации, поступают на коллектор.

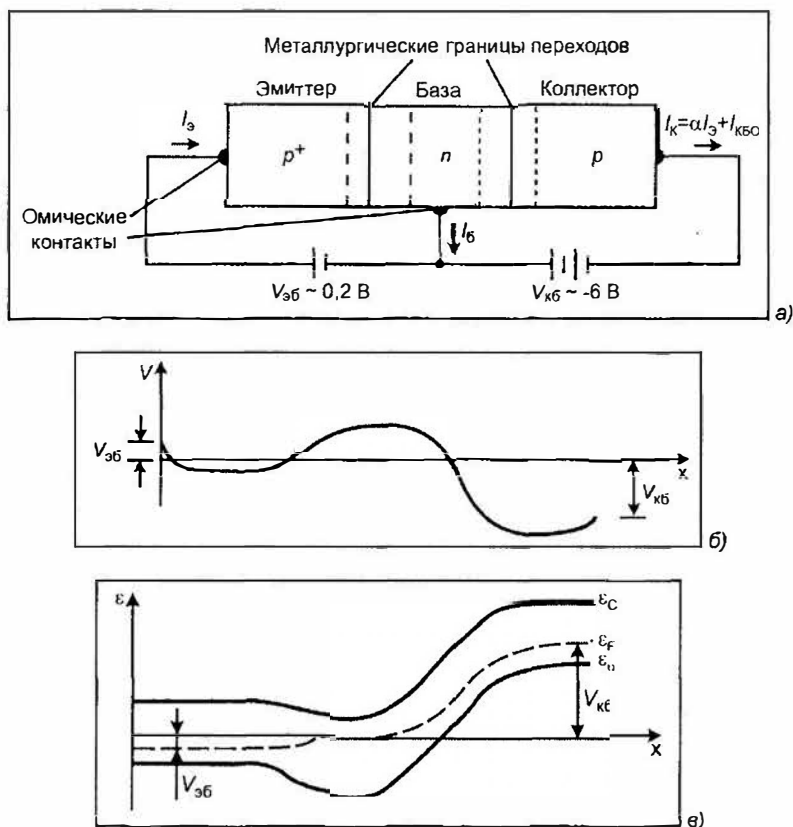


Рис. 3.28. Профиль транзистора (а) и распределение напряжения (б) и энергии (в) при нормальном смещении

Транзисторный эффект имеет место в том случае, если толщина базы между переходами меньше диффузионной длины неосновных носителей заряда (дырок) в базе. Таким образом, дырки, инжектируемые в область базы, диффундируют непосредственно к коллектору.

Поскольку типичные значения L_p и W равны 10^{-4} и 10^{-5} м, то рекомбинация незначительна.

Коэффициент усиления по постоянному току α транзистора в схеме с общей базой зависит главным образом от трех факторов:

- ☐ эффективности эмиттера γ ;
- ☐ коэффициента переноса β ;
- ☐ коэффициента ударной ионизации M .

Можно записать

$$\alpha = \gamma \beta M, \quad (3.5.1)$$

Но, поскольку $M = 1$,

$$\alpha = \gamma \beta. \quad (3.5.2)$$

Подставляя числовые данные в это выражение, получаем

$$\alpha = 0,99 \cdot 0,995 = 0,985.$$

Из рис. 3.28 видно, что ток коллектора определяется выражением

$$I_k = \alpha I_o + I_{кбо}, \quad (3.5.3)$$

где $I_{кбо}$ — ток утечки (ток коллектора, когда на базу подано обратное смещение относительно коллектора, а эмиттерная цепь разомкнута).

Ток эмиттера равен

$$I_o = I_b + I_k.$$

Подставим это равенство в выражение (3.5.3):

$$I_k = \alpha(I_b + I_k) + I_{кбо}$$

или

$$I_k(1 - \alpha) = \alpha I_b + I_{кбо}.$$

Отсюда

$$\begin{aligned} I_k &= \frac{\alpha}{1 - \alpha} I_b + \frac{1}{1 - \alpha} I_{кбо} = \\ &= \frac{0,985}{1 - 0,985} \cdot 20 \cdot 10^6 + \frac{1}{1 - 0,985} \cdot 1 \cdot 10^6 = 1,38 \text{ мА}. \end{aligned}$$

3.6. Электронная и дырочная компоненты тока

Начертите схему, показывающую разделение токов на электронную и дырочную компоненты, и укажите направление этих компонент в различных областях р—п—р-транзистора, работающего в условиях нормального смещения. Исходя из диаграммы, покажите, что ток базы $I_b = I_o(1 - \alpha) - I_{кбо}$, где α — произведение трех величин: эффективности эмиттера γ , коэффициента переноса β и коэффициента ударной ионизации M .

Пусть в транзисторе, включенном по схеме с общим эмиттером, ток утечки $I_{кбо} = 100$ нА для получения общего тока коллектора 1 мА ток базы должен быть равен 10 мкА. Вычислите коэффициент усиления по постоянному току для этого транзистора и ток утечки $I_{кбо}$.

Решение

На рис. 3.13 (см. выше) схематически показано распределение токов на электронную и дырочную компоненты и их направление в различных областях р—п—р-транзистора при нормальном смещении: эмиттерный переход смещен в прямом направлении, а коллекторный — в обратном.

Наглядно показаны величина и направление этих компонент.

Токи $\gamma(1 - \beta)I_o$ и $\beta\gamma I_o + I_{кбо}$ являются дырочными компонентами, а $(1 - \gamma)I_o$ и $I_{кбо}$ — электронными компонентами.

Если I_e — ток эмиттера, тогда ток, достигающий перехода "эмиттер—база", равен γI_e (напомним, что γ — эффективность эмиттера). Следовательно, электронная компонента равна $(1 - \gamma)I_e$. Дырочная компонента тока, достигающего перехода база — коллектор, равна $\gamma\beta I_e$, где β — коэффициент переноса. Поэтому дырочный ток в базе, обусловленный рекомбинацией носителей, равен $\gamma(1 - \beta)I_e$. Учитывая ток I_{kbo} обусловленный тепловой генерацией, ток коллектора можно записать в виде

$$I_K = \beta\gamma I_e + I_{kbo}$$

или

$$\alpha I_e + I_{kbo},$$

где $\alpha = \beta\gamma$ (в предположении, что коэффициент ударной ионизации $M = 1$). Таким образом, в р—п—р-транзисторе ток обусловлен в основном дырками. Полный ток базы

$$I_b = I_e(1 - \gamma) + (1 - \beta)\gamma I_e - I_{kbo} = I_e(1 - \alpha) - I_{kbo}. \quad (3.6.1)$$

Преобразуя это выражение, получаем, что ток эмиттера

$$I_e = \frac{I_b}{1 - \alpha} + \frac{I_{kbo}}{1 - \alpha},$$

а ток коллектора

$$I_K = \frac{\alpha}{1 - \alpha} I_b + \frac{I_{kbo}}{1 - \alpha}. \quad (3.6.2)$$

Когда $I_b = 0$,

$$I_K = I_{kbo} = \frac{I_{kbo}}{1 - \alpha}.$$

Подставив в (3.6.2) числовые данные, находим

$$\alpha = 0,99,$$

$$I_{kbo} = 100 \cdot 100 \cdot 10^{-9} = 10 \text{ мкА}.$$

3.7. Распределение избыточной концентрации в базе и вычисление тока эмиттера и дифференциального сопротивления между коллектором и базой

Рассмотрите р—п—р-транзистор, в котором площадь как коллекторного, так и эмиттерного переходов равна $1 \times 10^{-6} \text{ м}^2$, а коэффициент диффузии дырок в базе $D_p = 5 \times 10^{-3} \text{ м}^2/\text{с}$.

При $V_{kb} = 1,0 \text{ В}$ распределение концентрации дырок в базе имеет вид, показанный на рис. 3.29.

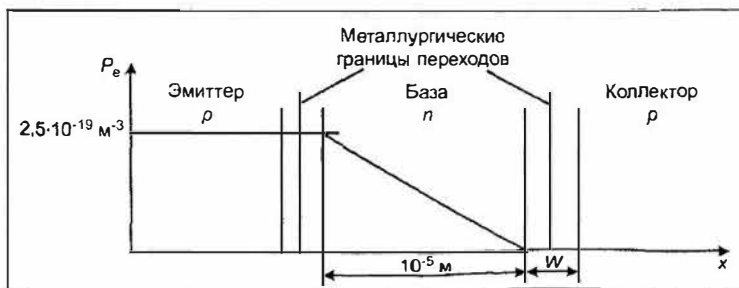


Рис. 3.29. Распределение концентрации дырок в базе

Определите ток эмиттера, обусловленный дырками, пренебрегая токами утечки. Вычислите дифференциальное сопротивление между коллектором и базой при $V_{кб} = 9,0$ В, если толщина обедненного слоя коллекторного перехода $W' = (1 + \sqrt{V_{кб}}) \times 10^{-6}$ м; предположите, что условия на переходе "эмиттер—база" не изменяются и соответствуют рисунку.

Решение

Плотность тока неосновных носителей (дырок) в базе определяется выражением

$$J_p = -q \cdot D_p \cdot dp_c / dx.$$

Следовательно, дырочный ток эмиттера

$$I_p = -q \cdot D_p \cdot dp_c / dx \cdot S,$$

где S — площадь перехода.

Наклон распределения избыточной концентрации дырок (рис. 3.29)

$$\frac{dp_c}{dx} = -\frac{p_c}{W'} = -\frac{2,5 \cdot 10^{19}}{10^{-5}} = -2,5 \cdot 10^{24} \text{ дырок} \cdot \text{м}^{-3} / \text{м},$$

$$I_p = -1,6 \cdot 10^{-19} \cdot 50 \cdot 10^{-4} \cdot (-2,5 \cdot 10^{24}) \cdot 10^6 = 2 \cdot 10^{-3} \text{ А}.$$

Пусть $W_{\text{колл}}$ — полное расстояние между границей обедненного слоя коллектора со стороны коллектора и границей обедненного слоя эмиттера со стороны базы.

Ток коллектора

$$I_k = q D_p S p_c / (W_{\text{колл}} - W) = q D_p S p_c / W'_b,$$

где W'_b — эффективная ширина базы. Таким образом,

$$W_{\text{колл}} = W'_b + W = 10^{-5} + (1 + \sqrt{V_{кб}}) \cdot 10^{-6}.$$

Если $V_{кб} = 1$ В, то $W_{\text{колл}} = 1,2 \times 10^{-5}$ м. Ток коллектора вычисляется по формуле

$$I_k = q D_p S p_c / [W_{\text{колл}} - (1 + \sqrt{V_{кб}}) \cdot 10^{-6}].$$

Продифференцируем это выражение по $V_{кб}$:

$$\frac{dI_k}{dV_{кб}} = -\frac{e D_p A p_c (-V_{кб}^{-1/2} / 2) \cdot 10^6}{[W_{\text{колл}} - (1 + \sqrt{V_{кб}}) \cdot 10^{-6}]^2}.$$

Если $V_{кб} = 9$ В, то

$$\frac{dI_k}{dV_{кб}} = \frac{10^{-4}}{3 \cdot 0,64} \text{ см}.$$

Следовательно, дифференциальное сопротивление $r = dV_{кб} / dI_k = 19,2$ кОм.

3.8. Модель транзистора по Эберсу — Моллу

Эквивалентная схема транзистора по Эберсу — Моллу для постоянного тока и вывод выражений для $V_{эб}$, $V_{кб}$ и $V_{кэ}$.

На рис. 3.30 приведена эквивалентная схема р—п—р-транзистора, включенного по схеме с общей базой.

С помощью модели идеальных диодов Эберса — Молла покажите, что

$$V_{эб} = V_T \ln[I_3(1 - \alpha_R) - \alpha_R I_6 + I_{эб0} / I_{кб0}]$$

$$V_{кб} = V_T \ln[I_3(1 - \alpha_F) - I_6 + I_{кб0} / I_{эб0}],$$

где $V_T = kT/q$ — термо-ЭДС, α_R — коэффициент усиления по току в режиме короткого замыкания в схеме с общей базой для больших сигналов в обратном направлении, α_F — коэффициент усиления по току в режиме короткого замыкания для больших сигналов в прямом направлении, $I_{эб0}$ — ток насыщения эмиттерного перехода с разомкнутым коллектором, $I_{кб0}$ — ток насыщения коллекторного перехода с разомкнутым эмиттером.

Предположите, что

$$\alpha_R = I_{кбS} = \alpha_F I_{эбS},$$

где $I_{эбS}$ — ток насыщения эмиттерного перехода с коллектором, замкнутым на базу, $I_{кбS}$ — ток насыщения коллекторного перехода с эмиттером, замкнутым на базу.

Найдите выражение для $V_{кэ}$.

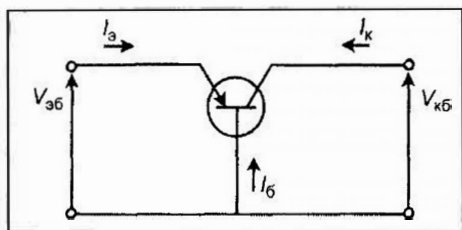


Рис. 3.30. Схема транзистора, включенного по схеме с общей базой

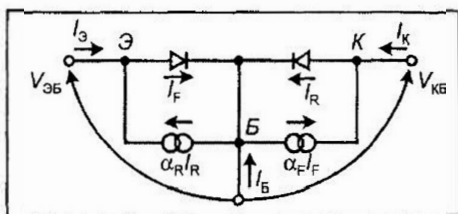


Рис. 3.31. Эквивалентная схема транзистора

Решение

Эквивалентная схема р-п-р-транзистора по Эберсу—Моллу для постоянного тока состоит из двух диодов и показана на рис. 3.31.

Уравнение идеального диода для прямого или обратного напряжения смещения имеет вид:

$$I = I_0 [\exp(V/V_T) - 1].$$

Уравнения для прямого и обратного токов в переходах "эмиттер—база" и "коллектор—база" записываются следующим образом:

$$I_F = I [\exp(V_{эб}/V_T) - 1]$$

$$I_R = I [\exp(V_{кб}/V_T) - 1].$$

Закон Кирхгофа для точек Э и К (рис. 3.31) дает:

$$I_3 = I_{эбS} [\exp(V_{эб}/V_T) - 1] - \alpha_R I_{кбS} [\exp(V_{кб}/V_T) - 1] \quad (3.8.1)$$

$$I_к = -\alpha_F I_{эбS} [\exp(V_{эб}/V_T) - 1] - I_{кбS} [\exp(V_{кб}/V_T) - 1]. \quad (3.8.2)$$

Кроме того, сумма токов равна нулю:

$$I_б + I_3 + I_к = 0. \quad (3.8.3)$$

Подставляя второй член правой части уравнения (3.8.2) в (3.8.1), получаем:

$$I_3 = I_{эбS} [\exp(V_{эб}/V_T) - 1] \cdot (1 - \alpha_R \alpha_F) - \alpha_R I_к. \quad (3.8.4)$$

уравнение (3.8.4) можно переписать в виде:

$$I_{\bar{\sigma}} = I_{\bar{\sigma}00} [\exp(V_{\bar{\sigma}} / V_T) - 1] - \alpha_R I_K, \quad (3.8.5)$$

где $I_{\bar{\sigma}00} = I_{\bar{\sigma}0S} \cdot (1 - \alpha_R \alpha_F)$.

Последнее равенство имеет место, поскольку при $I_K = 0$.

$$I_{\bar{\sigma}} = I_{\bar{\sigma}00} = I_F - \alpha_R I_R = I_F - \alpha_R (\alpha_F I_F) = I_F (1 - \alpha_R \alpha_F),$$

откуда получаем

$$I_{\bar{\sigma}00} = I_{\bar{\sigma}0S} \cdot (1 - \alpha_R \alpha_F).$$

В случае, когда $I_{\bar{\sigma}} = 0$, справедливо следующее равенство:

$$I_K = I_{\bar{\sigma}00} = I_R - \alpha_F I_F = I_R - \alpha_F (\alpha_R I_R) = I_R (1 - \alpha_R \alpha_F),$$

откуда получаем

$$I_{\bar{\sigma}00} = I_{\bar{\sigma}0S} \cdot (1 - \alpha_R \alpha_F).$$

Выразим I_K в выражении (3.8.5) через токи $I_{\bar{\sigma}}$ и $I_{\bar{\sigma}}$ согласно (3.8.3)

$$I_{\bar{\sigma}} = \alpha_R I_{\bar{\sigma}} + \alpha_R I_{\bar{\sigma}} + I_{\bar{\sigma}00} [\exp(V_{\bar{\sigma}} / V_T) - 1],$$

или

$$I_{\bar{\sigma}} (1 - \alpha_R) = \alpha_R I_{\bar{\sigma}} + I_{\bar{\sigma}00} [\exp(V_{\bar{\sigma}} / V_T) - 1],$$

откуда получаем

$$\exp(V_{\bar{\sigma}} / V_T) = [I_{\bar{\sigma}} (1 - \alpha_R) - \alpha_R I_{\bar{\sigma}} + I_{\bar{\sigma}00}] / I_{\bar{\sigma}00}.$$

Таким образом, выражение для $V_{\bar{\sigma}}$ имеет вид

$$V_{\bar{\sigma}} = V_T \ln \frac{I_{\bar{\sigma}} (1 - \alpha_R) - \alpha_R I_{\bar{\sigma}} + I_{\bar{\sigma}00}}{I_{\bar{\sigma}00}}. \quad (3.8.6)$$

Подставляя первый член правой части уравнения (3.8.1) в (3.8.2), находим ток коллектора

$$I_K = I_{\bar{\sigma}00} \left[\exp \left(\frac{V_{\bar{\sigma}}}{V_T} \right) - 1 \right] (1 - \alpha_F \alpha_R) - \alpha_F I_{\bar{\sigma}}. \quad (3.8.7)$$

Выражение (3.8.7) можно записать в виде

$$I_K = I_{\bar{\sigma}00} \left[\exp \left(\frac{V_{\bar{\sigma}}}{V_T} \right) - 1 \right] - \alpha_F I_{\bar{\sigma}}, \quad (3.8.7a)$$

где

$$I_{\bar{\sigma}00} = I_{\bar{\sigma}0S} (1 - \alpha_R \alpha_F),$$

поскольку при $I_{\bar{\sigma}} = 0$ имеем

$$I_K = I_{\bar{\sigma}00} = I_R - \alpha_F I_F = I_R - \alpha_F (\alpha_R I_R) = I_R (1 - \alpha_F \alpha_R), \quad (3.8.8)$$

$$I_{\bar{\sigma}00} = I_{\bar{\sigma}0S} (1 - \alpha_F \alpha_R).$$

Выражая снова I_K в уравнении (3.8.7) через $I_{\bar{\sigma}}$ и $I_{\bar{\sigma}}$ согласно (3.8.3), получаем

$$-I_{\bar{\sigma}} - I_{\bar{\sigma}} + \alpha_F I_{\bar{\sigma}} + I_{\bar{\sigma}00} = I_{\bar{\sigma}00} \exp \left(\frac{V_{\bar{\sigma}}}{V_T} \right).$$

Отсюда находим

$$\exp\left(\frac{V_{кб}}{V_T}\right) = \frac{I_{\beta}(a_F - 1) - I_{\bar{\beta}} + I_{кб0}}{I_{кб0}}.$$

Таким образом,

$$V_{кз} = V_T \ln \left[\frac{I_{\beta}(a_F - 1) - I_{\bar{\beta}} + I_{кб0} + I_{кб0}}{I_{кб0}} \right] \\ V_{кз} = V_{кб} - V_{\bar{\beta}} = V_T \ln \left\{ \frac{[I_{\beta}(a_F - 1) - I_{\bar{\beta}} + I_{кб0} + I_{кб0}] / [I_{\beta}(1 - \alpha_R) - \alpha_R I_{\bar{\beta}} + I_{кб0}]}{I_{кб0}} \right\}. \quad (3.8.9)$$

Предположим, что $a_F I_{кб0} = a_R I_{кб0}$ так что

$$\frac{I_{кб0}}{I_{кб0}} = \frac{I_{кб0}}{I_{кб0}} = \frac{a_R}{a_F} \\ I_{\beta} = -I_{\bar{\beta}} - I_{кб0}.$$

Используя эти соотношения, выражение (3.8.9) можно записать в виде

$$V_{кз} = V_T \ln \left[\frac{a_R I_K (1 - a_F) - a_F I_{\bar{\beta}} + I_{кб0}}{a_F I_K (a_R - 1) - I_{\bar{\beta}} + I_{кб0}} \right]. \quad (3.8.10)$$

Заметим, что в режиме нормального смещения переход "эмиттер—база" смещен в прямом направлении, а переход "коллектор—база" смещен в обратном направлении. Это означает, что $|V_{кб}| \gg V_T$, т. е. $\exp(V_{кб}/V_T) \rightarrow 0$. Поэтому уравнение (3.8.7а) можно привести к виду

$$I_K = -a_F I_{\beta} - I_{кб0}.$$

Отсюда мы видим, что ток I_K течет в направлении, противоположном указанному на рис. 3.29 току

$$I_K = a_F I_{\beta} + I_{кб0}.$$

Уравнение (3.8.7, а) определяет выходные (I_n / $V_{кб}$) характеристики р—п—р-транзистора, включенного по схеме с общей базой.

3.9. Транзистор в режиме усиления постоянного тока

Чем транзистор отличается от двух диодов, включенных навстречу друг другу?

Рассмотрите р—п—р-транзистор, включенный по схеме, показанной на рис. 3.32. Положим, площадь эмиттера равна 0.01 см^2 , концентрация дырок в базе 10^{12} см^{-3} , а коэффициент диффузии $50 \text{ см}^2/\text{с}$.

1. Напряжение, измеренное вольтметром, равно 25 мВ. Какое напряжение покажет вольтметр, если коллектор накоротко замкнуть с базой?
2. Оцените эффективную толщину базы.
3. Пусть источником постоянного тока служит батарея напряжением 25 мВ. Вычислите ток эмиттера, если коллектор накоротко замкнут с базой.

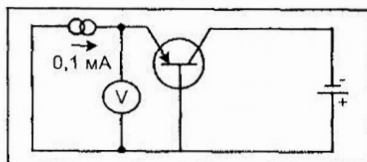


Рис. 3.32. Эквивалентная схема транзистора

решение

Транзистор отличается от двух диодов, включенных навстречу друг другу, тем, что в последнем соединении отсутствуют генераторы тока, которые существенным образом характеризуют работу транзистора, обеспечивая усиление тока.

Из эквивалентной схемы транзистора (рис. 3.31) на основе идеальных диодов Эберса — Молла ток эмиттера можно записать в виде:

$$I_e = I_{\text{эбс}} [\exp(V_{\text{эб}} / V_T) - 1] - \alpha_R I_{\text{кбс}} [\exp(V_{\text{кб}} / V_T) - 1], \quad (3.9.1)$$

где $I_{\text{эбс}}$ — ток насыщения эмиттерного перехода с коллектором, замкнутым на базу, $I_{\text{кбс}}$ — ток насыщения коллекторного перехода с эмиттером, замкнутым на базу, α_R — коэффициент усиления по току в режиме короткого замыкания в обратном направлении в схеме с общей базой.

Переход "коллектор—база" обычно смещен в обратном направлении на несколько вольт, поэтому $V_{\text{кб}} \gg V_T$, где V_T — термо-ЭДС, взятая равной 25 мВ.

Экспоненциальный член в уравнении (3.9.1) $\exp(V_{\text{кб}} / V_T) = \exp(-a) \rightarrow 0$, где a — большая величина.

Поэтому уравнение (3.9.1) можно записать в виде:

$$I_e \approx I_{\text{эбс}} [\exp(V_{\text{эб}} / V_T) - 1] + \alpha_R I_{\text{кбс}}. \quad (3.9.2)$$

Ток коллектора выражается аналогичным образом

$$I_k \approx -\alpha_F I_{\text{эбс}} [\exp(V_{\text{эб}} / V_T) - 1] - I_{\text{кбс}} \quad (3.9.3)$$

где α_F — коэффициент усиления по току в режиме короткого замыкания для прямого направления.

Согласно условиям задачи $V_{\text{эб}} = 25$ мВ, $V_T = 25$ мВ. Уравнение (3.9.2) можно записать в виде:

$$I_e \approx I_{\text{эбс}} [\exp(1) - 1] + \alpha_R I_{\text{кбс}} \approx I_{\text{эбс}} (2,7183 - 1) + \alpha_R I_{\text{кбс}} \approx I_{\text{эбс}} 1,7183 + \alpha_R I_{\text{кбс}}. \quad (3.9.4)$$

Если коллектор замкнут с базой, т. е. $V_{\text{кб}} = 0$, то

$$I_e = I_{\text{эбс}} [\exp(V_{\text{эб}} / V_T) - 1]. \quad (3.9.5)$$

Используя в этом выражении равенство

$$\alpha_F \cdot I_{\text{эбс}} = \alpha_R I_{\text{кбс}}$$

получим:

$$I_e = I_{\text{эбс}} (1,7183 + \alpha_F). \quad (3.9.6)$$

Приравнявая уравнения (3.9.5) и (3.9.6), т. е.

$$1,7183 + \alpha_F = \exp(V_{\text{эб}} / V_T) - 1$$

находим

$$V_{\text{эб}} = V_T \ln(2,7183 + \alpha_F).$$

Если предположить, что $\alpha_F = 0,98$, то

$$V_{\text{эб}} = 2,5 \cdot 10^{-3} \cdot \ln(3,6983) \approx 32,6 \text{ мВ}.$$

Контрольные вопросы

1. Что такое униполярный транзистор?
2. Нарисуйте схему и ВАХ униполярного транзистора с индуцированным каналом.

3. Нарисуйте схему и ВАХ униполярного транзистора со встроенным каналом.
4. Что такое пороговое напряжение отсечки?
5. В чем заключается "эффект поля"?
6. Какие малосигнальные параметры МОП-транзистора вы знаете?
7. Что такое КМОП-транзисторы?
8. Нарисуйте схему арсенид-галлиевого ПТШ-транзистора. Объясните принцип его работы.
9. Нарисуйте схему транзистора на гетероструктурах. Как он работает?
10. Какие конструкции транзисторов с вертикальным токопереносом вы знаете?
11. Что такое биполярный транзистор?
12. Нарисуйте схему биполярного эпитаксиально-планарного транзистора. Как он работает?
13. Какие режимы работы биполярного транзистора вы знаете?
14. Что такое многомиттерный транзистор?
15. Что такое многоколлекторный транзистор?
16. Нарисуйте схему транзистора с инжектором (I^2L).
17. Какую роль выполняет диод Шоттки в транзисторе? Как он выполнен конструктивно?
18. Нарисуйте схему транзистора с проницаемой базой.

Рекомендуемая литература

1. Авасян Н. А., Наумов Ю. Е., Фролкин В. Т. Основы микроэлектроники: Учебное пособие для вузов. — М.: Радио и связь, 1991.
2. Ефимов И. Е., Козырь И. Я., Горбунов Ю. Я. Микроэлектроника. Физические и технологические основы, надежность: Учебное пособие. — 2-е изд. — М.: Высшая школа, 1986.
3. Зи С. Физика полупроводниковых приборов: В 2-х книгах. — 2-е изд. — М.: Мир, 1984.
4. Колесов Л. А. Технология и конструкции микросхем, микропроцессоров и микросборок. Учебник для вузов. — М.: Радио и связь, 1989.
5. Линч П., Николайдес А. Задачи по физической электронике. Под ред. проф. Г. В. Скромного. — М.: Мир, 1975.
6. Росадо Л. Физическая электроника и микроэлектроника. Под ред. проф. В. А. Терехова. — М.: Высшая школа, 1991.
7. Степаненко И. П. Основы микроэлектроники: Учебное пособие для вузов. — 2-е изд. — М.: Лаборатория базовых знаний, 2000.
8. Терехов В. А. Задачник по электронным приборам: — 2-е изд. — М.: Энергоатомиздат, 1983.

4. Элементная база интегральных схем

4.1. Изоляция элементов

Обработка информации в дискретной электронике, использующей традиционную дискретную элементную базу, осуществлялась схемами различного функционального назначения. В микроэлектронике также исповедуются принципы схемотехнической обработки информации. В качестве элементов схемы используются *активные* элементы — транзисторы и *пассивные* элементы — диоды, резисторы, конденсаторы.

Под *элементами интегральных схем* (ИС) понимают неделимые и составные части ИС, которые нельзя автономно специфицировать и поставлять их как отдельные изделия.

Отличительной технологической особенностью элементов ИС по сравнению с дискретными приборами или электрорадиоэлементами является то, что они органически связаны общей полупроводниковой подложкой и друг с другом. Другой особенностью является то, что транзисторные структуры и пассивные элементы ИС производятся в едином технологическом процессе.

При этом номинал элемента (емкость конденсатора, сопротивление резистора) определяется как геометрическими размерами (топологией), так и заданными электрофизическими свойствами используемых материалов (толщиной диэлектрика, проводимостью полупроводника и т. д.). С другой стороны, геометрические параметры изготавливаемых элементов оптимизируются в соответствии с основной тенденцией микроэлектроники — ростом степени интеграции и уменьшением топологических норм.

В этой связи надо отметить факт создания принципиально новых элементов, ранее неизвестных в дискретной электронике: многоэмиттерные и многоколлекторные транзисторы, диоды Шоттки, встраиваемые в транзисторную структуру, и т. п.

В гибридных интегральных схемах, которым в этом курсе будет уделено минимальное внимание, используются также *компоненты*. Компоненты, в отличие от элементов ИС, выполняют те же функции, что и элементы, но их можно отдельно сертифицировать и поставлять в виде отдельных изделий.

Компоненты гибридных ИС представляют собой, как правило, навесные детали. Их отличие от дискретных элементов заключается в конструктивном решении (бескорпусные диоды, транзисторы, сборки и т. д.). И еще одна особенность пассивных элементов полупроводниковых (монокристаллических) ИС: в них отсутствуют аналоги индуктивностей, дросселей, трансформаторов. Если же встает острая необходимость использования индуктивных элементов, то индуктивный эффект реализуют схемным путем, используя операционные усилители с RC-цепями обратной связи, активные фильтры и т. д.

В гибридных СВЧ-микросхемах широко используют микрополосковые линии и их элементы. При острой необходимости применения резисторов или конденсаторов больших номиналов используются дискретные миниатюрные элементы, подключение которых осуществляется через специальные выводы интегральных схем.

Транзисторные структуры и элементы интегральных схем, расположенных на одной подложке, необходимо изолировать друг от друга, а соединение осуществлять в соответствии с принципиальной схемой путем металлической развязки или путем использования подложки.

На рис. 4.1 приведены три способа изоляции транзисторных биполярных структур.

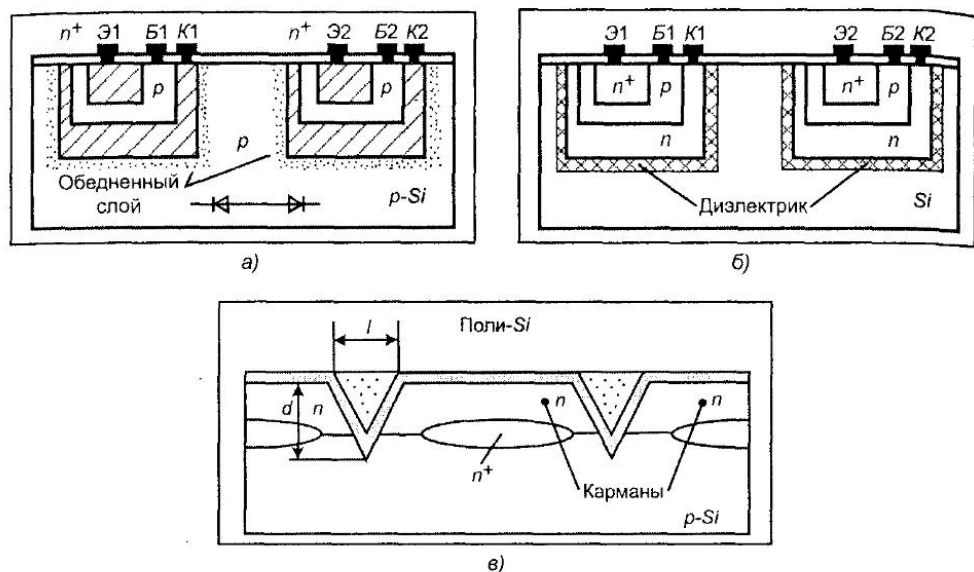


Рис. 4.1. Изоляция р—п-переходом (а); изоляция диэлектриком (б); комбинированная изоляция (в)

Метод изоляции обратносмещенным р—п-переходом базируется на свойстве такого перехода иметь очень высокое удельное сопротивление при обратном смещении. Изоляция р—п-переходом является однофазным способом потому, что материал по обе стороны и в пределах изолирующего слоя один и тот же.

Изоляция р—п-переходом по существу сводится к формированию двух встречно включенных диода между изолируемыми элементами (рис. 4.1, а).

Для того чтобы изолирующие диоды находились под обратным смещением, на подложку подают максимальный, отрицательный потенциал от источника питания.

Изоляция р—п-переходом органически вписывается в основной технологический цикл производства кремниевых интегральных схем. Используют изолирующую диффузию, методы тройной диффузии, встречной диффузии. К недостаткам этого способа изоляции следует отнести наличие обратных токов в р—п-переходах и наличие барьерных емкостей.

Метод изоляции диэлектриком сводится к созданию кармана из диэлектрика, в котором располагается транзисторная структура. Это более совершенный, чем предыдущий метод, прежде всего из-за чрезвычайно малых токов утечки, которые на 3—5 порядков меньше обратных токов в р—п-переходах.

увеличивая толщину диэлектрика и выбирая материал с малой диэлектрической проницаемостью, можно снизить и значения паразитных емкостей. На рис. 4.1, б показан один из способов изоляции диэлектриком транзисторных структур. Он получил названия *КВД — кремний в диэлектрике*. Одним из технологических процессов полной диэлектрической изоляции является эпик-процесс, обеспечивающий изоляцию элементов оксидным слоем SiO_2 .

Наибольшее распространение получили процессы, связанные с созданием транзисторных структур на диэлектрической подложке — *КНД — кремний на диэлектрике*. В качестве диэлектрической подложки часто используют сапфир, и такой способ изоляции получил название *КНС — кремний на сапфире*. На сапфировой подложке выращивается эпитаксиальный слой кремния, в котором методом прецизионного травления формируются кремниевые карманы. Карманы снизу изолированы сапфиром, сбоку и сверху — воздухом. В изолированных карманах и размещаются транзисторные структуры, которые затем коммутируются пленочной металлической разводкой.

Изоляцию диэлектриком относят к двухфазному способу потому, что используются одновременно две фазы — диэлектрик и полупроводник.

К недостаткам этого способа изоляции следует отнести необходимость совмещения нескольких разнородных технологических процессов.

Комбинированный метод, при котором сочетаются изоляция диэлектриком и изоляция р—п-переходом, является самым распространенным методом изоляции транзисторных структур.

Основным технологическим процессом является изопланарная технология, в основе которой лежит локальное окисление тонкого эпитаксиального слоя кремния. Результатом этого является образование карманов, которые сбоку изолированы диэлектриком, а от подложек изолируется р—п-переходом. В таких карманах и располагаются транзисторные структуры, а также элементы интегральных схем.

В изопланарном процессе для локального прокисления используются маски из нитрида кремния. Этот технологический процесс позволяет обеспечить большую плотность упаковки элементов на кристалле и получить высокие частотные и переходные характеристики транзисторных структур.

Большое распространение получил *метод боковой диэлектрической изоляции V-канавками*. В этом технологическом процессе вместо сквозного прокисления эпитаксиального слоя используется локальное анизотропное травление поверхности кристалла, ориентированной по плоскости (100).

В этом случае травление идет в плоскости (111) так, что грани (111) сходятся ниже границы эпитаксиального слоя. Образовавшиеся V-образные канавки заполняются диоксидом кремния, либо поликристаллическим кремнием (рис. 4.1, в).

Используя метод реактивного ионного травления, можно уменьшить ширину канавки и превратить ее из V- в U-образную.

Недостатком такого способа изоляции является использование плоскости (100), что сопряжено с повышенной плотностью поверхностных дефектов.

К изоляции МДП-транзисторных структур и элементам интегральных схем требования менее жестки в силу физических особенностей их работы. Эти же методы изоляции используются и в униполярных интегральных схемах.

4.2. Интегральные диоды

Диоды в интегральных микросхемах предназначены для выполнения ряда логических функций переключения электрических сигналов, выпрямления электрического тока, детектирования сигналов.

Любой из р—п-переходов транзисторной структуры, а также их комбинация могут быть использованы в качестве интегрального диода.

Эквивалентные схемы включения транзисторных структур в качестве диодов содержат собственную емкость диода и паразитные емкости, которые оказывают существенное влияние на характеристики диодов.

Пробивные напряжения диодов зависят от типа используемого перехода. Если применяется небольшой эмиттерный переход с сильно легированной областью эмиттера, то пробивные напряжения небольшие. Напротив, при использовании протяженного, слаболегированного коллекторного перехода пробивные напряжения достаточно велики.

Обратные токи $I_{обр}$ являются по существу токами термогенерации, зависящие от объема р—п-перехода. Поэтому они имеют большие значения у диодов, в которых используется большой коллекторный переход (табл. 4.1).

Таблица 4.1

Параметры	Тип диода				
	Б — Э	Б — К	БЭ — К	БК — Э	Б — ЭК
$U_{пр}$, В	7—8	40—50	40—50	7—8	7—8
$I_{обр}$, нА	0,5—1	15—30	15—30	0,5—1	20—40
C_0 , пФ	0,5	0,7	0,7	0,5	1,2
C_0 , пФ	1,2	3	3	3	3
t_0 , нс	50	75	50	10	100

Время восстановления обратного тока t_0 определяет время переключения диода в открытое или закрытое состояния.

Сравнительный анализ параметров биполярных интегральных диодов показывает, что в зависимости от функционального назначения диода можно выбрать нужную структуру.

В целом оптимальным вариантом для интегральных схем являются структуры типа БК—Э на основе перехода "база—эмиттер" с закороченным на базу коллектором и тип Б—Э на основе перехода "база—эмиттер" с разомкнутой цепью коллектора.

В интегральных схемах используются *интегральные стабилитроны*, представляющие собой полупроводниковый диод с быстрым нарастанием обратного тока при пробое р—п-перехода и нормированным значением пробивного напряжения. Они предназначены для стабилизации напряжения на нагрузке (рис. 4.2).

Интегральные стабилитроны формируются на базе структуры биполярного транзистора в зависимости от необходимого напряжения. Так обратное включение перехода "база—эмиттер" позволяет получить стабилизированное напряжение в пределах 5—10 В, обратное включение перехода БЭ—К применяют, когда нужно получить стабилизированное напряжение 3—5 В.

Температурная чувствительность таких стабилитронов лежит в пределах нескольких милливольт на градус.

В интегральных схемах используются также диоды Шоттки, представляющие собой контакт металла с кремнием, легированный донорной примесью ($< 10^{17} \text{ см}^{-3}$).

На рис. 4.3 приведены конструктивные решения планарных диодов Шоттки:

- конструкция с охранным кольцом из p^+ -области кремния позволяет исключить сильные электрические поля на краях (а);
- диод Шоттки с расширенным электродом позволяет избежать пробоя (б);
- конструкция с выпрямляющими и омическими контактами (в).

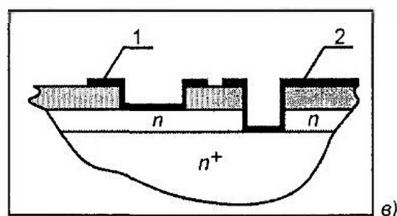
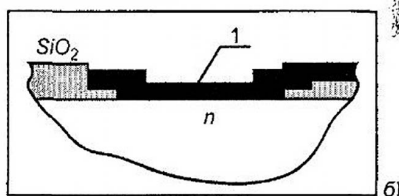
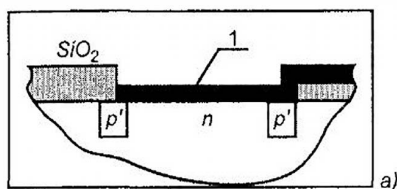


Рис. 4.3. Конструктивные решения планарных диодов Шоттки:
1 — металл, образующий барьер Шоттки;
2 — металл, образующий омический контакт

В качестве материала чаще всего используют алюминий. Для качественных диодов Шоттки в качестве материала используют сплав платины и никеля $\text{Ni}_x\text{Pt}_{1-x}$, образующий с кремнием силицидный слой. Меняя значения x , можно получить высоту барьеров от 0,64 эВ при $x = 0$ (или 100% Ni) до 0,84 эВ при $x = 1$ (или 100% Pt).

Интегральные МДП-транзисторные диоды формируются также на базе p — n -переходов транзисторов с индуцированным каналом в подложках разного типа электропроводности (рис. 4.4).

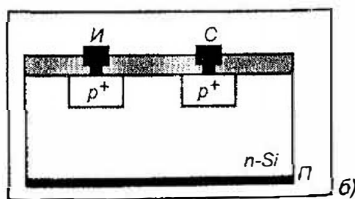
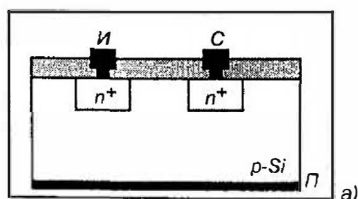


Рис. 4.4. Диоды в МДП-транзисторных структурах формируются типа И — П и С — П ($n^+ - p$) в p -кремневой подложке (а) и И — П и С — П ($p^+ - n$) в n -кремневых подложках (б)

Вольтамперные характеристики интегральных МДП-транзисторов аналогичны ВАХ интегральных биполярных транзисторов.

4.3. Интегральные резисторы

Интегральный резистор представляет собой элемент интегральных схем с заданным электросопротивлением и топологией, который используется в электрических цепях для обеспечения требуемого распределения тока и напряжений между отдельными участками цепи.

В интегральных схемах роль резисторов играют участки легированного полупроводника одной из областей транзисторной структуры. В гибридных интегральных схемах используются металлические пленки и пасты. Резисторы выполняются в одном технологическом процессе вместе с интегральными транзисторами и диодами.

Интегральные резисторы на биполярных структурах подразделяются на *диффузионные резисторы, пинч-резисторы, ионно-легированные резисторы, пленочные резисторы на основе поликристаллического кремния* (рис. 4.5).

Обычно тело резистора отождествляется с полоской длиной l , шириной b и толщиной d . Если ток протекает вдоль полоски параллельно ее плоскости с удельным сопротивлением материала ρ , то его сопротивление

$$R = \rho l / bd = k_f R_s, \quad (4.1)$$

где k_f — коэффициент формы, равный отношению длины полоски к ее ширине, $R_s = \rho / d$ — удельное сопротивление слоя.

При коэффициенте формы $k_f \leq 1$ резисторы изготавливаются в виде полоски. Если необходимы большие номиналы, то резистор выполняют в виде зигзагообразной конструкции. Максимальное сопротивление диффузионных резисторов не превышает 60 кОм.

Таким образом, номинальное значение резистора может быть получено выбором топологических параметров, коэффициентом формы или отношением длины l тела резистора к его ширине b , а также технологическими параметрами — выбором материала резистора и его толщины.

Диффузионные резисторы изготавливаются в эпитаксиальном слое транзисторной структуры. В зависимости от требуемого номинала и точности изготовления диффузионные резисторы могут изготавливаться в эмиттерной, базовой или коллекторной областях.

Чаще всего диффузионный резистор формируют в базовой области транзисторной биполярной структуры. Выбор этого слоя является компромиссом между большими геометрическими размерами, которые потребовались бы при изготовлении в эмиттерной области, и высоким температурным коэффициентом сопротивления резистора, если бы резистор выполнялся в слаболегированной коллекторной области.

В табл. 4.2 приведены параметры удельного поверхностного сопротивления в различных диффузионных областях и характеристики реальных резисторов.

Если необходимые номиналы превышают 60 кОм, используют конструкцию *пинч-резистора* (рис. 4.5, б). Большое удельное сопротивление достигается за счет использования донной части слаболегированной р-области.

Максимальное сопротивление пинч-резистора может достигать значения 200—300 кОм при простой полосковой конфигурации. Недостатком пинч-резисторов является большой разброс параметров изготавливаемых структур, а также большой температурный коэффициент сопротивления. Структура пинч-резистора сходна со структурой полевого транзистора, и именно этот факт позволяет получить большие значения сопротивления.

Таблица 4.2

Тип диффузионного слоя	Удельное поверхностное сопротивление, Ом/см ²	Разброс номинальных значений сопротивлений, %	Температурный коэффициент сопротивления резисторов, град ⁻¹
Эмиттерный	2—3	± 20	$(1-5) \times 10^{-4}$
Базовый	100—300	± (5—20)	$(1.5-3) \times 10^{-3}$
Базовый ограниченный эмиттерным (пинч-резистор)	$(5-10) \times 10^{-3}$	± (30—50)	$(3-6) \times 10^{-3}$
Коллекторный на эпитаксиальном слое	5×10^3	± 30	$(5-7) \times 10^{-3}$

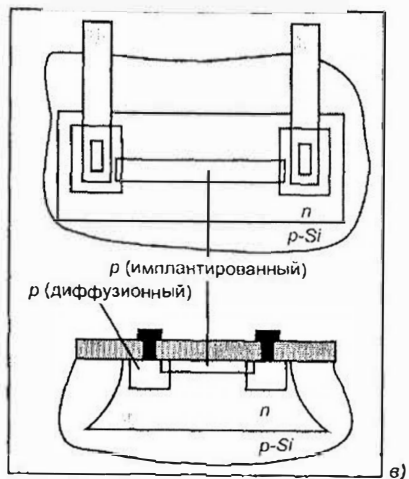
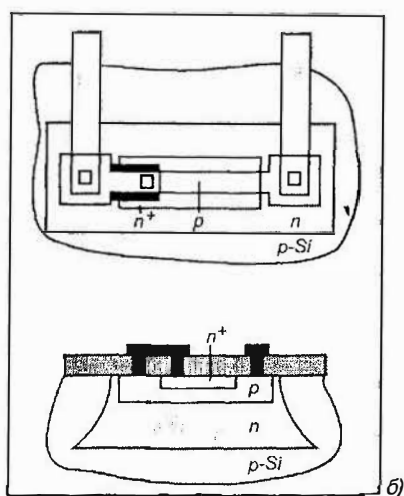
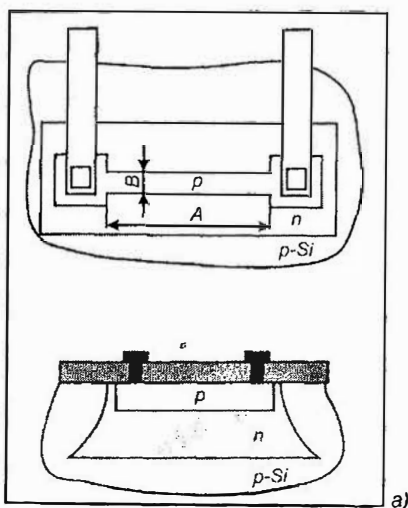


Рис. 4.5. Интегральные резисторы на основе биполярных транзисторных структур:
 а — диффузионный резистор; б — пинч-резистор;
 в — ионно-легированный резистор

Эмиттерная сильно легированная низкоомная область позволяет получить сопротивления в несколько Ом с температурным коэффициентом $0,01—0,02\%/^{\circ}\text{C}$.

Высокие удельные сопротивления могут быть обеспечены конструкцией *ионно-легированных резисторов*. Их структура аналогична диффузионным резисторам. Глубина залегания легированного и резистивного слоев составляет $0,2—0,3$ мкм. Поскольку толщина имплантированного слоя мала и к резистивному слою трудно приладить омические контакты, формируют диффузионные слои, осуществляющие омический контакт (рис. 4.5, в).

Тонкопленочные резисторы применяются в полупроводниковых биполярных интегральных схемах в основном СВЧ-диапазона, а также в схемах на арсениде галлия. Резистивный слой наносится непосредственно на поверхность нелегированной подложки.

В кремниевых цифровых БИС используются резистивные слои поликристаллического кремния толщиной $0,2 \div 0,3$ мкм. Такие резисторы размещаются над транзисторами, чтобы уменьшить площадь кристалла.

Интегральные резисторы МДП-транзисторных структур представляют собой, как правило, встроенные между истоком и стоком каналы (рис. 4.6).

Номиналы резистора определяются как топологией резистивных структур, так и технологией его изготовления. Обычно канал встраивается методом ионной имплантации, по своим свойствам аналогичен ранее рассмотренному ионно-легированному резистору.

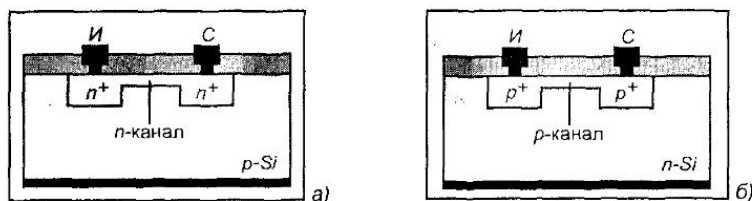


Рис. 4.6. Структура интегрального МДП-резистора на основе И — канал — С в p- (а) и n-подложках (б)

4.4. Интегральные конденсаторы

Интегральные конденсаторы представляют собой элементы интегральных схем, состоящие из проводящих электродов (обкладок), разделенных диэлектриком и предназначенные для использования в электрических цепях для обеспечения требуемого распределения тока и напряжения между отдельными элементами цепи.

В интегральных схемах роль конденсаторов играют обратносмещенные p—n-переходы, выполненные на основе транзисторной структуры в едином технологическом процессе.

В биполярных транзисторных структурах в конструкции интегрального транзистора используется один из переходов: "эмиттер—база", "база—коллектор", "коллектор—подложка". Эти переходы формируются диффузией и поэтому часто называются *диффузионными конденсаторами*.

Емкость конденсатора определяется емкостью перехода, имеющего диффузионную и барьерную составляющие. Основную роль играет барьерная емкость. Барьерная емкость связана с образованием области объемного заряда и потенциального барьера между p- и

р-областями перехода. Область объемного заряда р—п-перехода можно интерпретировать как аналог диэлектрика обычного конденсатора, если считать, что в нем отсутствуют подвижные носители зарядов.

Ширина этой области и плотность объемных зарядов неподвижных ионов донорных и акцепторных примесей зависят от напряжения обратного смещения, а также от контак-

ного потенциала $\varphi_k \geq 10 \frac{kT}{q}$.

С повышением напряжения обратного смещения ширина области объемного заряда увеличивается, что приводит к уменьшению барьерной емкости.

Величина барьерной емкости может быть определена из соотношения:

$$C = \frac{\varepsilon \varepsilon_0 S}{x_n - x_p}, \quad (4.2)$$

где ε — относительная диэлектрическая проницаемость, ε_0 — абсолютная диэлектрическая проницаемость, S — площадь р—п-перехода, x_n, x_p — границы области объемного заряда в материалах п- и р-типов.

Величина $x_n - x_p$ для р—п-перехода со ступенчатым распределением концентрации примесных атомов может быть вычислена из соотношения:

$$x_n - x_p = \left[\frac{2\varepsilon \varepsilon_0}{q} \frac{N_a + N_q}{N_a N_q} (U_{обп} + \varphi_k) \right]^{\frac{1}{2}}.$$

Тогда удельную барьерную емкость перехода запишем в виде:

$$C_0 = \left[\frac{\varepsilon \varepsilon_0 q}{2(U_{обп} + \varphi_k)(N_a - N_q)} \right]^{\frac{1}{2}}. \quad (4.3)$$

Если переход несимметричен и концентрация по одну сторону перехода много больше, чем по другую, например, $N_a \gg N_q$, то

$$C_0 = \sqrt{\frac{\varepsilon \varepsilon_0 q N_q}{2(U_{обп} + \varphi_k)}}. \quad (4.4)$$

В общем виде при любом распределении концентрации примесей можно записать:

$$C_0 \sim K \left(\frac{1}{U} \right)^m, \quad (4.5)$$

где K — коэффициент пропорциональности, зависящий от закона распределения концентрации примесных атомов, в окрестности р—п-перехода, m — коэффициент, значения

которого расположены в интервале: $\frac{1}{2} \geq m \geq \frac{1}{3}$. Типичные значения удельной барьерной емкости для различных р—п-переходов транзисторной структуры приведены в табл. 4.3.

На рис. 4.7 представлены конструктивные решения интегральных конденсаторов, выполненных на основе биполярной структуры.

Таблица 4.3

$U_{обр}, В$	$C_0 = C_{об},$ пФ/мм ²	$C_0 = C_{бк},$ пФ/мм ²	$C_0 = C_{кп},$ (без n^+ -слоя), пФ/мм ²	$C_0 = C_{кп},$ (с n^+ -слоем), пФ/мм ²
0	1400	300	190	260
5	1000	120	60	90
15	—	90	40	55

Наибольшую удельную барьерную емкость C_0 имеет переход "эмиттер—база" (рис. 4.7, а). Низкое пробивное напряжение этого перехода ограничивает возможность его широкого применения.

Конденсатор, сформированный на базе перехода "база—коллектор" (рис. 4.7, б) имеет более высокое пробивное напряжение ($\sim 50 В$). Обычно используют переход "база—коллектор" при его низком обратном смещении, но одновременно высоком обратном смещении перехода "коллектор—подложка". В этом случае соотношение $C_{бк} / C_{кп}$ имеет большее значение.

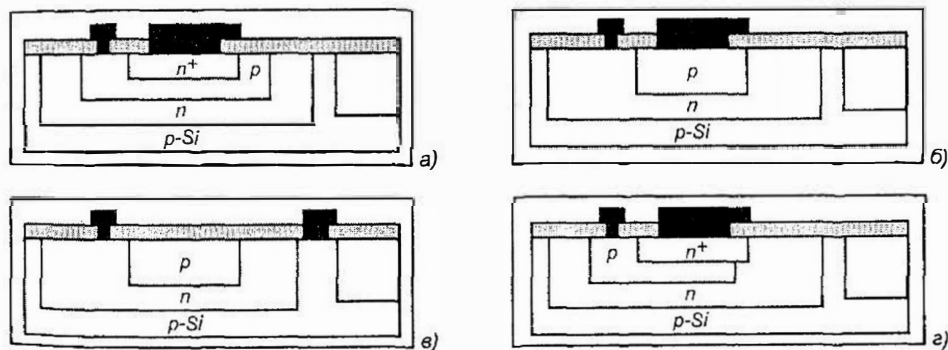


Рис. 4.7. Структуры интегральных биполярных конденсаторов: а — на основе эмиттерного перехода; б — на основе базового перехода; в — на основе перехода "коллектор—подложка"; г — на основе параллельно включенных емкостей эмиттерного и коллекторного переходов

Конденсаторы, сформированные на основе перехода "коллектор—подложка" (рис. 4.7, в) имеют ограниченное применение, потому что подложка обычно заземлена по переменной составляющей тока. Однако конденсатор этого типа является неотъемлемой частью интегральных схем с изоляцией по р—п-переходу.

Иногда в интегральных схемах используют *комбинированный конденсатор*. На рис. 4.7, г приведена структура конденсатора на основе параллельно включенных емкостей эмиттерного и коллекторного переходов.

Оптимальной конфигурацией конденсатора является квадрат. Общая емкость складывается из боковой и данной составляющих р—п-перехода. Обычно боковая составляющая в десятки раз меньше данной и ею пренебрегают.

При проектировании интегральных схем задаются условием, чтобы общая площадь всех конденсаторов интегральной схемы не превышала 25% площади кристалла.

Важным параметром интегрального конденсатора является *добротность* Q , которая определяется соотношением:

$$Q = \frac{1}{2\pi fCR}, \quad (4.6)$$

где f — рабочая частота, C — емкость конденсатора, R — сопротивление резистора, последовательно включенного с конденсатором.

Обычно в качестве резистора имеют в виду сопротивление диффузной области, непосредственно прилегающей к области объемного заряда перехода.

Добротность характеризует потерю мощности при протекании емкостного тока. Добротность возрастает с уменьшением частоты и с уменьшением сопротивления нижних слоев транзисторной структуры. Типичное значение добротности на частоте 500 Гц составляет 50—100.

Интегральные конденсаторы, сформированные на основе биполярной транзисторной структуры, имеют ряд недостатков.

Прежде всего, на основе таких структур невозможно создать конденсатор большой емкости. Для этого необходимо использовать большую площадь подложки. Конденсаторы такой конструкции имеют малую добротность, и их емкость существенно зависит от приложенного напряжения.

Недостатки диффузионных конденсаторов могут быть в схеме устранены, если воспользоваться конструкцией МДП-конденсатора (рис. 4.8). Над эмиттерным n^+ -слоем, служащим нижней обкладкой конденсатора, выращивается тонкий слой кремния SiO_2 , а затем наносится верхняя металлическая обкладка.

Важным достоинством МОП-конденсаторов является возможность их работы при любой полярности напряжения на обкладках.

Другой важной особенностью является независимость номинального значения емкости от приложенного напряжения.

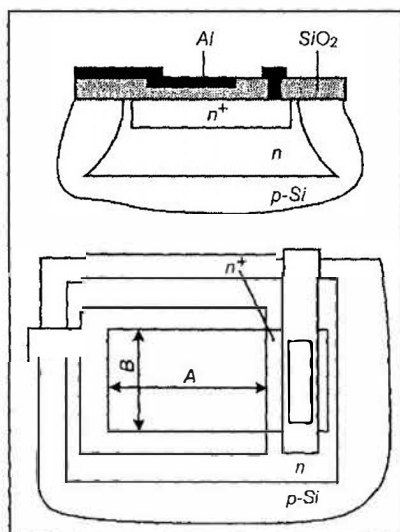


Рис. 4.8. Структура интегрального МДП-конденсатора и его топология:
 A — длина обкладки,
 B — ее ширина

И наконец, добротность МДП-конденсаторов значительно превосходит добротность интегральных конденсаторов, выполненных в биполярных транзисторных структурах.

В табл. 4.4 приведены типичные параметры интегральных конденсаторов: технологический разброс δ , температурный коэффициент емкости ТКЕ, пробивное напряжение $U_{пр}$ и добротность Q .

Таблица 4.4

Тип конденсатора	C_0 , пФ/мм ²	$C_{макс}$, пФ	δ , %	ТКЕ, %/градус	$U_{пр}$, В	Q (1 МГц)
Переход БК	150	300	± 20	$-0,1$	50	50—100
Переход БЭ	1000	1200	± 20	$-0,1$	7	1—20
МОП-структура	300	500	± 25	0,02	20	200

Индуктивные элементы в основном пленочные и применяются в гибридных интегральных СВЧ-схемах.

Задачи и упражнения

4.1. Расчет диффузионных резисторов

Рассчитать геометрические размеры диффузионных резисторов (рис. 4.9, а) с $R = 1,5$ кОм с контактными площадками, показанными на рис. 4.9, б).

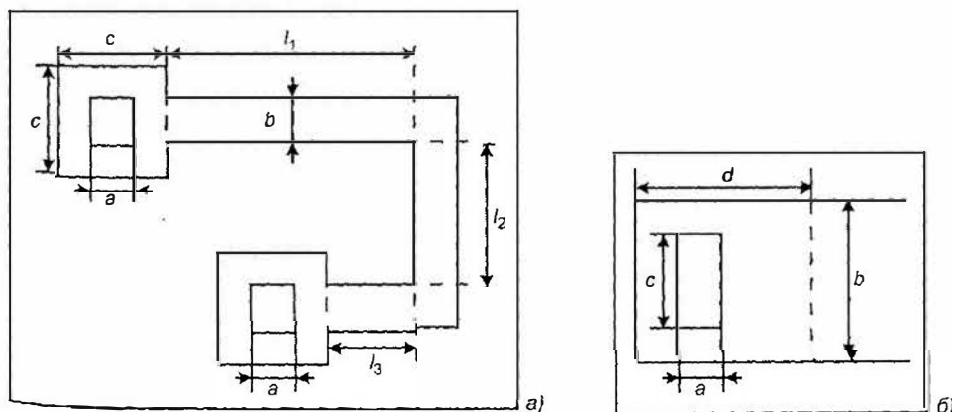


Рис. 4.9. Топология диффузионного резистора

Решение

Примем рассеиваемую мощность $P = 1,5$ мВт и разброс сопротивления 10% при $|\Delta R_s|/R_s \approx 7\%$. При найденных геометрических размерах определить разброс сопротивления вероятностным методом. Будем считать, что минимальная ширина окна для изготовления резистора $b_{min} = 4$ мкм; систематическая ошибка, связанная с растравлением

оксида $\Delta b_{\phi} = 0,75$ мкм; ошибка в линейных размерах при изготовлении фотошаблонов $\Delta b_{\phi} = \Delta b_{\tau} = \pm 0,4$ мкм; $R_y = 180$ Ом. Примем, что эффективная ширина $b_{эф} = b + 2\Delta b_{\phi}$, а $\Delta b_{\phi} \approx d = 1$ мкм.

Тогда минимальная эффективная ширина, определяемая технологическими ограничениями, будет

$$b_{эф0} = b_{\min} + 2(\Delta b_c + \Delta b_{\phi}) \approx 7,5 \text{ мкм.}$$

Будем считать, что $l \gg b$, тогда минимальную ширину, определяемую по рассеиваемой мощности, $b_{эфm}$ вычислим из соотношения:

$$b_{эфm} = 2\Delta b_{\phi} l \left(\frac{\Delta R}{R} - \frac{2\Delta b_{\tau}}{l} - \frac{\Delta R_y}{R_y} \right) \approx 27 \text{ мкм.}$$

С другой стороны, минимальная ширина, определяемая точностью, будет

$$b_{эфp} = \sqrt{R_y p / R p_0} \approx 6 \text{ мкм.}$$

Тогда

$$b_{эф} = \max \{ b_{эф0}; b_{эфp}; b_{эфm} \} = 27 \text{ мкм.}$$

Ширина окна вскрытия по фотошаблону

$$b = b_{эф} - 2(\Delta b_c + \Delta b_{\phi}) = 23,5 \text{ мкм.}$$

Проведем округление и примем $b = 24$ мкм. Из формулы для сопротивления диффузионного резистора зигзагообразной формы имеем:

$$\sum_{i=1}^{n+1} l_i = b_{эф} (R / R_0 - 0,55n - K_1 - K_2).$$

Примем $c = 16$ мкм; $a = 4$ мкм; $d = 12$ мкм (см. рис. 4.9, б).

Таким образом,

$$d_{эф} = d + 2(\Delta b_c + \Delta b_{\phi}) = 15,5 \text{ мкм;}$$

$$a_{эф} = a + 2\Delta b_c = 5,5 \text{ мкм.}$$

Значения

$$b_{эф} / d_{эф} = 1,8,$$

$$a_{эф} / d_{эф} = 0,35,$$

$$K_1 = K_2 = 0,21.$$

При этом суммарная длина линейных участков $\sum_{i=1}^3 l_i = 184$ мкм.

Разброс сопротивлений, полученных вероятностным методом при найденных значениях геометрических размеров резистора, позволяет записать

$$\sigma_r = \delta_r / 3,$$

где δ_r — относительный разброс какого-либо параметра. Тогда

$$\delta = \Delta R \cdot 100\% / R \approx 7,5\%.$$

Полученное значение, как и следовало ожидать, немного отличается от заданного (на 2,5%).

4.2. Расчет пинч-резистора

Рассчитать длину пинч-резистора l с $R \approx 200$ кОм при напряжении на клеммах, равном нулю.

Решение

Примем: $b = 20$ мкм; $\Delta b_c = 0,75$ мкм; напряжение $U_s = -2$ В. Будем считать, что $R_s = 17$ кОм; $d_k = 2,4$ мкм; $\omega = 0,3$ мкм; $\Delta d_{np} = 0,24$ мкм; $d_s = 1,7$ мкм.

Сопротивление резистора найдем из формулы:

$$R_{cp} = \frac{R_s l}{b}.$$

С учетом Δb_c получим $l = 253$ мкм.

4.3. Расчет конденсатора биполярной структуры

Рассчитать геометрические размеры квадратного конденсатора емкостью $C = 55$ пФ (при напряжении на переходе $U = 0$) с напряжением пробоя $U_{np} \geq 30$ В.

Решение

Для коэлекторного перехода, в котором будет изготовлен конденсатор, примем $d_k = 2,4$ мкм; $C_0 = 220$ пФ/мм²; $U_{np} \approx 40$ В. Площадь перехода $S_k = C / C_0 = 0,25$ мм². С погрешностью менее 5% величина $S_k = L^2$. Отсюда $L_1 \approx 500$ мкм.

4.4. Расчет МОП-конденсатора

Рассчитать удельную емкость $C_{0л}$; отношение $C'_{0min} / C_{0л}$; пороговое напряжение U_0 ; напряжение пробоя U_{np} МДП-конденсатора, структура которого приведена на рис. 4.8.

Решение

Примем, что диэлектриком является SiO₂, толщина которого $d_l = 10^{-5}$ см и диэлектрической проницаемостью $\epsilon_l = 3,9$. Поверхностную концентрацию донорной примеси N_{dl} примем равной 10^{19} см⁻³, контактная площадка выполнена из Al, кристаллографическая ориентация кремния $\langle 111 \rangle$. Считая, что d_n мало, положим $N_{dl} = N_{\delta l} = \text{const}$.

Из соотношения для области пространственного заряда толщиной

$$d_n \approx \sqrt{2\epsilon\epsilon_0\phi / qN_{dl}}$$

имеем

$$C_{0л} = 345 \text{ пФ/мм}^2.$$

Удельная емкость, определяемая диэлектриком

$$C_{0л} = \epsilon_l \epsilon_0 / d_l$$

и емкость идеального конденсатора определяется как

$$C_{0min} = \epsilon_l \epsilon_0 / [d_l + (\epsilon_l / \epsilon) d_{n \max}].$$

Следовательно,

$$C'_{0min} / C_0 = [1 + d_{n \max} \epsilon_l / (d_l \epsilon)]^{-1}.$$

Ширина слоя пространственного заряда достигает максимального значения

$$d_{n\max} = 2\sqrt{\epsilon\epsilon_0 |\Delta\varphi_F| / q N_A},$$

где

$$\Delta\varphi_F = -\varphi_F \ln \left(\frac{N_A}{n_i} \right).$$

Тогда $\Delta\varphi_F = -0,53$ В; $d_{n\max} = 1,7 \times 10^{-6}$ см, $C'_{0\min} / C_{0\lambda} = 0,96$.

Отсюда

$$\varphi_M = 4,3$$
 В

$$\varphi_{\text{мод}} = \varphi_M - \varphi_{n-g} - \frac{\varphi_d}{2} - \Delta\varphi_F = -0,08$$
 В,

где $\varphi_{n-g} = 4,35$. Тогда для ориентации кремния $\langle 111 \rangle$ плотность заряда поверхностных состояний равна $(Q_{n.s.} / q) \times 10^{-1} \text{ см}^{-2} = 5,0$ и $U_0 = -58$ В (при $Q_{n.s.} = 5 \times 10^{11} q$), $U_{np} = E_{кр} d_{\lambda} \approx 60$ В при $N_{\omega\lambda} \geq 10^{19} \text{ см}^{-3}$ конденсатор обладает высокой стабильностью ($C'_{0\min} / C_{0\lambda} = 0,96$).

Контрольные вопросы

1. Что такое элементы и компоненты интегральных схем?
2. Какие методы изоляции транзисторных структур вы знаете?
3. Расскажите о методе изоляции типа КИС.
4. Каковы функции диодов в интегральных схемах.
5. Какие схемы диодного включения транзистора вы знаете?
6. Дайте определение интегрального резистора.
7. Как связано сопротивление резистора с коэффициентом формы?
8. Что такое диффузионный резистор?
9. Что такое пинч-резистор?
10. Как устроен интегральный резистор МОП-структуры?
11. Дайте определение интегральному конденсатору.
12. Как рассчитать емкость интегрального конденсатора?
13. Дайте определение добротности интегрального конденсатора.

Рекомендуемая литература

1. Аваев Н. А., Наумов Ю. Е., Фролкин В. Т. Основы микроэлектроники: Учебное пособие для вузов. — М.: Радио и связь, 1991.
2. Березин А. С., Мочалкина О. Р. Технология и конструирование интегральных микросхем: Учебное пособие для вузов под ред. И. П. Степаненко. — М.: Радио и связь, 1983.
3. Ефимов И. Е., Козырь И. Я., Горбунов Ю. Я. Микроэлектроника. Физические и технологические основы, надежность: Учебное пособие. — 2-е изд. — М.: Высшая школа, 1986.
4. Колесов Л. А. Технология и конструкции микросхем, микропроцессоров и микросборок. Учебник для вузов. — М.: Радио и связь, 1989.
5. Степаненко И. П. Основы микроэлектроники: Учебное пособие для вузов. — 2-е изд. — М.: Лаборатория базовых знаний, 2000.

5. Технология производства интегральных схем

5.1. Технологические процессы изготовления ИС

Производство кремниевых интегральных схем базируется на технологических процессах формирования в кристалле кремния большого количества транзисторных структур, соединенных в схемы с заданными функциями.

В настоящее время значительный интерес представляют процессы производства монокристаллических полупроводниковых схем, как на кремнии, так и на арсениде галлия. Производство интегральных схем состоит из более чем нескольких сотен отдельных или повторяющихся технологических процессов и операций, являющихся предметом ноу-хау, которые защищены патентами.

Этапы технологических процессов производства интегральных схем условно можно разделить на четыре группы:

- ☐ процессы первичной обработки материалов, результатом которых являются отполированные кремниевые пластины;
- ☐ процессы формирования топологии интегральных схем, в результате которых формируются физические и геометрические параметры активных и пассивных элементов заданной схемы заданного функционального назначения;
- ☐ процессы локального изменения физических свойств подложки, в результате которых формируются технологические и топологические параметры элементов заданной схемы заданного функционального назначения;
- ☐ сборка микросистемных устройств и процессы межоперационного контроля.

В результате всех этих процессов кремниевая пластина трансформируется во множество отдельных интегральных схем, каждая из которых содержит миллионы транзисторных структур. Каждая интегральная схема выполняет отдельную функцию или набор функций.

Каждый этап технологического процесса основывается на ключевых физических или химических явлениях, имеющих целью локально преобразовать полупроводниковую структуру так, чтобы выполнялась заданная целевая функция, связанная с процессами обработки или хранения информации.

Кратко рассмотрим основные технологические операции и процессы, связанные с наиболее широко применяемой кремниевой технологией.

Арсенид-галлиевая и кремний-германиевая технологии используют в основном процессы, хорошо разработанные в кремниевой технологии.

5.2. Процессы первичной обработки материалов

Основным материалом для изготовления интегральных схем служит кремний — полупроводниковый материал серого цвета, один из наиболее распространенных в природе химических элементов.

Свойства кремния приведены в табл. 5.1.

Таблица 5.1

Атомный номер	14
Атомный вес	28,086
Плотность атомов	$4,96 \cdot 10^{22}$ ат/см ³
Плотность вещества	2,328 г/см ³
Диэлектрическая постоянная	$11,7 \pm 0,2$
Ширина запрещенной зоны	$1,115 \pm 0,008$ эВ
Температурный коэффициент ширины запрещенной зоны	$-2,3 \times 10^{-4}$ эВ/°C
Температура плавления	1417 ± 4 °C
Подвижность электронов	1350 ± 100 см ² /В·с
Подвижность дырок	480 ± 15 см ² /В·с
Коэффициент преломления	3,420
Теплопроводность	157 (Вт/см)/°C
Линейное тепловое расширение	$(0,26 \pm 0,3) \times 10^{-6}$ /°C
Постоянная кристаллической решетки	5,4307 Å
Коэффициент объемного сжатия	$0,98 \times 10^{-12}$ см ² /дин
Фотоэмиссионная работа выхода	$5,05 \pm 0,2$ эВ
Твердость	7,0 (по шкале Моса)
Удельная теплота плавления	100 Дж/г
Удельная теплота сублимации	$(18 \pm 2) \times 10^3$ Дж/г
Концентрация основных носителей	$1,5 \times 10^{16}$ см ⁻³
Валентность	4

Кремний, как материал микроэлектроники, на сегодняшний день является основным и в обозримом будущем не будет вытеснен по следующим причинам:

- ☐ уникальное сочетание ширины запрещенной зоны и других электронных свойств;
- ☐ стабильность и диэлектрические свойства окисла;
- ☐ технологичность в различных физико-химических процессах;
- ☐ большие природные запасы.

Первым этапом в изготовлении кремниевой пластины является очистка сырого кремния от примесей.

Обычно используется *зонная плавка* — метод перекристаллизации материалов посредством создания в образце небольшого расплавленного участка — зоны ее перемещения по всему образцу. На рис. 5.1 приведена схема зонной очистки кремния.

В установках зонной очистки широко применяется индукционный нагрев материала токами высокой частоты. В процессе плавки либо образец, либо нагреватель перемещаются

со скоростью 0,1—10 мм/мин. Зонная очистка позволяет получить чистые материалы с содержанием примесей 10^{-7} — $10^{-9}\%$.

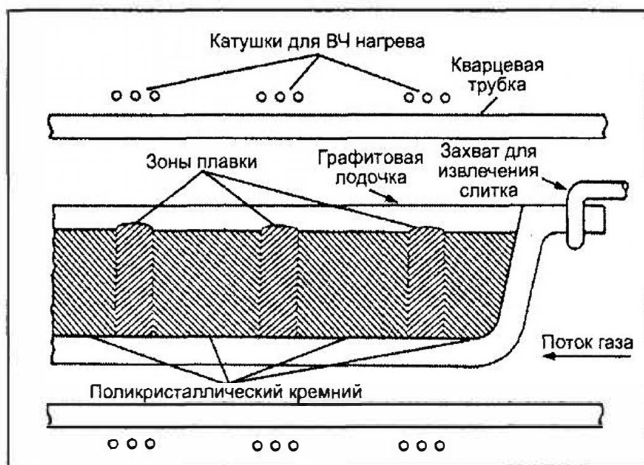


Рис. 5.1. Установка зонной очистки кремния

На втором этапе производится выращивание монокристалла кремния.

Слиток выращивается на основе монокристалла — *затравки*, помещенной на торце держателя. Кристалл-затравка помещается в расплав кремния и медленно вращаясь перемещается вверх.

Расплавление и последующее охлаждение расплава кремния стимулирует рост монокристалла в соответствии со структурой затравки кристалла.

Диаметр цилиндрического слитка кремния, получаемый в процессе роста по методу Чохральского, достигает 120, 150, 200 мм. Длина слитков зависит от непрерывности подачи исходного материала в вакууме и может достигать метровой длины и более (рис. 5.2).

Для получения высококачественных слитков используются автоматизированные системы, позволяющие задать оптимальную скорость вытягивания кристалла. Слитки шлифуются для получения плоской грани параллельной оси кристалла. Эта грань используется в качестве базовой на протяжении всего процесса изготовления интегральных схем при совмещении изображений и проведения электрических и оптических измерений.

Следующим этапом технологического процесса является разрезание слитка. С помощью алмазных дисков, вращающихся с большими скоростями, слиток расщепляется на отдельные пластины толщиной 0,5—1,0 мм. Поверхность подложки чаще всего ориентируют вдоль кристаллографической плоскости [111]. Отклонение плоскости пластины от кристаллографических плоскостей должно быть в пределах 20 секунд.

На следующем этапе механической обработки пластины полируют пастами различных классов. В итоге шероховатость поверхности кремниевых пластин составляет 0,5—0,05 мкм, что соответствует высшему 14 классу.

Толщина нарушенного абразивной обработкой слоя может достигать значения 1 мкм. Для подготовки пластин высокого класса поверхности используется *химическое травление* кремния. С его помощью можно получить поверхность оптической чистоты. Одновременно

менно с этим также удаляется нарушенный слой, оставшийся после механической обработки.

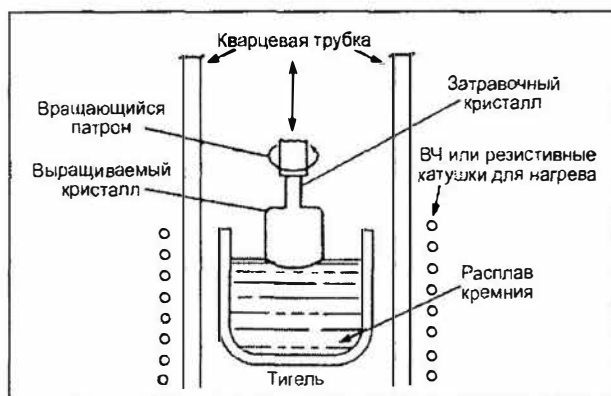


Рис. 5.2. Выращивание кристаллов по методу Чохральского

Монокристаллический кремний легируется и, соответственно, может быть электронного или дырочного типов проводимости.

В зависимости от легирующей примеси, каковой может быть алюминий, бор, сурьма, мышьяк, фосфор или золото, различают различные типы кремниевых пластин. Обозначения имеют три буквы и дробь, числители которой информируют об удельном объемном сопротивлении, а знаменатели — о диффузионной длине неравновесных носителей заряда (в миллиметрах). К примеру, КЕФ 2/0,2 — это кремний электронного типа проводимости, легированный фосфором, имеющий удельное сопротивление 2 Ом·см и диффузионную длину неравновесных носителей 0,2 мм.

Для восстановления разрушенного в процессе механической обработки приповерхностного слоя на пластине кремния выращивают тонкий (~ 10 мкм) эпитаксиальный слой.

5.3. Процессы литографии

Группа технологических процессов, предназначенная для формирования топологических параметров транзисторных структур, объединена понятием *литография*.

Литографические процессы формируют заданный рельеф (рисунок) в слое полупроводника, диэлектрика или металла с целью изготовления интегральных схем или других электронных приборов. Литографические процессы включают в себя следующие этапы:

- ☐ нанесение фоточувствительной полимерной пленки на кремниевую пластину;
- ☐ сушка и последующее экспонирование пластины через соответствующий фотошаблон;
- ☐ проявление изображения путем травления незащищенных полимерной пленкой областей со сформированным изображением.

Основной характеристикой литографического процесса является *разрешающая способность*, другими словами, способность разделять воспроизводить отдельные элементы топологии. Разрешающая способность определяется длиной волны воздействующего из-

лучения и параметрами резиста. Теоретический предел разрешающей способности, согласно критерию Релея, определяется величиной, пропорциональной половине длины волны экспонирующего излучения.

В зависимости от вида воздействующего излучения различают оптическую или фотолитографию, электронолитографию, рентгенолитографию, ионно-лучевую литографию, лазерную литографию. Приведем краткую их характеристику.

5.3.1. Фотолитография

Фотолитография представляет собой фотомеханический способ изготовления печатной формы плоской печати на заданной поверхности.

Целью фотолитографического процесса является перенос деталей рисунка фотошаблона на поверхность кремниевой пластины, покрытой слоем фоторезиста.

Под *фоторезистом* понимают резист, чувствительный к излучению в видимой или ультрафиолетовой области. Различают *позитивные* и *негативные* фоторезисты.

Фотошаблон представляет собой плоскопараллельную пластину из прозрачного материала (стекло), на которую нанесен рисунок в виде прозрачных и непрозрачных для излучения участков. Эти участки образуют топологию одного из слоев структуры. Фотошаблон может быть выполнен в виде позитивного или негативного изображения исходной топологии.

При облучении фоторезистов через шаблон и последующим проявлением удаляются облученные его участки, и на подложке формируется изображение фотошаблона. Фоторезист остается на участках подложки, соответствующих, например темным полям фотошаблона. Для отрицательных фоторезистов можно получить дополнительное (негативное) изображение. Далее в результате воздействия агрессивных веществ, например в процессе химического травления, материал, находящийся на участках, свободных от защитной маски фоторезиста, удаляется, что обеспечивает воспроизведение изображения элементов схем на подложке.

Фотолитографические процессы включают в себя множество операций, начиная от изготовления фотошаблонов и кончая формированием элементов схем на подложке.

Различают проекционную и контактную фотолитографии.

Метод *проекционной фотолитографии* заключается в проецировании фотошаблона на пластину кремния, покрытую фоторезистом. При этом между пластиной кремния и фотошаблоном имеется определенный воздушный зазор (рис. 5.3).

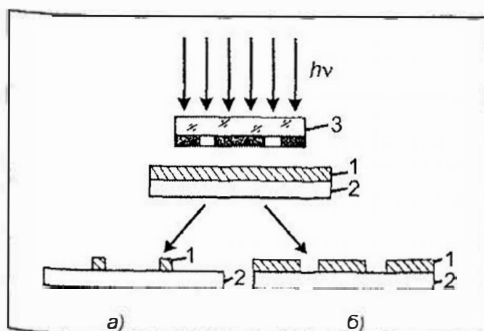


Рис. 5.3. Схема воспроизведения изображения методом проекционной фотолитографии:

а — с негативным фоторезистом;
б — с позитивным фоторезистом;
1 — фоторезист; 2 — подложка;
3 — фотошаблон

Метод проекционной литографии с сохранением масштаба позволяет обрабатывать пластины диаметром до 150 мм (рис. 5.4, кривая 2).

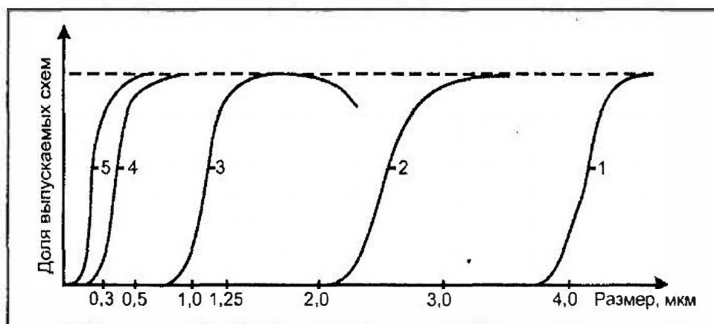


Рис. 5.4. Развитие методов литографии

Оптические характеристики таких систем ограничены и для повышения разрешающей способности используются источники излучения меньшей длины волны, например УФ-излучение.

Метод *контактной литографии* имеет две основные операции: совмещение изображений и экспонирование фоторезиста (рис. 5.5). При этом методе фотошаблон и пластина кремния с нанесенным фоторезистом плотно прижаты друг к другу. В методе контактной фотолитографии также используется излучение оптического диапазона. Этот метод применяют для получения рисунков с топологической нормой 2—3 мкм.

Для получения точной передачи размеров изображения фотошаблона необходимо одновременно и взаимосвязанно изменять как время экспонирования, так и время проявления рисунка.

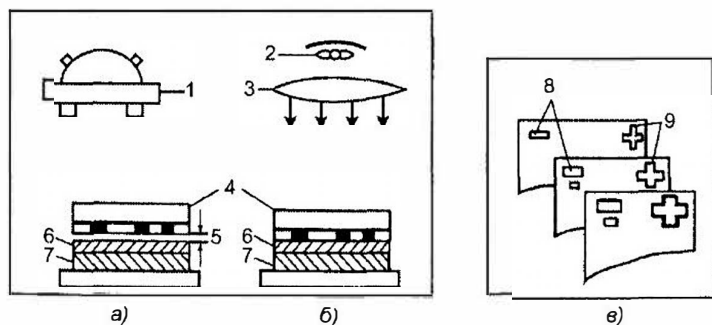


Рис. 5.5. Схема операции совмещения (а) и экспонирования (б), фигуры совмещения (в):
1 — микроскоп; 2 — источник света; 3 — объектив; 4 — фотошаблон; 5 — зазор; 6 — фоторезист;
7 — экспонируемая пластина; 8 — фигуры совмещения на пластине;
9 — фигуры совмещения на фотошаблоне

На установках с пошаговым репродуцированием и уменьшением размера, или *фото-мультипликаторах*, производится размножение рисунка шаблона по всей поверхности пластины (рис. 5.6).

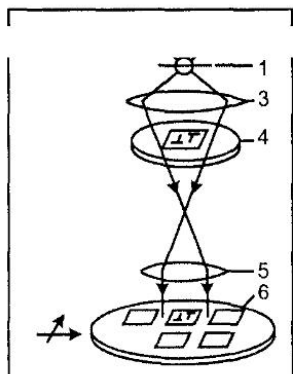


Рис. 5.6. Оптическая схема фотомультипликатора

Коротковолновое излучение от источника 1 формируется с помощью зеркала 2 и конденсора 3 на плоскость промежуточного фотошаблона 4. Рисунок промежуточного фотошаблона проецируется через уменьшающую линзу 5 на поверхность пластины 6, покрытой резистом.

После экспонирования стол автоматически перемещается в новое положение. Последующее экспонирование позволяет покрыть поверхность пластины изображениями промежуточного шаблона.

Проекционная установка с уменьшением масштаба характеризуется большей разрешающей способностью, чем установка с сохранением масштаба. Промежуточные фотошаблоны проще изготовить, чем шаблоны для установок проекционной литографии с сохранением масштаба (рис. 5.4, кривая 3). Для получения промежуточных фотошаблонов используется *фотоповторитель*, на котором происходит мультиплицирование изображения фотооригинала в уменьшенном масштабе.

Нынешние успехи фотолитографии обусловлены двумя причинами. Во-первых, непрерывным совершенствованием разрешающей способности и числовой апертуры оборудования и, во-вторых, использованием новых фоторезистов.

Среди других путей повышения эффективности фотолитографических процессов следует указать на использование *усиливающих контраст материалов* (УКМ). В таком технологическом процессе слой УКМ, не прозрачный в основном состоянии, наносят на верхний слой фоторезиста. При экспонировании он становится прозрачным и усиливает поглощение света фоторезистом. Неэкспонированные области остаются темными и с их помощью образуется оптическая маска. Использование УКМ существенно улучшает разрешающую способность фоторезиста. Применения УКМ в сочетании с источником ультрафиолетового диапазона позволяет получить на установках с пошаговым экспонированием (10:1) разрешающую способность лучше 0,35 мкм.

Дальнейшее развитие систем фотолитографии связано с использованием *эксимерных лазеров*. Такие системы отличаются высокой производительностью и субмикронные топологические нормы (см. рис. 5.4, кривая 4). Фотолитографические процессы включают в себя множество операций, начиная с изготовления фотошаблонов и кончая формированием элементов схем на подложке.

5.3.2. Электронолитография

Электронолитография представляет собой современный высокоразрешающий метод создания топологии микроструктур с использованием электронных пучков в качестве источника излучения.

Как технологический процесс больших потенциальных возможностей электронолитография получила признание во второй половине 60-х годов, а практические работы относятся к концу 70-х—началу 80-х годов (рис. 5.4, кривая 5).

В основе электронолитографии лежит явление нетермического воздействия электронных лучей на материал резиста, чувствительного к потоку электронов. Метод электронолитографии представляет собой комплекс технологических приемов, позволяющий создать технологический рисунок на пластине. Сделать это можно двумя путями. Первый путь заключается в последовательном экспонировании каждого элемента топологии. Этот метод получил название последовательной или *сканирующей электронной литографии*. Второй путь предусматривает передачу на пластину полного изображения шаблона с помощью электронных лучей. Этот метод называется *проекционной электронной литографией*.

В обоих методах используется экспонирование резиста пучком электронов с его последующим проявлением. При этом резист должен быть чувствителен к электронному потоку.

Известно, что длина волны движущегося электрона (в нм) определяется соотношением:

$$\lambda = 10 \sqrt{\frac{150}{U}},$$

где U — ускоряющее напряжение, В.

Меняя ускоряющее напряжение, можно получить заданную длину волны, определяемую из табл. 5.2.

Таблица 5.2

U, В	1	10	10^2	10^3	10^4	$1,5 \times 10^4$	$3,0 \times 10^4$	$4,0 \times 10^4$	$5,0 \times 10^4$
λ , нм	122,600	38,800	12,300	3,880	1,230	1,000	0,698	0,601	0,540

Так, уже при 15 кВ длина волны составляет 1 нм (или $0,1 \text{ \AA}$), что на четыре порядка меньше длины излучения УФ-излучателя. Именно поэтому в электронно-лучевой литографии пренебрежимо малы эффекты дифракции, ограничивающие разрешающую способность субмикронной литографии. Теоретически электронный пучок может быть сфокусирован в пятно, диаметр которого определяется дифракционным пределом:

$$d_{\text{диф}} = 1,22 \cdot \lambda / D \approx 10 \text{ \AA},$$

где D — апертурный угол, равный углу сходимости пучка, λ — длина волны излучения. Однако из-за ограничений, накладываемых абберрациями электронно-оптических систем, тепловым разбросом скоростей эмитируемых электроном, и их взаимодействием с материалом электронного резиста минимальные технологические линии составляют 30 \AA .

Для электронной литографии характерна значительная глубина фокуса электронно-оптической системы. Это позволяет значительно снизить требования к плоскостности и

положению обрабатываемых пластин, что особенно важно для технологии пластин большого диаметра.

В процессе электронной литографии из производства исключается изготовление промежуточных шаблонов, что существенно упрощает технологию, повышает производительность труда, расширяет конструктивные возможности и снижает стоимость производства. Дело в том, что при электронно-лучевом экспонировании локализация процесса взаимодействия безынерционных электронов на резисте осуществляется с помощью электрических и магнитных полей, управляемых встроенной ЭВМ. Поэтому в процессе обработки можно мгновенно и с большой точностью выбирать любую точку обрабатываемой поверхности.

Электроны в пучке характеризуются малой длиной волны излучения, но значительной кинетической энергией. Поэтому электронными потоками можно экспонировать как обычные фоторезисты, так и электронорезисты.

Сканирующая электронная литография является непосредственным методом генерирования топологии с высокой разрешающей способностью. В этом методе рисунок топологии вычерчивается с помощью электронного луча малого сечения, управляемого микропроцессором. Электронно-лучевые системы метода сканирования имеют две подсистемы: подсистему формирования и отклонения луча и подсистему генерирования рисунка и управления движением.

В подсистеме формирования луча создается круглый луч с гауссовым распределением плотности энергии или специально расформированный луч с квадратным сечением (рис. 5.7).

В подсистемах, формирующих электронный луч с гауссовым распределением (рис. 5.7, а), используется обычный подход, принятый для формирования электронного зонда в растровом сканирующем микроскопе. На поверхности пластины получается уменьшенное изображение. Возможности такого луча достаточно широки и определяют изменения фокусного расстояния линз. Диаметр луча выбирается равным примерно четверти минимальной ширины нужной топологической линии на рисунке.

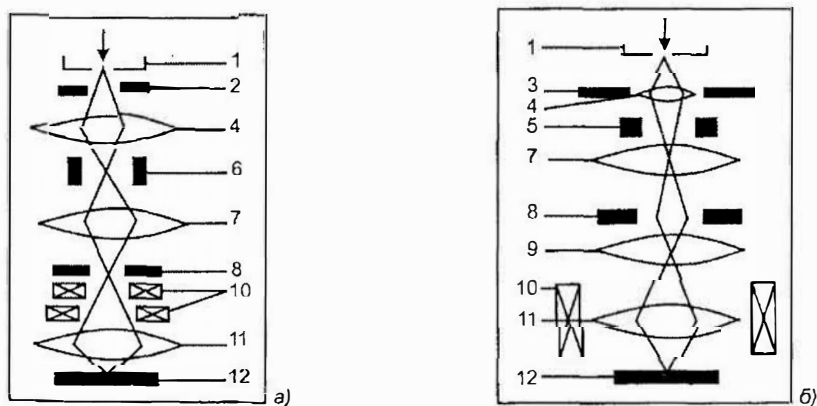


Рис. 5.7. Схемы формирования луча с гауссовым распределением энергии по сечению (а) и специальным сформированным лучом (б) квадратного сечения: 1 — электронная пушка; 2 — апертура рассеяния; 3 — квадратная формирующая апертура; 4 — первая конденсорная линза; 5 — первая уменьшающая линза; 6 — устройство гашения луча; 7 — вторая конденсорная линза; 8 — апертура, ограничивающая луч; 9 — вторая уменьшающая линза; 10 — отклоняющие устройства; 11 — оконечная линза; 12 — стол с пластиной

В подсистемах с квадратным сечением луча (рис. 5.7, б) источник электронов освещает квадратное отверстие в центре конденсированной линзы 3, которая создает изображение на входном зрачке конденсорной линзы 7. Использование уменьшающей линзы 9 и оконечной линзы 11 позволяет создать на пластине луч квадратного сечения. Размер квадратного луча обычно приближают к нормальной топологической ширине.

Методы сканирования луча приведены на рис. 5.8.

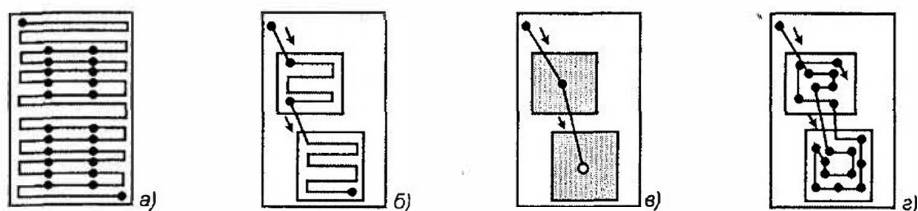


Рис. 5.8. Методы сканирования луча: растровый (а) и векторный (б), векторно-растровый (в), векторное сканирование с дополнительным оконтуриванием (г)

Высокое качество изображения на пластине постоянно контролируется и оперативно регулируется относительно электронного луча. Совмещение изображения осуществляется сканированием специальных маркерных знаков, представляющих собой прямоугольники или кресты из материала с большим атомным номером. Момент пересечения маркерного знака электрическим лучом фиксируется соответствующим детектором электронов.

Проекционная электронная литография использует широкий электронный пучок, проходящий через шаблон и воспроизводящий его изображение на покрытой слое резиста пластине.

Проекционные системы конструктивно проще потому, как данные о топологии для каждой экспозиции хранятся не в управляющем электронном устройстве, а заложены в маске. Разработаны два типа проекционных систем: с сохранением масштаба (1:1) и с редуцированием (10:1) и мультипликацией.

В проекционных системах с сохранением размеров используется принцип высокоразрешающего оптико-электронного преобразователя. Фотокатод или фотошаблон, с которого переносится изображение, предварительно изготавливается с помощью сканирующих систем (рис. 5.9, а).

Фотокатод, маскируемый тонким металлическим рисунком, эмитирует электроны, которые ускоряются высоким напряжением (20 кВ). Однородное магнитное поле фокусирует фотоэлектроны на анод (образец-пластина). Положение изображения определяется путем собирания характеристического рентгеновского излучения отметок на пластине. В этом случае фотокатод маскируют так, что освещенными остаются только метки совмещения.

Редукционная система с уменьшением масштаба изображения представляет собой электронно-оптический аналог редукционной проекционной оптической камеры (рис. 5.9, б). Пучок электронов из электронной пушки 7 формируется электродами 8 и конденсорными линзами 9 в широкий поток, попадающий на маску 10 из фольги. Проекционные линзы 11 и апертура позволяют сформировать на пластине 6 уменьшенное изображение. Размер поля ограничивается величиной порядка нескольких миллиметров, и для экспозиции очередного участка производится шаговое перемещение пластины.

Маска представляет собой свободно подвешенную металлическую фольгу. Для совмещения в установке используется сканирующий режим работы, при котором освещающий

луч фокусируется на маске и производит ее сканирование. Изготовление маски является весьма трудоемкой операцией и по существу определяет успех дела. В качестве маски можно использовать сочетания фотокатода с ускоряющей структурой.

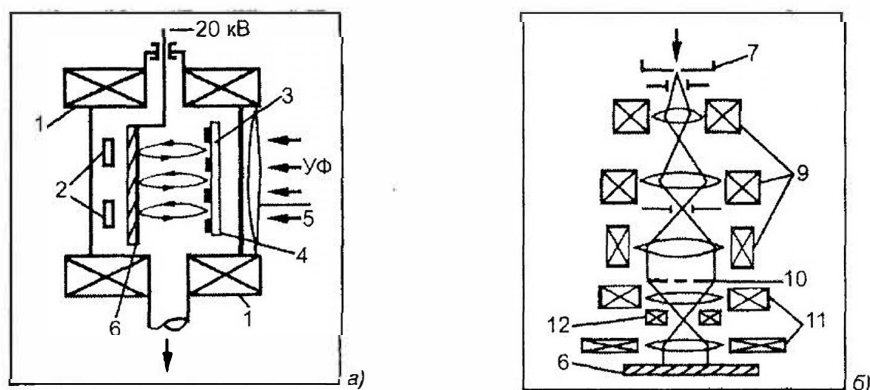


Рис. 5.9. Электронные проекционные системы с сохранением масштаба (а) и с редуцированием изображения (б): 1 — обмотка магнита; 2 — детекторы рентгеновского излучения; 3 — фотокатод с маской; 4 — электроды; 5 — УФ-источник; 6 — слой электронного резиста на Si-пластине; 7 — электронная пушка; 8 — электроды гашения луча; 9 — конденсорные линзы; 10 — маска; 11 — проекционные линзы; 12 — апертюра и катушка совмещения

При проекционном способе переноса изображения с уменьшением поле изображения невелико и трудно преодолеть искажения формы изображения, вызываемые внешними и внутренними паразитными электрическими и магнитными полями. Этот способ следует рассматривать как перспективный лишь для создания топологии в устройствах, где не требуется высокой точности совмещения.

Для проекционных установок радиус пятна R выражается соотношением

$$R = E_e / U,$$

где E_e — энергия эмитированных электронов ($\sim 0,1$ — $0,2$ эВ), U — напряженность электрического поля ($\sim 10^4$ В/см). В этом случае возможно получение размера пучка порядка 7×10^{-2} мкм и разрешения — $0,2$ мкм. Соответственно такого же размера можно получить и линии рисунка. Проекционная электронолитография более производительна, чем сканирующая. По разрешающей способности сканирующая электронолитография превосходит проекционную. К тому же шаблоны для проекционной электронолитографии конструктивно сложны и трудоемки в изготовлении.

Электронная литография в области субмикронной технологии может конкурировать с фотолитографией, несмотря на значительную стоимость установок (рис. 5.4, кривая 4).

5.3.3. Рентгенолитография

Рентгенолитография представляет собой современный высокоразрешающий метод создания топологии микроструктур с использованием рентгеновского излучения.

Этот метод напоминает фотолитографическую контактную печать сзором и использует коротковолновое рентгеновское излучение с длиной волны $\lambda \sim 0,4$ — $2,5$ нм. На рентгенолитографических установках может быть получено разрешение $\sim 0,05$ — $0,5$ мкм. Малая

энергия мягкого рентгеновского излучения позволяет избежать рассеяния и поглощения. Метод нечувствителен к частицам пыли вещества с малым атомным номером, к электрическому заряду на шаблоне или подложке.

Работы по созданию технологических установок рентгенолитографии развиваются по двум независимым направлениям.

Первое направление связано с конструированием установок пошагового мультиплицирования с использованием традиционных источников рентгеновских излучений в диапазоне 0,4—1,3 нм. Эти установки имеют относительно простые системы совмещения, где используются чувствительные ($\sim 5 \text{ Дж/см}^3$) негативные *рентгенорезисты* со временем экспонирования более 1 мин. На таких установках возможно получение линии рисунка с характерным размером 0,5—1,0 мкм.

Одновременно ведутся исследования, направленные на разработку установок литографии на основе синхротронных источников рентгеновского излучения очень высокой интенсивности. Их применение позволит резко повысить производительность. Однако в этом случае необходимы более сложные системы совмещения, а также менее чувствительные позитивные рентгенорезисты ($\sim 300 \text{ Дж/см}^3$). При этом возможно изготовление многослойных структур статических неоднородностей с размерами элементов 0,05—0,5 мкм.

Установки с *пошаговой мультипликацией* изображены на рис. 5.10. В установках этого класса используются излучатели с вращающимися анодами из меди, алюминия, молибдена, кремния, палладия, серебра. Сфокусированный на такой анод пучок электронов, взаимодействуя с материалом анода, генерирует рентгеновское излучение. При этом анод обязательно охлаждается водой. Пройдя через рентгеношаблон, излучение формирует на подложке рисунок.

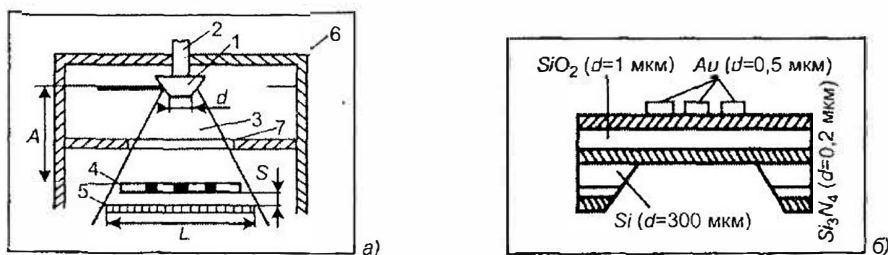


Рис. 5.10. Схема с пошаговой мультипликацией (а) и рентгеношаблон на основе кремния (б):
1 — анод; 2 — устройство вращения анода; 3 — пучок рентгеновского излучения; 4 — рентгеношаблон;
5 — подложка

Излучатель размещен в специально сконструированной вакуумной камере с бриллиевым окном. Ввиду отсутствия эффективных систем формирования рентгеновского излучения используется расходящийся пучок излучения. Геометрические параметры установки оптимизируют таким образом, чтобы сократить время экспонирования и получить минимальную размытость изображения ($\delta = Sd/A$) и искажение геометрии рисунка ($\Delta = SL/2.4$). Время экспонирования на каждом шаге в рентгенолитографии обычно на два порядка больше, чем в фотолитографии. В методе рентгенолитографии отсутствуют проблемы фокусировки и эффекты отражения и рассеяния, которые ограничивают возможности фото- и электронной литографии. Этот метод позволяет получать топологию с вертикальными стенками.

Одной из важных задач в рентгенографии является изготовление шаблонов (рис. 5.10, б). Материалы подложек и защитных слоев выбираются с учетом соответствия их спектральных характеристик свойствам рентгеновского излучения, применяемого в установке с заданным материалом анода. Так, материалом подложек могут служить Be, Si, Si_3N_4 , SiO_2 , Al_2O_3 и органические полимерные материалы типа майлара и полиамида. В качестве защитных слоев рекомендуется использовать Au, Pt, Re, Er и т. д.

Особенностью конструкции рентгеношаблона является использование в качестве подложек мембран микронной толщины. Это позволяет обеспечить достаточно хорошее пропускание излучения, но существенно усложняет технологию изготовления, а также использования рентгеношаблонов. Рентгеношаблоны изготавливаются методом электронной литографии.

В рентгенолитографии с *синхротронным источником излучения* вместо обычного источника X-лучей используется излучение релятивистских электронов, циркулирующих по замкнутой орбите в ускорителе-синхротроне или электронном накопительном кольце.

Синхротронное излучение возникает при движении релятивистских электронов в магнитном поле по круговым орбитам. Энергия электронов должна быть большой величины

$$E \gg E_0 = m_0 c^2,$$

где m_0 — масса покоя электрона, c — скорость света. Достигнув нужной энергии, электроны выводятся из синхротрона. Спектр синхротронного излучения зависит от энергии электронов. Синхротронное излучение характеризуется высокой интенсивностью $E = 10^4$ эВ и малой расходимостью пучка.

Установка рентгенолитографии с синхротронным источником показана на рис. 5.11.

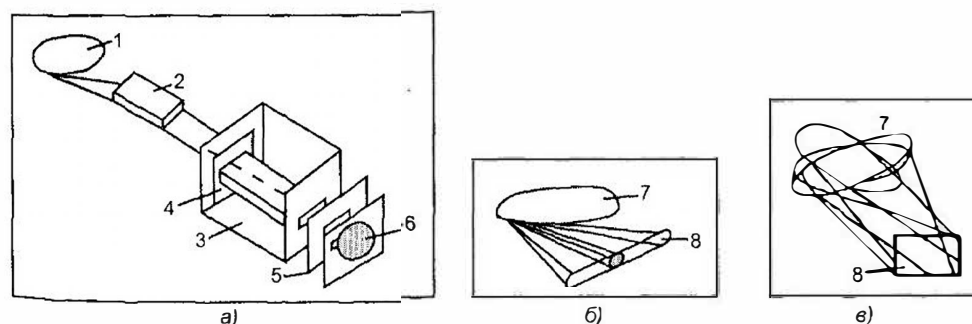


Рис. 5.11. Установка рентгенолитографии с синхротронным источником: 1 — накопительное кольцо с орбитами электронов; 2 — сканирующее зеркало; 3 — гелиевая камера; 4 — бериллиевое окно; 5 — рентгеношаблон; 6 — кремниевая пластина; 7 — электронные орбиты; 8 — зона экспонирования

Источником интенсивного X-излучения является электронное накопительное кольцо. Оно представляет собой замкнутые (как правило, круговые) орбиты электронов. По этим орбитам диаметром ≥ 10 м электроны циркулируют в течение десятков часов, накапливая энергию.

Накопительные кольца имеют октагональную конфигурацию и могут генерировать различные пучки излучения. Пучок генерируемого излучения — длительный и узкий (рис. 5.11, б). Если воздействовать на него магнитным полем и придать электронам волнообразное движение, то можно увеличить ширину пучка излучения (рис. 5.11, в).

Еще одним источником рентгеновского излучения может служить лазерная плазма, возникающая при воздействии интенсивного лазерного излучения на металлическую мишень. Сравнительные характеристики источников рентгеновского излучения приведены в табл. 5.3.

Таблица 5.3

Источник	λ , нм	Режим	Яркость, фотон λ (см ² ·с)	Размер, мкм
Трубка с Al катодом ($P \approx 0,25$ кВт)	0,83	Непрерывный	2×10^{16}	0,3—0,8
Лазерная плазма ($E \approx 100$ Дж, $t \approx 1$ нс)	0,8	Импульсный	10^{20}	0,3—0,8
Синхротронное излучение ($E \approx 1,5$ ГэВ)	0,3	Непрерывный	2×10^{25}	0,1—0,8

5.3.4. Ионная литография

Ионная литография представляет собой современный высокоразрешающий метод создания топологии микроструктур с использованием ионных или протонных пучков излучения.

Масса ионов значительно больше массы электронов, поэтому пучки протонов или ионов значительно меньше подвержены влиянию паразитных электрических полей. По этой же причине ионные пучки легче фокусировать и коллимировать. Из-за отсутствия эффекта близости ионно-оптическая фокусировка позволяет хорошо фокусировать изображение и исключать возникновение полутеневых искажений. Вследствие большого сечения пучка протонов и высокой чувствительности позитивного резиста к протонному излучению экспонирование происходит за доли секунды, что позволяет создать процессы высокой производительности.

Известны три метода ионно-лучевых литографических (ИЛЛ) процессов:

- ☐ использование коллимированных управляемых ионных пучков;
- ☐ использование сфокусированных ионных пучков (ФИП);
- ☐ использование ионно-проекционных систем.

Схема ИЛЛ с *коллимированным ионным пучком* аналогична системе рентгенолитографии с зазором и приведена на рис. 5.12.

Пластина, на которой необходимо создать заданную топологию, покрывается слоем резиста и экспонируется через шаблон, помещенный возле нее с небольшим зазором. Экспонирование производится путем сканирования коллимированным протонным пучком. Ток пучка с сечением 1 см² достигает 1 мкА, энергия протонов в пучке составляет 150—250 кэВ. При такой энергии протоны имеют малую длину пробега в материале ионошаблона и ионы сильно поглощаются. Наиболее удобен ионошаблон с отверстиями на месте экспонируемых участков.

Метод литографии острого фокусированного пучка напоминает сканирующую электронолитографию, однако отличается более точным управлением пучком с применением элек-

тростатических систем (рис. 5.12, б). Топология создается непосредственно на резисте, что позволяет исключить изготовление шаблонов. Основными проблемами реализации этого метода ИЛЛ являются эффекты расхождения пучка из-за высокой плотности тока, яркости ионных источников, а также хроматической aberrации.

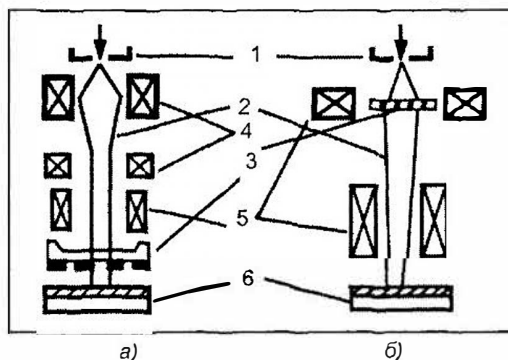


Рис. 5.12. Схема установок ИЛЛ с использованием коллимированных ионных пучков (а) и ионно-проекционной системой: 1 — источник ионов; 2 — ионный пучок; 3 — ионшаблон; 4 — коллимирующая система; 5 — фокусирующая система; 6 — подложка

Метод ИЛЛ *сфокусированных пучков* может быть использован для изготовления масок для рентгенолитографии и вполне конкурентоспособен, например, с электронно-лучевой литографией. Созданные технологические установки позволяют получить ширину линии 0,4—0,1 мкм.

Системы ИЛЛ со *сфокусированным пучком* могут объединяться с электронно-лучевыми системами литографии.

Принцип работы систем *ионно-лучевой проекционной литографии* аналогичен принципу работы проекционных установок фотолитографии с пошаговым экспонированием.

Ионные пучки могут быть использованы не только для экспонирования, но и для одновременного проведения процессов имплантационного легирования, ионного травления, а также в ряде других процессов создания субмикронных структур.

5.3.5. Лазерная литография

Среди рассмотренных видов излучения, используемых в литографии, весьма перспективным является лазерное когерентное излучение. Оно обладает высокой спектральной яркостью, возможностью концентрации энергии до высоких уровней ($\approx 10^6$ Вт/см).

Однако пионерские разработки лазерного оборудования в области литографии еще недостаточно зарекомендовали себя в производстве. В настоящее время применение лазеров в литографических технологических процессах и оборудовании происходит по нескольким направлениям:

- установки непосредственного экспонирования фоторезиста;
- лазерные установки для изготовления промежуточных шаблонов;
- голографические системы контроля и ретуши фотошаблонов.

Перспективными разработками стали фотолитографические установки на основе эксимерных лазеров, генерирующих в УФ-области спектра излучения.

5.4. Процессы локального изменения свойств полупроводников

5.4.1. Эпитаксия

Эпитаксия (от греческого *epi* — на, над и *taxis* — расположение) представляет собой процесс наращивания монокристаллических слоев вещества на подложку (кристалл), при котором кристаллографическая ориентация наращиваемого слоя повторяет кристаллографическую ориентацию подложки.

Различают гетероэпитаксию и гомоэпитаксию.

Процесс *гомоэпитаксии* или *автоэпитаксии* позволяет получать наращиваемые слои с составом, аналогичным подложке или отличающимся только от нее примесными свойствами.

Процесс *гетероэпитаксии* позволяет получать слои с различными относительно подложки свойствами как по химическому составу, так и по кристаллической структуре.

Процесс эпитаксии легко осуществляется, если разность параметров решетки лежит в пределах 10%. В этом случае тонкий эпитаксиальный слой продолжает структуру атомных плоскостей подложки.

Эпитаксиальное наращивание возможно из любой фазы вещества: газовой — газофазная эпитаксия (ГФЭ); жидкой — жидкофазная эпитаксия (ЖФЭ) и твердой — твердофазная эпитаксия (ТФЭ).

Методы *газофазной эпитаксии* делятся на физические и химические.

К физическим методам относятся методы термического осаждения из молекулярных пучков в вакууме, методы катодного распыления и осаждения, метод мгновенного расплавления.

Идеальные условия роста эпитаксиальной пленки заключаются в создании таких условий, при которых осаждение происходит на достаточно прогретую подложку. В этом случае подвижность падающих атомов или молекул обеспечивает формирование упорядоченной структуры. Температура подложки обычно лежит в пределах 576—673 К. Увеличение температуры приводит к переиспарению падающих молекул с подложки и их конденсации на более холодных частях установки. Создание пересыщенного пара приводит к созданию условий конденсации весьма далеких от термодинамического равновесия. Это особенно усложняет процесс напыления сложных соединений и сплавов. При напылении из одного источника испарение должно происходить конгруэнтно. Если же компоненты соединения диссоциируют, то их летучесть должна быть одинакова. Этого трудно добиться и поэтому получение однородных по толщине и составу эпитаксиальных пленок является весьма трудной задачей.

Следует отметить, что пленки, полученные классическим методом вакуумного испарения со многими источниками, позволяют получить монокристаллические структуры, причем их параметры будут близки к параметрам объемных кристаллов этих соединений.

Модификацией метода классического вакуумного напыления является метод мгновенного испарения. Сущность метода заключается в поддержании динамического равновесия между непрерывно и равномерно поступающим в испаритель веществом и составом паровой фазы. С этой целью в разогретый испаритель с постоянной скоростью подается гранулированный порошок испаряемого материала. Возгонка и конденсация его на подложке

происходят минуя жидкую фазу. Это возможно, когда температура испарения выше температуры, необходимой для испарения наименее летучего компонента, но ниже температуры плавления испаряемого материала. Недостатком метода мгновенного испарения является сравнительно высокая скорость напыления, зависящая существенным образом от температуры испарителя и размера зерен.

Дальнейшим развитием метода вакуумного испарения со многими источниками стала технология эпитаксии из молекулярных пучков или *молекулярно-лучевая эпитаксия* (МЛЭ).

Отличительными чертами технологии получения пленок методом молекулярно-лучевой эпитаксии являются:

- сверхвысокий вакуум в камере ($1,3 \times 10^{-8}$ Па);
- контроль структуры и состояния подложки в течение всего процесса напыления;
- контроль и регулирование состава конденсата остаточной атмосферы;
- контроль состава молекулярных пучков.

Прецизионное управление составом молекулярных пучков позволяет производить напыление любых структур как из отдельных элементов, так и из их произвольного сочетания. Этот метод позволяет создать структуры с любым заданным профилем легирования.

Установка МЛЭ представляет собой весьма сложную вакуумную систему, отвечающую вышеперечисленным требованиям (рис. 5.13).

Она состоит из двух камер, имеющих сверхвысоковакуумную безмаслянную откачку. В камере роста происходит формирование кристаллических пленок. Анализ свойств подложек и пленок происходит в камере анализа, конструкция которой предусматривает передачу механических усилий между камерами с целью перемещения держателя подложки из одной камеры в другую без их разгерметизации.

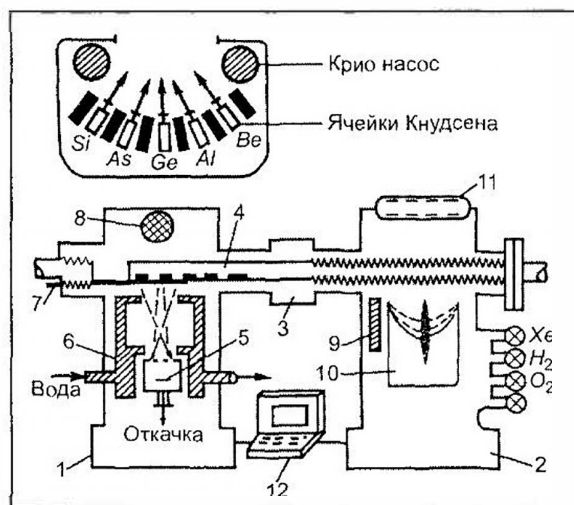


Рис. 5.13. Схема установки МЛЭ и вариант конструкции испарителя: 1 — камера роста; 2 — камера анализа; 3 — вакуумный клапан; 4 — держатель подложки; 5 — испарители; 6 — охлаждаемые экраны; 7 — заслонка; 8 — квадрупольный масс-спектрометр; 9 — ионная пушка; 10 — дифрактометр медленных электронов и оже-спектрометр; 11 — окно; 12 — микро-ЭВМ

Напыление пленок происходит из источников молекулярных пучков, в качестве которых используются испарители типа *ячеек Кнудсена*. Ячейка выполнена из спектрально-чистого графита, покрытого снаружи изолирующим слоем алунда (Al_2O_3). Нагрев осуществляется с помощью спирали из вольфрамовой проволоки. Чтобы исключить взаимное тепловое влияние источников, ячейки помещены в экраны из рифленой танталовой фольги и окружены общим медным экраном, который охлаждается жидким азотом. Точность поддержания температуры ячейки составляет один градус. Отличительной особенностью испарителя типа ячейки Кнудсена является возможность создания ламинарных, безвихревых потоков молекул. Перекрытие молекулярного потока осуществляется с помощью заслонки, управление которой происходит вне камеры без ее разгерметизации. Очистка поверхности подложки производится отжигом и бомбардировкой ионами инертных газов.

Контроль химического состава и рельефа поверхности производится с помощью оптического спектрометра и дифрактометра медленных электронов. Эти же приборы позволяют исследовать химический состав, кристаллическую структуру и ориентацию полученной пленки. Анализ остаточной атмосферы в камере, состав и интенсивность молекулярных пучков осуществляются с помощью квадрупольного масс-спектрометра. Контроль толщины эпитаксиальных пленок производится системой на основе кварцевого измерителя толщины. Но вполне понятным причинам управление процессом формирования статических неоднородностей необходимо проводить в строгом соответствии с измеренными параметрами структур, оперативным внесением корректив в процесс. Осуществление этого возможно только с помощью встроенной микро-ЭВМ, позволяющей управлять процессом в реальном масштабе времени. Поэтому этим методом выращиваются сверхтонкие (10—100 нм) эпитаксиальные пленки полупроводников, диэлектриков и металлов, создаются сверхрешетки, осуществляется многослойная застройка решеток.

К достоинствам метода МЛЭ следует также отнести:

- возможность оперативного контроля и управления стехиометрией состава структур;
- обеспечение формирования весьма однородных по составу и свойствам пленок практически любой толщины с заданным профилем легирования;
- исключительные возможности для автоматизации всего процесса напыления с помощью встроенной микро-ЭВМ.

Химические методы эпитаксии основаны на осаждении из газовой фазы вещества, полученного в результате ряда химических реакций. Например, восстановление хлоридов Si и Ge в водороде: $\text{SiCl}_4 + \text{H}_2 = \text{Si} + 4\text{HCl}$ или пиролитическое разложение моносилана: $\text{SiH}_4 = \text{Si} + 2\text{H}_2$ и др. Химические процессы проводятся в реакторах (рис. 5.14).

В реактор подается газовая смесь необходимого состава и соединения легирующих элементов.

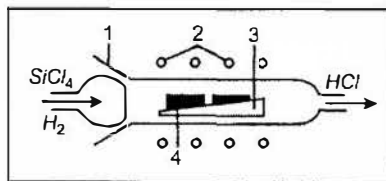


Рис. 5.14. Схема реакторной камеры для эпитаксии хлоридным методом: 1 — реакторная камера; 2 — нагреватель; 3 — кварцевая подложка; 4 — подложка

Методы *жидкофазной эпитаксии* основаны на процессах кристаллизации из раствора или расплава. Обычно жидкофазная эпитаксия производится при температуре 400—500 °С.

что позволяет получать многослойные эпитаксиальные структуры и пленки заданной конфигурации.

Методы твердофазной эпитаксии основаны на процессах ориентированного роста при изотермическом отжиге.

Методы ЖФЭ и ТФЭ распространены меньше, чем метод ГФЭ. Эпитаксиальные структуры для различных применений выпускаются в виде круглых пластин диаметром 80, 100, 150, 200 мм и толщиной от 200 до 400 мкм.

Марку эпитаксиальной структуры принято обозначать дробно, в числителе которой приводится характеристика эпитаксиального слоя, а в знаменателе --- характеристика кремниевой подложки. Перед дробью стоит цифра, указывающая на диаметр эпитаксиальной структуры.

К примеру, $125 \frac{10\text{КЭФ}0,5}{200\text{КДБ}10(110)}$ расшифровывается следующим образом: подложка кремния диаметром 125 мм, ориентированная в плоскости (110) с дырочной проводимостью, легированная бором с удельным объемным сопротивлением 10 Ом·см и толщиной 200 мкм, на которой выращена эпитаксиальная структура пленки кремния толщиной 10 мкм с электронной проводимостью, легированная фосфором с удельным сопротивлением 0,5 Ом·см.

5.4.2. Легирование полупроводников

Легирование полупроводников представляет собой процесс введения примесей или структурных дефектов с целью направленного изменения электрических свойств.

Электрические свойства легированных полупроводников зависят от природы и концентрации введенных примесей доноров или акцепторов. Примеси, как правило, образуют в полупроводниках твердые растворы замещения и обладают высокой растворимостью вплоть до значений 10^{18} — 10^{20} см⁻³. Эти примеси имеют малые сечения захвата носителей,

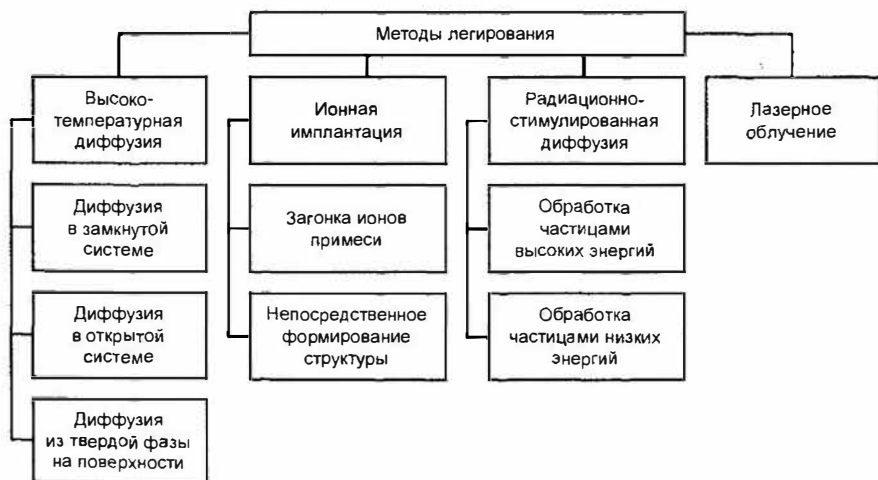


Рис. 5.15. Классификация методов локального легирования полупроводников

являются малозффективными центрами рекомбинаций и практически не влияют на время жизни носителей заряда.

Методы легирования делятся на две группы: либо непосредственно в процессах выращивания монокристаллов и эпитаксиальных структур, либо локальное легирование отдельных областей монокристаллов.

На рис. 5.15 приведена классификация методов легирования полупроводников. Дадим краткий анализ этих методов.

Высокотемпературное легирование

Метод высокотемпературной диффузии состоит в том, что легирующая примесь приводится в соприкосновение с поверхностью монокристалла кремния. Монокристалл разогревается, и атомы примеси проникают внутрь монокристалла, замещая атомы кремния в решетке.

Метод высокотемпературной диффузии является одним из наиболее развитых и широко используемых процессов создания локальных областей с заданным типом проводимости в полупроводниках.

Существует несколько механизмов внедрения примесей.

При *диффузии замещения* (диффузия по вакансиям) решетка испытывает при нагреве сильную тепловую вибрацию. Некоторые атомы покидают свои места и замещаются другими. Если поблизости окажется атом примеси с приблизительно теми же размерами и валентностью, то может произойти замещение ушедшего атома в узле решетки атомом примеси. Атомы примесей движутся по узлам решетки стохастически во всех направлениях, но в целом их движение направлено в сторону снижения концентрации примеси. Скорость диффузии замещения зависит от скорости возникновения вакансий в решетке.

При *диффузии внедрения* атом примеси находит себе место в кристаллической решетке в междоузлиях, не замещая атома исходного элемента. При высоких температурах атомы примеси могут перепрыгивать с одного междоузлия на другое, распространяясь с большей скоростью, чем при диффузии замещения.

Для изготовления барьеров и р—п-переходов используется процесс диффузии примесных атомов, которые внедряются в кристаллическую решетку полупроводника и меняют ее электрофизические свойства.

Процесс изотропной диффузии описывается с помощью коэффициента диффузии D и определяется из *первого закона Фика*:

$$j = -D \cdot dN / dx, \quad (5.1)$$

где j — плотность потока диффундирующих атомов, N — концентрация атомов, x — координата.

Этот феноменологический закон был сформулирован в 1855 году А. Фиком по аналогии с уравнением теплопроводности.

Коэффициент диффузии D зависит от температуры T следующим образом:

$$D = D_0 \cdot \exp(-\varepsilon / kT),$$

где D_0 — постоянная величина, ε — энергия активации для скачка атома, k — постоянная Больцмана.

Изменение концентрации диффундирующего вещества во времени при одномерной диффузии определяется *вторым законом Фика*:

$$\frac{dN}{dt} = \frac{d}{dx} \left(D \frac{dN}{dx} \right) = D \cdot \frac{d^2 N}{dx^2}. \quad (5.2)$$

Легирующая примесь в полупроводник может быть введена из газообразного, жидкого или твердого источника. Поэтому по способу подведения примеси из внешнего источника к среде, в которой необходимо создать неоднородности концентрации примесей, различают три метода.

В процессе проведения *диффузии в замкнутой системе* (метод "закрытой трубы") пластины проводника и источники примеси (диффузانت) загружают в кварцевую ампулу, вакуумируют, герметизируют и помещают в печь (рис. 5.16, а). В качестве диффузанта могут быть использованы вещества в любом агрегатном состоянии. Закрытая труба сначала разогревается до температуры разложения газообразного источника. На легируемых пластинах формируется локальный источник диффузии, место которого определяется соответствующей маской. Затем температура повышается, благодаря чему создаются необходимые условия диффузии легирующей примеси в объеме полупроводника. В результате формируются области с заданной концентрацией примеси и заданным типом проводимости. Метод имеет ряд недостатков:

- электрофизические параметры легированных областей существенно зависят от состава атмосферы в изолируемом рабочем объеме;
- двухступенчатый процесс термообработки увеличивает продолжительность процесса легирования;
- подготовка закрытой трубы требует много времени;
- получение образца сопровождается разрушением трубы;
- низкая воспроизводимость электрофизических параметров при низкой поверхностной концентрации примеси.

Чаще всего этот метод применяется с целью глубокого легирования.

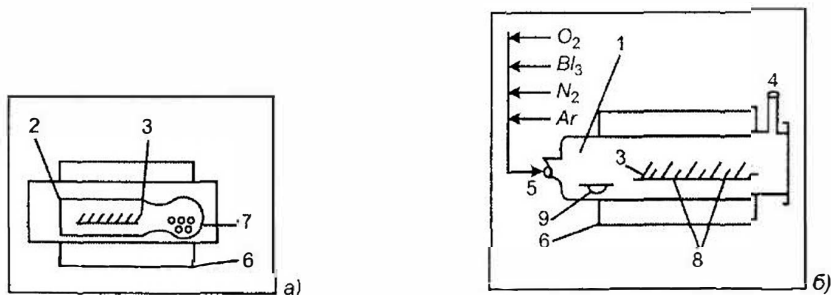


Рис. 5.16. Схема способов проведения диффузии: 1 — открытая диффузионная труба; 2 — закрытая диффузионная труба; 3 — пластины; 4 — вытяжка; 5 — выходной конец для введения газа-носителя; 6 — диффузионная печь; 7 — ампулы с диффузантами; 8 — твердый диффузонт; 9 — жидкий диффузонт

Наиболее широкое применение получил метод *диффузии в открытой системе* (метод "открытой трубы"). Диффузия в этом случае может проводиться из твердых, жидких или газообразных источников. Основными диффузантами при диффузии в кремний являются бор и фосфор. Создание локальных неоднородностей проводится в две стадии. На первой

стадии (*стадия загонки*) в необходимых местах поверхности при невысоких по сравнению с диффузией температурах создают тонкий диффузионный слой примеси. На второй стадии (*стадия разгонки*) пластины нагревают в атмосфере, не содержащий диффузанта. При этом происходит диффузионная перераспределение примесей, приводящая к созданию локальной области в объеме пластины.

Двустадийная диффузия в этом методе имеет два основных преимущества. Во-первых, разделение процесса на две стадии делает технологический процесс более управляемым, повышает воспроизводимость результатов и упрощает контроль. Во-вторых, низкотемпературный процесс стадии загонки облегчает маскирование будущих статических неоднородностей. Стадия разгонки происходит в отсутствии паров диффузанта.

Высокотемпературная диффузия возможна с использованием твердых планарных источников легирующей примеси. Существует несколько способов использования твердых источников. Наиболее привлекательным является расположение источника примеси параллельно между легируемыми полупроводниковыми пластинами. Расстояние между пластинами и диффузантом определяется соотношением

$$L > 2\sqrt{Dt},$$

где D — коэффициент диффузии примеси в газовой среде, t — время обработки пластин. Обычно значение L (пластина-диффузант) подбирают близким к размеру полупроводниковых пластин, а материалом служит стеклокерамическая композиция с использованием окиси бора (B_2O_3), нитрид-бора (BN), фосфата алюминия ($Al_2O_3 \cdot 3P_2O_5$).

Использование твердых планарных источников позволяет улучшить однородность легирования, поднять процент выхода годных структур, рационально использовать рабочий объем диффузионной печи, исключить токсичные продукты реакции и т. д.

Метод высокотемпературной диффузии из твердой фазы позволяет также сократить в 2—4 раза длительность высокотемпературной обработки, вводить донорные и акцепторные примеси одновременно, сократить трудоемкость операции легирования в 2—3 раза при одновременном улучшении электрофизических параметров структур.

Решение второго закона Фика для заданных начальных и граничных условий имеет вид:

$$N(x, t) = N \cdot \operatorname{erfc} \frac{x}{2\sqrt{Dt}}, \quad (5.3)$$

где $N(x, t)$ — концентрация примесей на расстоянии x от поверхности, N — поверхностная концентрация примесей, t — время диффузии.

Формирование областей транзистора происходит путем задания концентрационных и временных параметров диффузионного процесса.

На рис. 5.17 приведены расчетные кривые (α) и полученный профиль распределения примесей (\bullet) в областях биполярного транзистора.

Подводя итоги анализа методов высокотемпературной диффузии, отметим, что эти технологические процессы обеспечивают создание локальных неоднородностей в полупроводниковых континуальных средах в виде областей с широким диапазоном концентрации и глубин залегания примесей.

Вместе с тем, методом высокотемпературной диффузии присущ ряд недостатков:

- изотропность, приводящая к ограничению размеров формируемой неоднородности в кремниевой пластине;

- изменение электрофизических характеристик, а также генерация дислокаций и механических напряжений в континуальной среде из-за обработки пластин при высоких температурах;
- ограниченный набор диффузانتов, в диапазоне температур 900—1250 °С, обладающих высоким коэффициентом диффузии и хорошей растворимостью;
- появление используемых примесей в активной области при формировании структуры вследствие протекания газов-носителей или проникающих через трубы газов из окружающей среды;
- трудность получения тонких легированных слоев и редких p—n-переходов.

Все это снижает эффективность методов высокотемпературной диффузии, затрудняет высокоэффективную автоматизацию и интеграцию всего технологического процесса. Поэтому поисковые работы в области технологии создания областей заданной проводимости и концентрации направлены на увеличение эффективности процессов, рациональном сочетании этих процессов с другими методами легирования.

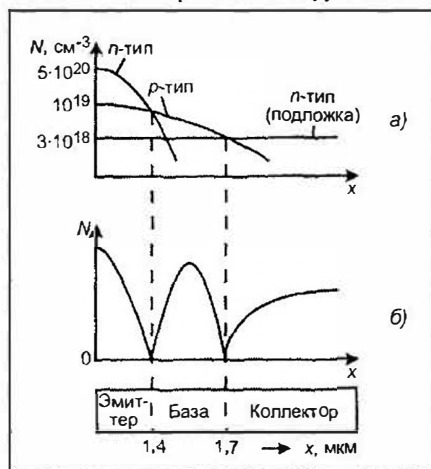


Рис. 5.17. Формирование методом высокотемпературной диффузии областей транзисторной биполярной структуры

Ионная имплантация

Ионная имплантация (ионное внедрение, ионное легирование) представляет собой процесс введения примесных атомов в твердое тело путем бомбардировки его поверхности ускоренными ионами.

Практически метод ионной имплантации состоит в бомбардировке твердых тел пучками ускоренных ионов с энергией в пределах от 10 кэВ до 1 МэВ с целью создания в них локальных неоднородностей и p—n-переходов. Ускоренные ионы проникают в кристаллическую решетку, преодолевая отталкивающее противодействие положительных заряженных ядер атомов. Ионы более легких веществ проникают глубже, чем ионы тяжелых веществ, однако траектория движения последних более прямолинейна.

Глубина проникания ионов в среду характеризуется понятием пробега. Средний нормальный пробег λ представляет собой проекцию траектории среднего полного пробега на направление первоначальной скорости иона. При точной ориентации направления падения пучка ионов вдоль кристаллографической оси происходит *каналирование* ионов в кри-

сталле. В этом случае пробег ионов существенно больше, чем при неориентированном облучении. Ион может вызвать в решетке зоны разупорядочения, размер которых может составлять 3—10 нм. Это происходит в том случае, когда энергия, переданная ионом атому решетки, превышает энергию связи атомов в твердом теле. Разупорядоченные зоны или радиационные дефекты при облучении накапливаются, и когда число их превосходит критическое значение, происходит локальный переход кристаллической структуры в аморфный слой. Большинство внедренных ионов находится в междоузлиях и не является электрически активным. Для восстановления кристаллической структуры облученную среду отжигают, в результате чего происходит распад радиационных дефектов. В это же время ионы примеси занимают вакантные узлы, внедряются, образуя области с заданной концентрацией примеси или р—n-переходы. Конструктивно установка для ионной имплантации представлена на рис. 5.18.

Источником ионов служит газоразрядная камера, в которой с помощью электронной бомбардировки происходит ионизация молекул и атомов имплантируемого вещества. Дополнительная ионизация осуществляется с помощью магнитного сепаратора. Обычно используются два или более источника ионов примеси для получения областей р- и n-типов. Например, в качестве источника акцепторной примеси используется трехфтористый бор (BF_3), поступающий из баллона с натекателем. В качестве донорной примеси часто используют пары фосфора, образуемые в нагревателе с красным фосфором. Ионы имитируются с помощью вытягивающих электродов, подключенных к источнику высокого напряжения. Ионный пучок формируется в канале с помощью магнитного сепаратора и систем фокусирования и сканирования. После прохождения магнитного сепаратора, в котором происходит отбор ионов по массе, они попадают в систему фокусировки.

Сформированный ионный пучок сканирует поверхность, расположенную в вертикальной плоскости. Заслонка в канале предназначена для перекрытия ионного пучка при наборе заданной дозы облучения. Ускорительная трубка с источником высокого напряжения (~ 200 кВ) предназначена для ускорения ионов до нужной энергии. Облучаемый материал помещается в кассету барабанного типа, расположенную в приемной камере, которая откачана до уровня 10^{-3} Па.

Однородность легирования конструктивно обеспечивается вращением барабана и сканированием ионного пучка в заданное время облучения t (с), определяемое из соотношения

$$t = \frac{2\pi Q}{\alpha j},$$

где α — угол сектора облучаемой окружности барабана; Q — доза облучения, Кл/м²; j — плотность ионного тока в пучке, А/м².

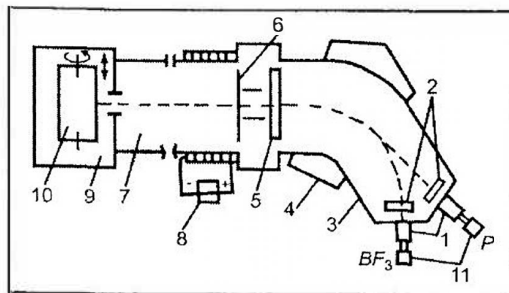


Рис. 5.18. Схема установки ионной имплантации:
1 — источник ионов;
2 — вытягивающие электроды;
3 — канал; 4 — магнитный сепаратор;
5 — система фокусирования и сканирования; 6 — заслонка;
7 — ускорительная трубка;
8 — источник высокого напряжения;
9 — приемная камера;
10 — барабан; 11 — баллон с натекателем

Доза облучения определяет число частиц, внедренных в единицу поверхности:

$$N = \frac{Q}{q} = \frac{It}{en},$$

где q — заряд одного иона, n — кратность ионизации, e — заряд электрона.

Отличительной особенностью метода ионной имплантации является возможность создания максимальной концентрации не на поверхности, что было характерно для диффузионных процессов, а в глубине среды. Это происходит потому, что с увеличением энергии ионов максимум концентрации располагается в глубине полупроводника. Поверхностная концентрация при этом падает. С повышением энергии ионов до 2,5 МэВ повышается глубина проникновения. Инверсионный слой концентрации примеси в этом случае формируется в глубине среды, образуя р—п- и п—р-переходы (рис. 5.19).

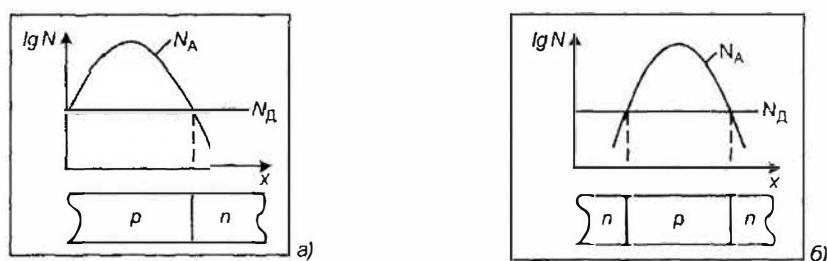


Рис. 5.19. Образование р—п-переходов при внедрении ионов малых (а) и больших (б) энергий

В режиме каналирования пучков возникают нежелательные слабо контролируемые эффекты. Следует отметить, что при полном легировании концентрация свободных носителей заряда в легированном слое меньше концентрации внедренных примесей. Это происходит из-за попадания ионов в междоузельное положение, где они электрически неактивны. Коэффициент использования примесей, например, для акцепторной примеси составляет

$$K_U^A = n_p + N_D / N_A,$$

где n_p — концентрация дырок в инверсном слое; N_D — концентрация доноров в исходном кремнии; N_A — средняя концентрация введенных атомов акцепторной примеси. Аналогично выводится коэффициент использования K_U^D при внедрении донорной примеси.

Электрические свойства полупроводника зависят как от количества имплантированных ионов, так и от числа радиационных дефектов.

На практике при создании транзисторных структур используют оба способа внедрения примесей: высокотемпературную диффузию и ионную имплантацию. В технологическом процессе на стадии загонки примеси используется ионная имплантация, а в дальнейшем диффузионная разгонка. В этой весьма перспективной технологии сочетаются достоинства методов: точность дозировки ионной имплантации и глубокие уровни залегания переходов, характерные для процесса диффузии.

В процессе ионного легирования важна операция отжига, при которой устраняются радиационные дефекты и активизируются внедренные атомы. Режим отжига зависит от дозы облучения. Для малых доз температура отжига составляет $T_{отж} \leq 260^\circ \text{C}$, для боль-

ших — $T_{отж} \geq 570^\circ\text{C}$. Заметим, что граница между понятиями *отжиг* и *диффузия* лежит вблизи 1000°C . При температурах $< 1000^\circ\text{C}$ процессами диффузии можно пренебречь.

Метод ионной имплантации, используемый для создания статических неоднородностей в микроэлектронных приборах и устройствах, имеет ряд преимуществ:

- ☐ нетермический характер введения примесей обеспечивает большой выбор легирующих примесей;
- ☐ чистота вводимой примеси в технологическом процессе;
- ☐ широкий диапазон концентраций вводимой примеси;
- ☐ эффективный контроль дозы вводимой примеси, что позволяет достичь высокой воспроизводимости результатов;
- ☐ высокая однородность легирования;
- ☐ возможность создания неоднородности проводимости в объеме континуальной среды путем подбора энергии ионов;
- ☐ возможность создания неоднородностей субмикронных размеров;
- ☐ сокращение длительности производственного процесса по сравнению с методом диффузии;
- ☐ возможность автоматизации технологического процесса.

Тенденции в развитии метода ионной имплантации заключаются в преодолении ограничений физического и технического характера, среди которых:

- ☐ возникновение радиационных нарушений кристаллической решетки, растущих с дозой внедрения примесей;
- ☐ трудность создания примесей на глубине более 2 мкм ;
- ☐ сложность эксплуатации технологического оборудования и его высокая стоимость.

Радиационно-стимулированная диффузия

Радиационно-стимулированная диффузия (РСД) представляет новое направление, являющееся комбинацией высокотемпературной диффузии и ионной имплантации. Этот метод сконцентрировал в себе ряд достоинств обоих ранее описанных методов. Сущность метода РСД заключается в бомбардировке кристалла легкими ионами, энергия которых передается атомам подложки. Вследствие этого наблюдается смещение атомов в междоузельное пространство и образуются вакансии. В определенных условиях вакансии могут мигрировать в кристалле, меняясь положением в решетке с соседними атомами основного кристалла или атомами примеси. Эта часть процесса носит диффузионный характер и аналогична термической генерации дефектов. Ионная бомбардировка увеличивает коэффициент диффузии любой примеси. При стимулировании диффузии в полупроводниках методом ионной бомбардировки главную роль играют неравновесные вакансии и междоузельные атомы. Ионизационные и полевые эффекты существенной роли не играют. Процесс РСД обычно описывается диффузионными уравнениями вида:

$$\frac{dN}{dt} = D \frac{d^2N}{dx^2} + v \frac{dN}{dx},$$

где D — коэффициент радиационно-стимулированной диффузии, v — скорость перемещения поверхности вследствие ее распыления. Приведенные уравнения несколько отличаются от известных уравнений Фика, описывающих процесс высокотемпературной диффузии.

Процесс РСД осуществляется ионами электрически неактивных примесей (H_2 , He, N, Ar и др.) высокой или низкой энергии. Обработка среды высокоэнергетичными ионами (10—100 КэВ), например, протонами обеспечивает целенаправленное управление концентрацией и типом распределения примеси в уже сформированных структурах. Глубина проникновения примеси зависит от длительности протонной бомбардировки, энергии протонов и интенсивности облучения.

К преимуществам метода РСД по отношению к методам высокотемпературной диффузии и ионной имплантации следует отнести:

- ☐ сравнительно низкую температуру подложки в процессе обработки;
- ☐ прямую зависимость скорости генерации дефектов от параметров луча, а не температуры подложки;
- ☐ формирование направления диффузии не градиентом концентрации примеси, а слоем нарушений, создаваемым ионным пучком;
- ☐ незначительные радиационные нарушения в легированном слое;
- ☐ активацию введенной примеси в узлы решетки уже в процессе обработки.

Лазерный отжиг

В настоящее время в микроэлектронике все большее распространение получает лазерное облучение полупроводниковых материалов с целью создания в них локальных неоднородностей и р—п-переходов.

К достоинствам лазерной обработки следует отнести:

- ☐ возможность локального изменения свойств материалов;
- ☐ локальный безынерционный нагрев практически до любой заданной температуры;
- ☐ более высокую активацию примеси, чем при обычном отжиге;
- ☐ получение качественной поверхности пластины после отжига.

В процессе легирования лазерное облучение среды используется как для непосредственного селективного легирования, так и для отжига пластин после проведения ионной имплантации, а также диффузии, эпитаксиального наращивания и т. д.

Особенно большое внимание в настоящее время уделяется лазерному отжигу полупроводниковых материалов после процесса ионной имплантации. Специфичность лазерного излучения позволяет сфокусировать световой пучок строго определенного спектрального состава на ранее облученную ионами поверхность. При этом кратковременно до необходимой температуры нагревается лишь очень тонкий слой или его выбранные участки.

Различают два механизма отжига имплантированных структур: отжиг мощным импульсным излучением и отжиг непрерывным излучением.

Благодаря фокусировке лазерного излучения удается локально поднять температуру. При этом происходят процессы рекристаллизации полупроводниковой структуры аналогично эпитаксии из твердой фазы. И в этом случае неповрежденная имплантацией основа полупроводника используется в качестве затравки, на которой наращивается материал той же самой структуры и ориентации. В этом случае скорость рекристаллизации составляет 0,001—0,01 м/с, что на три порядка меньше скорости рекристаллизации при облучении лазерными гигантскими импульсами.

Особо отметим, что при отжиге непрерывным лазерным излучением профиль распределения примеси остается прежним (рис. 5.20, а), в то время как при облучении мощным

лазерным импульсом профиль распределения примеси зависит от мощности гигантского импульса (рис. 5.20, б).

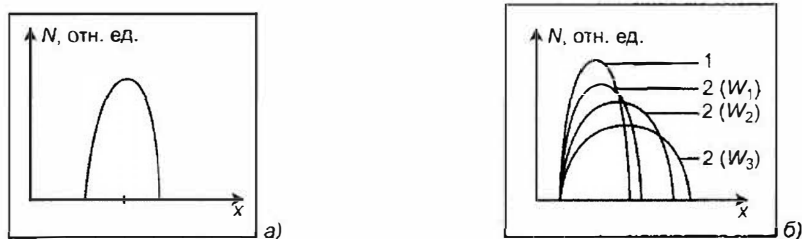


Рис. 5.20. Характер профилей распределения примесей в процессе лазерного отжига:

а — облучение непрерывным излучением малой мощности (до и после отжига);

б — облучение гигантскими импульсами;

1 — до отжига; 2 — отжиг излучением с изменяемой плотностью мощности $W_1 < W_2 < W_3$

С помощью лазерной рекристаллизации можно обеспечить строго ориентированный рост всех кристаллов, формируемых на аморфной подложке. Этот процесс может сопровождаться ростом пленки поликристаллического кремния, текстурированного в направлении (100) с размером зерен 1—100 мкм. Этот метод ориентированного выращивания пленок на аморфных подложках получил название *графоэпитаксии*.

5.5. Процессы обработки поверхности

5.5.1. Окисные пленки кремния

Окисные пленки кремния широко применяются в качестве маскирующих покрытий в процессах литографии при локальной эпитаксии и локальном травлении, для защиты и пассивации поверхности полупроводников, а также в качестве составного элемента в биполярных и МОП-транзисторных структурах. Легированные окисные пленки используются как источники примеси в процессах диффузии.

Из существующих методов окисления отметим следующие.

- ❑ *Окисление при высоком давлении* ($\sim 10^5$ Па) происходит при невысоких температурах процесса. Это позволяет уменьшить концентрацию дефектов. Толщина пленки в таком процессе возрастает пропорционально давлению. Этот процесс эффективен для небольших толщин окисла от десятых долей до двух микронетров.
- ❑ *Осаждение окисла из газовой фазы* происходит в результате пиролиза алкосиланов, окисления и гидролиза силана, гидролиза галогенов кремния. Этот метод позволяет нарастить на поверхность нагретого до 700—1000 К кремния пленки окисла заданной толщины.
- ❑ *Осаждение диоксида кремния из пленкообразующих растворов* позволяет производить контролируемое введение примесей в полупроводники. В состав пленкообразующих растворов входят кремнийорганические соединения, азотнокислые соли легирующих элементов. В качестве растворителей используют ацетон, этиловый спирт. Источником кислорода в процессах окисления служит кислород воды.
- ❑ *Нанесение окисла в вакууме* происходит путем вакуумного испарения или катодного расплавления. Исходным материалом является SiO_2 , наносимый в присутствии кислорода.

Скорость покрытия составляет 0,01—0,03 мкм/мин, что позволяет достаточно точно рассчитать значения необходимых толщин окисных пленок.

Контроль толщины окисных пленок, как правило, осуществляется *цветовым* методом, что позволяет осуществлять неразрушающий контроль толщины пленки.

5.5.2. Травление

Одним из важнейших процессов в технологическом маршруте создания интегральных схем является процесс травления.

Под *травлением* понимают растворение и последующее удаление заданной части материала с поверхности в целях полировки, изменения формы, очистки от загрязнений, а также для выявления структуры поверхности материала. Первые процессы из перечисленных выше относят к технологическим, а вторые — к структурным.

В микроэлектронике используют процессы *технологического травления*, которые делят на химические, физические и физико-химические методы.

Химическое травление основано на процессах растворения исходных материалов. Травление может осуществляться как в жидких, так и в газообразных средах. При этом обрабатываемый материал частично претерпевает химические изменения.

Различают следующие виды химического травления:

- изотропное растворение полупроводника с одинаковой скоростью травления по всем направлениям, которое применяют для удаления нарушенного слоя или полирования поверхности (рис. 5.21, а);
- анизотропное травление — растворение полупроводника с различной скоростью по разным направлениям монокристалла (рис. 5.21, б);
- селективное травление — растворение полупроводника с различной скоростью на разных участках поверхности с различным химическим составом;
- локальное травление — удаление материала со строго ограниченных и заданных участков полупроводника, которое обеспечивает получение заданного рельефа поверхности;
- послойное травление — равномерное последовательное снятие тонких поверхностных слоев.



Рис. 5.21. Изотропный с боковым подтравом ($L_{исх} \neq L$) (а) и анизотропный ($L_{исх} = L$) характер травления (б)

К физическим процессам травления относятся процессы травления, в которых удаление ненужного материала происходит путем физического распыления с помощью ионов газа.

Этот процесс называют *сухим травлением* и основан на использовании компонентов низкотемпературной газоразрядной плазмы — ионов, электронов, возбужденных атомов.

На рис. 5.22 приведена классификация процессов в зависимости от сочетания давления и энергии ионов плазмы.

Характер процессов меняется от чисто физического процесса расплавления до химического процесса сухого травления.

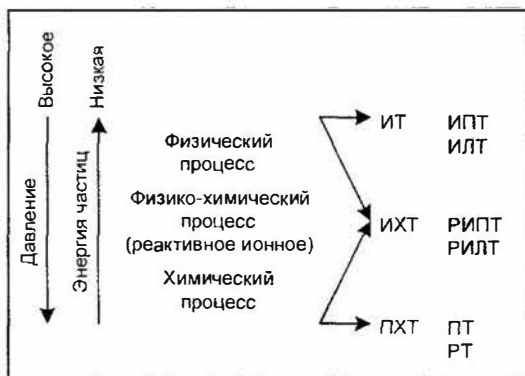


Рис. 5.22. Классификация процессов плазменного травления

К процессам травления предъявляются требования высокой селективности, отсутствия деградирующего влияния на свойства и размеры защитных масок, низкого уровня загрязненности поверхности материала и искажения полученного рельефа, высокой воспроизводимости и равномерности травления, минимального уровня загрязнения окружающей среды.

Ионное травление

Ионное травление осуществляется ионами плазмы, не реагирующими с обрабатываемым материалом. Этот процесс также называют травлением-распылением.

Ионное травление полупроводниковых структур осуществляется в реакторах со встроенными ВЧ-системами. Конструктивное решение таких реакторов может быть самым разнообразным с использованием диодной (рис. 5.23, а) и триодной системы (рис. 5.23, б), с распылителем магнетронного типа (рис. 5.23, в) и т. д.

Для обеспечения процесса ионного травления необходимо соблюсти следующие условия:

- ☐ травление структур должно осуществляться без существенной деградации электрофизических характеристик;
- ☐ травление большого количества подложек желательно максимального размера с высокой скоростью и достаточной равномерностью;
- ☐ минимальное загрязнение образцов инородными материалами;
- ☐ максимальная эффективность использования подводимой мощности;
- ☐ контроль скорости травления с возможностью фиксации границ слоев в многослойных структурах.

К основному недостатку процесса ионного травления относится отсутствие селективности травления. В связи с этим ограничивается глубина травления, которая должна быть

соизмерима с толщиной маски. Увеличение же толщины маски приводит к ухудшению передачи размеров. Вместе с тем в процессе ионного травления происходит разогрев пластин, в результате чего маска либо "задубливается", либо разрушается вообще. Этот процесс не позволяет получать субмикронные неоднородности в слоях рабочих материалов толщиной 0,1—0,3 мкм через органические маски. Эти недостатки, а также конструктивно-технологическая сложность реализации процесса ионного травления сдерживают широкое распространение этого метода в микроэлектронике.

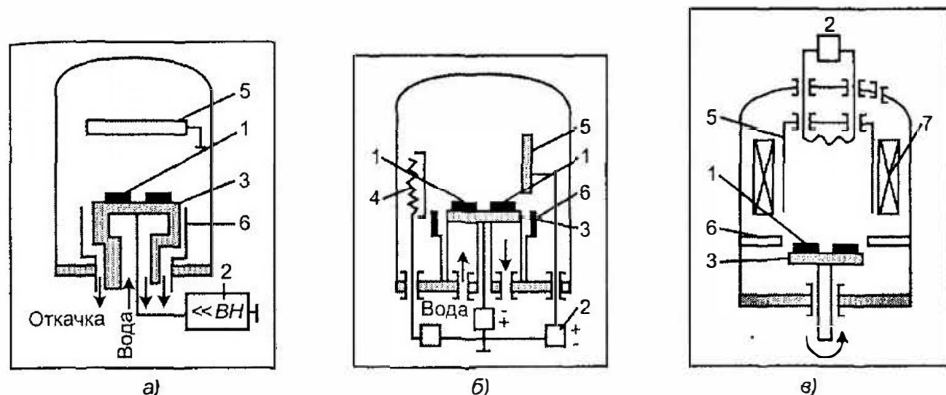


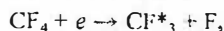
Рис. 5.23. Схема конструктивного решения установок ионного травления и распыления:

- 1 — мишень, в которой формируются статические неоднородности;
2 — источник питания, в т. ч. ВЧ-источники; 3 — мишень; 4 — термокатод;
5 — анод; 6 — экран; 7 — электромагнит

Ионно-химическое травление

Ионно-химическое травление (ИХТ) представляет собой физико-химический процесс, который происходит при достаточно высоком давлении газов и значительной энергии частиц.

Поверхностные слои материала удаляются с помощью физического распыления ионами, а также в результате химической реакции между активными частицами и поверхностными атомами. В этом процессе различают *реактивное ионно-плазменное травление (РИПТ)*, в процессе которого обрабатываемый материал находится в области плазмы, и *реактивное ионно-лучевое травление (РИЛТ)*, при проведении которого материал помещают в вакуумной зоне обработки. В первом случае химически активные частицы могут доставляться к поверхности из плазмы разряда и образовываться на поверхности при ударной диссоциации молекулярных ионов или нейтрализации атомарных ионов. Во втором случае на поверхность воздействуют только молекулярные или атомные ионы, способные образовывать химически активные частицы при ударной диссоциации или нейтрализации. В процессах ИХТ используются различные фреоны: CF_4 , CClF_3 , CClF_2 , CHClF_2 , C_2F_6 и т. д. В высокочастотном разряде молекулы фреона переходят в возбужденное состояние и образуют реакционноспособные радикалы, например:



где звездочкой помечен активный радикал.

Взаимодействуя с поверхностными частицами образца, эти радикалы формируют локальные неоднородности, а продукты реакции в виде летучих соединений удаляются из объема. Конструкции реакторов ИХТ различны и зависят от ионного источника бомбардирующих частиц (рис. 5.24). В большинстве систем используется ВЧ-разряд и диодная (рис. 5.24, а), триодная (рис. 5.24, б), планарная магнетронная (рис. 5.24, в) и другие типы систем электродов.

Процесс физического распыления возникает при энергиях ионов, превышающих 100 эВ. Физическое распыление активизирует поверхность материала, повышает скорость химических реакций, которые в свою очередь ослабляют химические связи поверхностных атомов и увеличивают скорость их физического распыления.

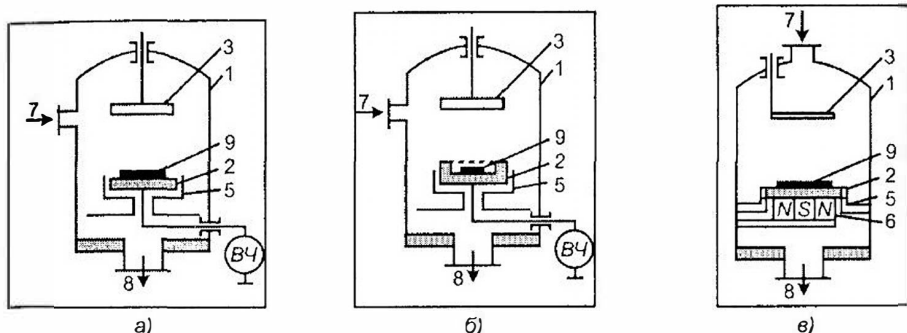


Рис. 5.24. Схемы установок для ионно-химического травления:

1 — камера; 2 — ВЧ-электрод; 3 — заземленный электрод; 4 — третий электрод; 5 — экран, 6 — магнитная система; 7 — натекатели; 8 — откачка; 9 — образец

Процессы ИХТ, например РИИТ, обладают высокой анизотропией и используются в качестве универсального процесса травления нитрида кремния, поликремния, фосфоросиликатного стекла, алюминия и других материалов.

Процессы ИХТ обладают способностью воспроизвести с шаблонов субмикронные (0,3—0,5 мкм) структуры. Наличие химических реакций позволяет значительно, по сравнению с ИТ, увеличить скорость и селективность травления при одновременном снижении теплового и радиационного воздействия на органические резистивные материалы. В процессе ИХТ удается с большой точностью переносить угловые и линейные размеры с маскирующего материала на рабочий, потому что отсутствует эффект переосаждения удаляемого материала, характерного для ионного травления.

Плазмохимическое травление

Плазмохимическое травление (ПХТ) происходит в результате химических реакций между химически активными частицами и поверхностными атомами материала.

Если обрабатываемый материал находится в области плазмаразряда, то процесс травления называется *плазменным* (ПТ). В этом процессе химические реакции травления будут активизироваться низкоэнергетической бомбардировкой как электронами, так и ионами. Если же материал находится в вакуумной (реакционной) зоне обработки, то травление производят только химически активные частицы без дополнительной электронной или ионной бомбардировки. Это так называемый процесс *радикального травления* (РТ).

Рассматривая процесс ПТ, отметим, что химически активные частицы (свободные атомы и радикалы) вступают в химическую реакцию с поверхностными атомами образца и удаляют поверхностные слои в результате образования летучих продуктов реакции. Роль электронов и ионов, присутствующих в плазме, заключается в увеличении скорости травления. Активирующее действие электронов и ионов определяется их энергией.

Генерация энергетических и химически активных частиц для процесса плазменного травления осуществляется в реакторе диодного типа, типовая конструкция которого приведена на рис. 5.25. Плазменное травление осуществляется при энергиях ниже 100 эВ.

В реакторах, использующих процессы радикального травления, подложки вынесены из области плазмы, и обработка идет по механизму гетерогенной химической реакции. При этом реакция не осложняется воздействием загрязненных частиц.

Современный уровень технологии плазменного травления отличается высокой однородностью и воспроизводимостью. Это возможно только при условии автоматизации управления такими параметрами реактора, как давление, мощность разряда, состав газов, скорость его протекания. С этой целью широко применяются встроенные микропроцессорные системы управления процессом.

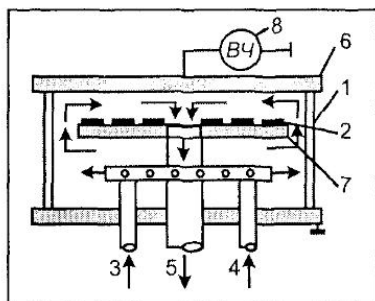


Рис. 5.25. Схема конструкций для плазменного травления:
1 — вакуумная камера; 2 — образцы;
3 — газ-травитель; 4 — газ;
5 — откачка продуктов реакции;
6 — верхний электрод; 7 — нижний электрод и стол; 8 — генератор ВЧ

Процессы плазмохимического травления могут обеспечить обработку поликремниевых структур, а также удаление масок с фоторезистов. Процессы ПТХ (особенно РТ) обладают наибольшей селективностью и оказывают незначительное тепловое и радиационное воздействие на обрабатываемые структуры по сравнению с процессами ИТ и ИХТ. Это позволяет осуществлять травление толстых слоев материалов (до 10 мкм) через тонкие "незадубленные" резисты. Низкий показатель анизотропии не позволяет с помощью процессов ПТХ получать статические неоднородности субмикронных размеров в слоях рабочих материалов толщиной более 0,3 мкм. Поэтому процессы ПХТ используются в основном для изготовления тонких (0,1—0,2 мкм) неорганических масок для процессов ИТ и ИХТ.

В последнее время появилось сообщение о разработке системы сухого плазменного травления с использованием электронного циклотронного резонанса. На этой системе получены локальные неоднородности на кремниевой пластине в виде линий с шириной 0,2—0,3 мкм. Эта система предназначена для производства первого поколения интегральных схем.

Лазерно-стимулированное травление

Лазерно-стимулированное травление — перспективный метод травления поверхности, в котором используется лазерное излучение с энергией фотонов от 0,1 эВ (ИК-излучения) до 6 эВ (УФ-излучения) для активизации травящих частиц на поверхности твердого тела.

Такие возбужденные частицы способствуют разрыву поверхностных связей или десорбции в газовую фазу. Они также способствуют ускорению химического воздействия между различными адсорбатами, адсорбатом и адсорбентом. Все это позволяет использовать процессы травления с возбуждением и нагревом поверхности для проекционного травления металлов и диэлектриков.

Процессы лазерно-стимулированного травления представляют собой качественно новый вид обработки материалов. Эти процессы основаны на управлении локальными химическими реакциями и фазовыми изменениями на границах раздела "твердое тело—газ" и "твердое тело—жидкость" с помощью лазерного излучения.

Этот метод позволяет получать требуемый рисунок с топологическими нормами субмикронных размеров.

5.5.3. Металлизация поверхности

Для соединения отдельных элементов транзисторных структур в электрическую схему используются тонкие пленки металла. Наибольшее распространение получили тонкие пленки алюминия. В последние годы внимание исследователей привлекает использование медных межсоединений, а также межсоединений из тугоплавких металлов, таких как молибден.

В основе процессов получения тонких пленок методом осаждения в вакууме лежат явления кинетики, испарения и конденсации газов.

Количество вещества M , испаренного с единицы поверхности в момент времени t , определяется потоком энергии Q , подводимым в данную точку поверхности, а также физическими характеристиками испаряемого материала. Изучение состава пучков, образующихся в процессе испарения, показывает, что поток испаренного вещества имеет широкий энергетический спектр, который не зависит от способа испарения. В этом потоке содержатся различные компоненты, такие как атомы и ионы, электроны и совокупности молекул (кластеры) исходного материала. Способ испарения определяет вклад каждого компонента потока. В свою очередь каждый компонент потока испаренного вещества определяет состав испаряемого материала.

Для выбранного способа испарения энергия частиц в потоке, их количество, степень ионизации зависит от теплоты испарения, коэффициента распыления, энергии ионизации и работы выхода материала. Качество получаемых пленок в сильной степени зависит от энергии частиц испаренного вещества.

Частицы с энергией, близкой к нулю, имеют малую подвижность на поверхности. Собственной энергии частице не хватает, чтобы найти свое устойчивое положение на поверхности. В условиях больших потоков происходит только рост островков в центрах кристаллизации. Пленка получается рыхлой, а адгезия плохой. Характеристики пленки могут быть улучшены за счет снижения плотности потока и подогрева подложки.

При повышении энергии частиц (≥ 10 эВ) адгезионные свойства улучшаются, обеспечивается сплошность пленки при невысоких температурах подложки. Частицы с высокой энергией (~ 300 эВ) могут уже удалять с поверхности адсорбированные молекулы газа, исключая из состава пленочной структуры газовые включения.

При увеличении энергии частиц свыше 300 эВ у них проявляется способность создания новых центров кристаллизации, нарушающих поверхностную структуру. Высокоэнергетические частицы такой энергии могут имплантироваться вглубь структуры (рис. 5.26).

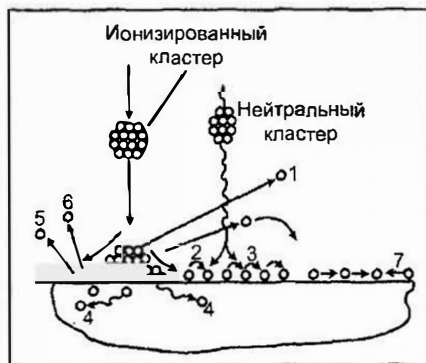


Рис. 5.26. Процессы, происходящие на подложке: 1 — частица конденсированного материала; 2 — сконденсированная частица; 3 — мигрирующая частица; 4 — имплантированная частица; 5 — частица распыленной подложки; 6 — переиспаренные частицы; 7 — образованная статическая неоднородность в виде тонкой пленки

Кластеры в молекулярном пучке имеют энергию не больше средней энергии частиц потока. Их скорость передвижения из-за большой массы незначительна по сравнению со скоростью перемещения ионов и электронов. Попадая на поверхность материала, кластеры образуют островки пленки с пористой структурой и плохой адгезией. Наличие кластеров в потоке приводит к образованию капель, которые попадают на поверхность и образуют там воронки, проколы в пленочных структурах, разрушают ранее созданные неоднородности.

Электроны, входящие в состав потока, образуются в результате ионизации части испаряемого материала при термоионной эмиссии. Возможен и механизм вторичной электронной эмиссии при некоторых способах термического испарения. Если в целом поток вещества электрически нейтрален, то электронная составляющая не играет существенной роли в процессе конденсации. Высокоэнергетичные электроны способны нагревать и очищать поверхность, на которой формируется статическая неоднородность. Если же поток электрически не скомпенсирован, то электронная составляющая может зарядить поверхность. Следствием образования статического заряда являются пробой, особенно частые в случае образования статических неоднородностей из диэлектрических или полупроводниковых материалов. Особое значение имеет ионный компонент в потоке испаряемого материала. Доля ионной составляющей должна быть тем больше, чем ниже энергия ионизации испаряемого материала. Заметим, что энергия ионизации тугоплавких материалов составляет 7—8 эВ, а для легкоплавких — 8—10 эВ. Поэтому наиболее ионизированным будет пучок, образованный из тугоплавких элементов. Именно ионный компонент ответствен за формирование на поверхности тонкой пленки.

Поток может быть сформирован различными способами. Различают четыре основных способа формирования потока испаряемого вещества (рис. 5.27).

Термический способ использует разогрев вещества до температуры испарения и последующее формирование потока. Энергия частиц в таком потоке составляет 0,2—5 эВ при плотности потока $10^{18} \text{ см}^{-2} \cdot \text{с}^{-1}$ и незначительной ионизации (2%). В настоящее время известны различные способы термического разогрева. Резистивный нагрев осуществляется за счет выделения джоулева тепла в испаритель в виде проволочного нагревателя, лодочки, ячейки Кнудсена при пропускании через него электрического тока. В качестве материала нагревателя используется обычно тугоплавкие материалы (Ta, Mo, Pt, Ir, C и др.). В соответствующих установках с помощью таких испарителей можно распылять различные материалы (Al, Au, Cr, Ni, Fe, Sn, Cu), а также соединения (SiO_2 , SiO, MgF_2 , ZnS, Sn, Te, JnPb, GaAs и др.).

Для термического испарения применяется также и ВЧ-нагрев. В этом случае может достигаться высокая температура нагрева испаряемого вещества ($\sim 2800^\circ\text{C}$). Это вещество обычно помещается в тигель из Ta, Mo, Jr, Pt или C. Таким способом возможно получение пленок Al_2O_3 , SiO_2 , ZnO_2 , TiO_2 , Y_2O_3 , которые широко применяются в приборах микроэлектроники.

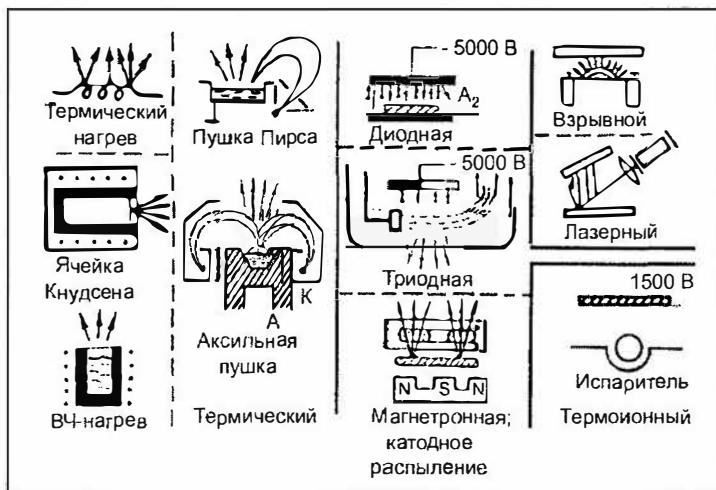


Рис. 5.27. Способы формирования потоков испаряемого вещества

Среди термических способов электронно-лучевой нагрев исходного материала получил самое широкое применение в промышленных установках. Это связано с тем, что он позволяет поднять температуру поверхности испарения до 3500°C . Это позволяет испарить практически все классы веществ. Луч можно сканировать по поверхности испаряемого вещества, что позволяет улучшить равномерность покрытия и увеличить запасы распыляемого материала в установке. Электронно-лучевой разогрев позволяет проводить испарение из различных источников. Стабильность характеристик всей системы обеспечивает получение высококачественных пленок. Конструкция электронных пушек и тигли могут быть различными, однако из них можно испарять: Ta, Cu, W, Mo, Nb, Ni, Ti, Fe, Al_2O_3 , TiO_2 и другие материалы.

Весьма перспективным способом термического напыления является лазерный метод, использующий лучистый нагрев вещества в камере от источника энергии, вынесенного из камеры. Высокая энергия излучения, отсутствие источника тепла в камере позволяет получать статические неоднородности практически из любого материала с сохранением стехиометрии состава исходного вещества.

В методе *катодного распыления* используются диодные, триодные, дуговые и магнетронные источники для получения сильно ионизированных потоков испаряемого вещества, а также электрические и магнитные поля для управления этими потоками в испарительной камере. Энергия частиц в потоке лежит в пределах $0,5\text{--}20\text{ эВ}$, плотность потока — в пределах $10^{17}\text{ см}^{-2}\cdot\text{с}^{-1}$, а доля ионизированных частиц доходит до 100%. С помощью диодной и триодной системы, а также триодной системы с продольным магнитным полем можно испарять практически все материалы.

Особое внимание специалистов привлекает магнетронное распыление, основанное на локализации аномально тлеющего разряда в разреженном газе со скрещенными электрическими и магнитными полями. Магнетронные системы позволяют:

- ☐ достичь высоких скоростей формирования пленочных структур;
- ☐ уменьшить уровень газовых включений в формируемую неоднородность;
- ☐ формировать структуры более 10 мкм с сохранением адгезии и однородности;
- ☐ обеспечить стабильность параметров;
- ☐ нанести пленки окислов с металлических мишеней без использования радиочастотных источников электропитания системы.

К недостаткам магнетронной системы следует отнести низкие значения коэффициента использования материала.

К катодному способу относится и вакуумная дуга, представляющая собой сильнооточный разряд в парах материала катода. В результате эрозии катода формируется поток испаряемого материала. Метод отличается высокой производительностью.

Основным недостатком является наличие капельной фракции в потоке, которую можно уменьшить при работе в области коротких импульсов.

Взрывной метод позволяет формировать поток с энергией частиц до 7000 эВ с плотностью потока частиц до $10^{20} \text{ см}^{-2} \cdot \text{с}^{-1}$ и долей ионизации частиц 30%. Этот метод пока не нашел широкого применения, хотя и позволяет получить структуры из металлов, полупроводниковых соединений и диэлектриков. В этом методе особое место занимает лазерный источник для импульсного испарения. Высокая плотность потока и энергии испаряемых частиц, максимальная температура испарения исходного материала позволяет исключить фракционные испарения легкоплавких и летучих компонентов сложных соединений и получить статические неоднородности заданного размера и стехиометрию состава. Метод лазерного импульсного испарения имеет высокую производительность.

Термоионное испарение характеризуется высоким значением энергии частиц в потоке (0,2—7000 эВ), высокой плотностью $10^{20} \text{ см}^{-2} \cdot \text{с}^{-1}$ и практически полной ионизацией потока. Однако этот метод пока не нашел широкого практического применения в промышленных установках.

Проведя анализ различных методов создания потоков и основных возможностей испарителей, используемых в этих методах, скажем несколько слов о структуре установки для получения пленочных статических неоднородностей.

Конструктивно эти установки весьма разнообразны и определяются производительностью, требованиями к вакуумной гигиене, качеством получаемых структур и их стоимостью и другими технико-экономическими параметрами.

В основе конструкции любой вакуумной напылительной установки лежит испарительная система. Все способы осаждения пленок в вакууме требуют определенных вакуумных условий, состава газовой среды, которые и обеспечиваются вакуумной системой. Для поддержания условий конденсации пленки и формирования потока частиц необходимо построить систему нагрева, контроля и регулирования температуры подложки, а также создания электрических и магнитных полей. Для контролируемой газовой среды необходимо создать в установке систему измерения состава газа, их дозированного напуска, предварительной очистки газов, формирования газовой смеси.

Контроль получаемых структур должен осуществляться системами, позволяющими оперативно измерять параметры, например, толщину, удельное сопротивление, спектральные характеристики и т. п.

5.6. Сборка интегральных схем

К основным процессам сборки интегральных схем относятся следующие технологические процессы:

- ☐ скрайбирование и разделение;
- ☐ монтаж кристаллов на основание корпуса;
- ☐ разварка выводов;
- ☐ герметизация корпусов;
- ☐ тестовые испытания;
- ☐ маркировка.

Скрайбирование производится либо диском, либо лазерным излучением. Наклеенные на мембрану пластины после скрайбирования ломаются на отдельные кристаллы. Межоперационный контроль позволяет определить пригодные к дальнейшему корпусированию кристалл. Негодные помечаются специальными метками и в монтаже на основании корпуса не участвуют.

Монтаж кристаллов на основание корпуса осуществляется методом приклейки, например эпоксидной смолой.

На посадочном месте кристаллы ИС могут быть закреплены и методом пайки легкоплавкими припоями.

Разварка выводов производится от контактных площадок интегральной схемы к стандартным выводам корпуса заданной конструкции. Один из наиболее распространенных методов — *термокомпрессия*. Этот метод соединения металлов с металлами микросваркой давлением с подогревом соединяемых деталей. Температура соединения при термокомпрессии не превышает температуру образования эвтектики соединяемых материалов. Термокомпрессия ограничена числом сочетаний свариваемых материалов.

Другим методом разварки является сварка давлением с косвенным импульсным нагревом. При этом способе сварки инструмент (пуансон) импульсно нагревается проходящим по нему током. Локальный прогрев позволяет приваривать проводники из относительно малопластичных металлов к тонким пленкам на керамических подложках.

В процессе ультразвуковой сварки происходит возбуждение в свариваемых деталях упругих колебаний и создания определенного давления. В этом способе сварки необходимые условия для образования соединения создаются в результате механических колебаний УЗ-преобразователя.

С помощью межоперационного контроля оценивается качество микросвязных соединений, в основу которых положены следующие признаки:

- ☐ прочность на отрыв при испытании соединений;
- ☐ степень деформации проводника в зоне соединений;
- ☐ электрическое сопротивление зоны сварного соединения;
- ☐ стойкость к термическому старению;
- ☐ стойкость к градиентам температур.

Герметизация интегральных схем может быть осуществлена либо путем их корпусирования, либо бескорпусной герметизацией.

Герметизация в корпусе предполагает предварительное изготовление корпусов определенной конструкции с основанием для крепления кристалла, изолированными выводами, крышкой и вспомогательными деталями.

В процессе бескорпусной герметизации кристалл с внешними выводами заливается полимерным компаундом.

В методе комбинированной герметизации предварительно изготавливают капсулы, в которые устанавливают сборку кристалла с выводами, а затем эту конструкцию заливают герметизирующим компаундом.

Вакуумно-плотную герметизацию обеспечивают металлические, стеклянные и керамические корпуса, детали которых соединяются пайкой или сваркой, включая лазерную сварку.

Тестовые испытания перед маркировкой предназначены для комплексного испытания интегральных схем. Измерения электрических параметров производится в статических и динамических режимах. В результате тестовых испытаний схемы разбраковываются по параметрам, допускающим определенное назначение. В дальнейшем они маркируются для применения в коммерческих и военных целях.

5.7. Типовые технологические маршруты производства интегральных схем

Технологический маршрут представляет собой строго выверенную, научно обоснованную последовательность технологических операций, направленных на формирование конструкции заданного типа интегральной схемы и обеспечения ее надежного функционирования.

Процессы формирования микросхем включают в себя большое число различных по своей физико-химической природе операций, таких как: окисление, травление, осаждение пленок, диффузии, литографии и т. д. Каждая из этих операций, как правило, сопровождается межоперационным контролем. Межоперационный контроль необходим для отбраковки негодных изделий на определенной стадии технологического маршрута. Это позволяет существенно повысить экономичность процесса, потому как бракованная продукция снимается с производства на любой стадии. Это позволяет не расходовать значительные средства на материалы, высококвалифицированный труд, использование дорогостоящего оборудования для дальнейшего производства заведомого брака.

Эффективность любого производства интегральных схем определяется процентом выхода годowych. Поэтому отбраковка на любой стадии технологического маршрута позволяет повысить показатель выхода годowych интегральных схем.

Несмотря на то, что технологические маршруты бывают типовыми, характерными для определенного типа интегральных схем, они в то же время индивидуальны для каждого предприятия. Это зависит от наличия технологической оснастки, используемых материалов, технологического оборудования, квалификации персонала и т. д.

Приведем несколько типовых технологических процессов, достаточно полно описанных профессором Л. А. Коледовым.

На рис. 5.28 приведен технологический маршрут производства микросхем на биполярных транзисторах по широко распространенной изопланарной технологии, относящейся к комбинированному типу изоляции элементов. Комбинированная изоляция сочетает технологичность изоляции р—n-переходом и высокое качество изоляции диэлектриком в приповерхностном слое и с боков.

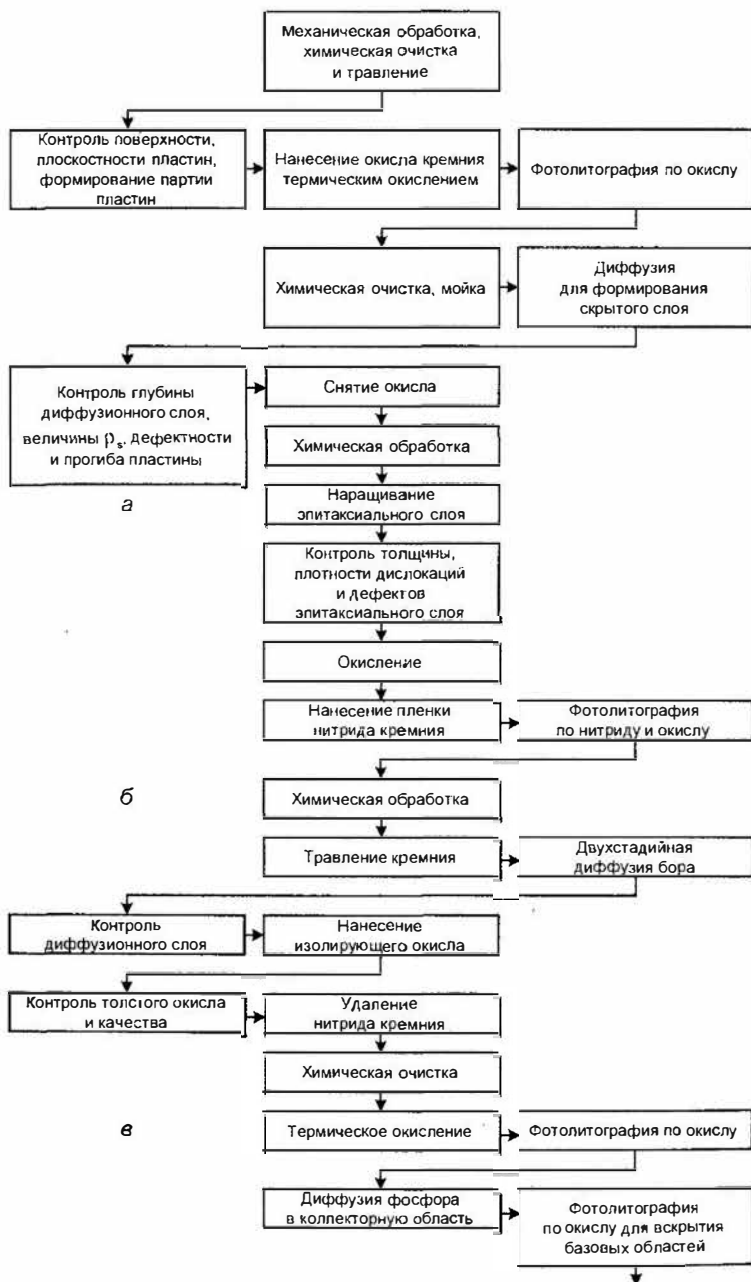


Рис. 5.28. Технологический маршрут производства микросхем на биполярных транзисторах по широко распространенной изопланарной технологии, относящейся к комбинированному типу изоляции элементов (см. продолжение)

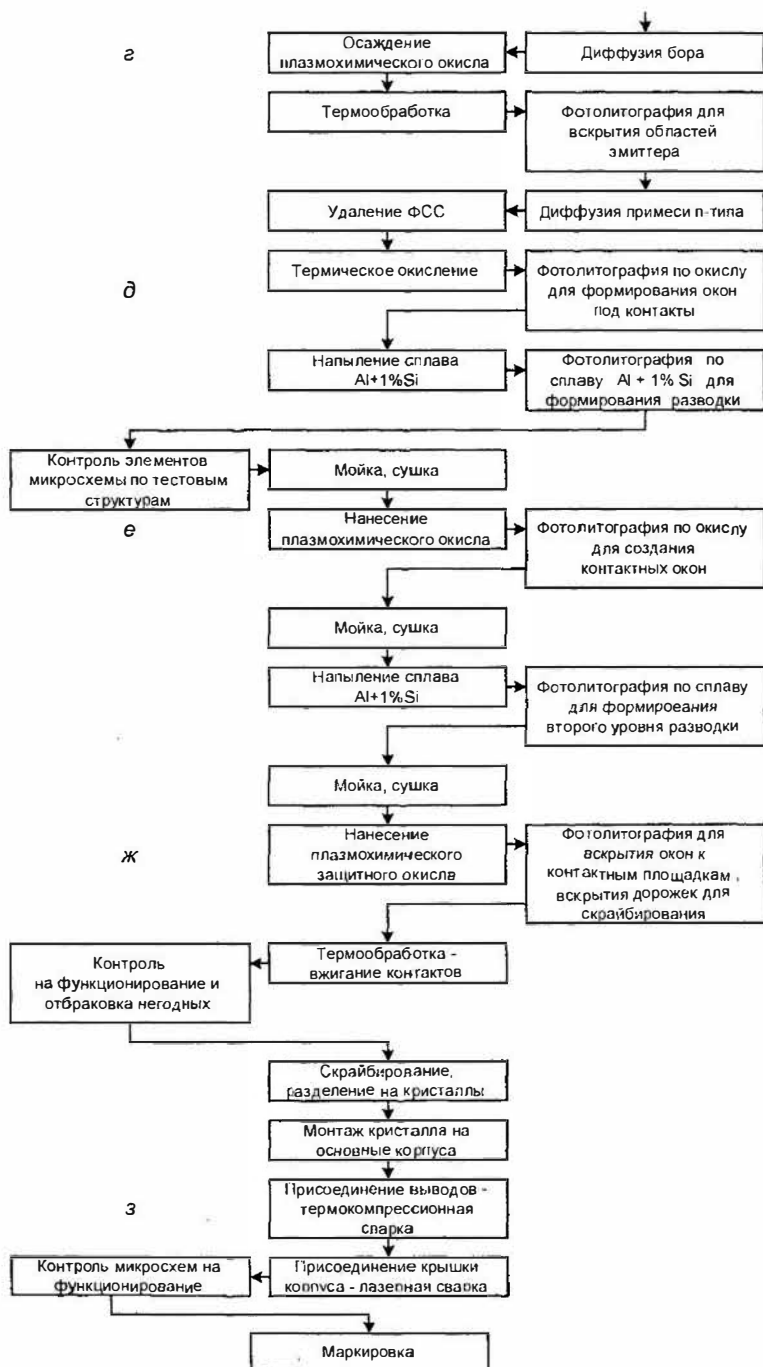


Рис. 5.28. Продолжение

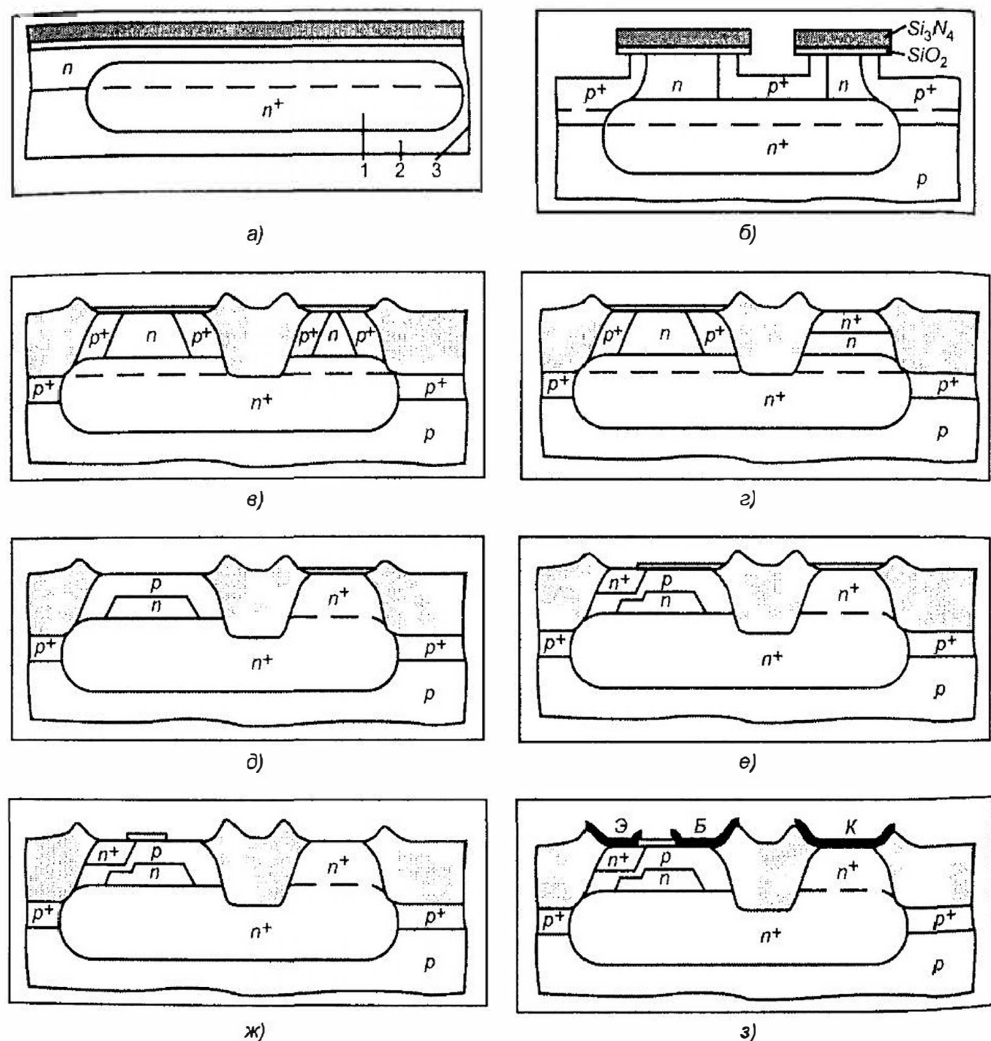


Рис. 5.28. Окончание

На рис. 5.28 введены все фотолитографические операции, операции диффузии и главные этапы межоперационного контроля в процессе производства.

Исходная пластина p-Si со скрытым n⁺-слоем создается в процессе I фотолитографии и диффузии для формирования скрытого слоя. Затем на подложке наращивается тонкий эпитаксиальный n-слой (рис. 5.28, а).

Далее производится контроль диффузионного слоя, а затем наносится изолирующий окисел и маскирующая пленка нитрида кремния.

Затем проводят II фотолитографию, травление эпитаксиального слоя кремния, диффузию акцепторной примеси для формирования p-слоев, которые образуют дно и стенки канавок. В результате этих операций получается структура, представленная на рис. 5.28, б.

В рамках межоперационного контроля производится контроль диффузионного слоя.

Полученную структуру окисляют в парах воды с тем, чтобы образующийся в канавках изолирующий слой SiO_2 достиг уровня подложки. Образуются карманы, которые с боков изолированы слоем окисла Si_3N_4 , который защищает кремний от окисления в парах воды (рис. 5.28, в). Далее формируется тонкий слой окисла.

В результате процесса III фотолитографии и последующей диффузии фосфора формируется коллекторная область транзистора $\text{n}^+\text{-Si}$ (рис. 5.28, г).

В процессе IV фотолитографии по окислу, предназначенной для вскрытия базовых областей, и последующей диффузией бора формируется базовая область транзистора p-Si (рис. 5.28, д). После ряда процессов осаждения фосфорсиликатного стекла (ФСС), окисла и термообработки производится V фотолитография. Она предназначена для вскрытия области эмиттера. Далее осуществляется диффузия соответствующей примеси n -типа (фосфора) и формируется эмиттерная область транзистора (рис. 5.28, е).

Процесс VI фотолитографии по окислу предназначен для формирования окна под металлизацию. На поверхность кристалла напыляется сплав на основе алюминия и производится VII фотолитография. Эта литография производится непосредственно по металлу с целью формирования электрической разводки интегральной схемы.

Межоперационным контролем производится испытание встроенных тестовых структур. С целью создания контактов к областям транзистора производится VIII фотолитография по окислу (рис. 5.28, ж). После процесса напыления алюминиевого сплава проводится IX фотолитография. С ее помощью формируется второй уровень разводки, и наносится пленка сплава Al (рис. 5.28, з). И наконец, X фотолитография предназначена для вскрытия окон к контактным площадкам, вскрытия дорожек для скрайбирования пластины.

По линии межоперационного контроля производится отбраковка негодных микросхем. Заключительные операции связаны с оформлением отдельных кристаллов, вырезаемых из целой пластины, в промышленную интегральную схему (рис. 5.28).

В табл. 5.4 приведен пример типового маршрута изготовления интегральных схем, где описаны конкретные параметры отдельных технологических процессов.

Таблица 5.4

№ п/п	Операция и параметры
1	Двухстадийная химическая обработка пластин в перекисно-аммиачном растворе
2	Окисление кремния во влажном кислороде при 1000°C в течение 2 час. до получения окисла толщиной $(0,6 \pm 0,006)$ мкм
3	Фотолитография под локальную диффузию сурьмы
4	Химическая обработка пластин в перекисно-аммиачном растворе
5	Диффузия сурьмы для формирования скрытого n^+ -слоя в две стадии: загонка при 1220°C до $\rho_s = (45 \pm 5)$ Ом/см ² из растворимых источников диффузии; обработка осажденного сурьмяно-силикатного стекла во влажном кислороде при 1000°C ; снятие стекла и окисла в растворе HF ; вторая стадия (разгонка) при 1200°C до $\rho_s = (35 \pm 5)$ Ом/см ² и глубины залегания p-n -перехода $(3,5 \pm 0,5)$ мкм
6	Снятие окисла в растворе HF
7	Химическая обработка пластин в перекисно-аммиачном растворе

Таблица 5.4 (продолжение)

№ п/п	Операция и параметры
8	Эпитаксиальное наращивание монокристаллического слоя кремния n-типа из газовой смеси $\text{HCl} + \text{H}_2$ (хлоридный метод) при 1100°C , толщиной $(2,0 \pm 0,2)$ мкм, с плотностью дефектов не более $5 \times 10^2 \text{ см}^{-2}$, легированного мышьяком
9	Окисление поверхности эпитаксиального слоя при 1000°C в течение 40 мин. в сухом кислороде для получения окисла толщиной (60 ± 10) нм
10	Осаждение нитрида кремния из газовой фазы при 1050°C до толщины (180 ± 20) нм с пористостью не более 10 см^{-2}
11	Фотография для вскрытия окон в пленках нитрида и окисла кремния для формирования разделительных областей
12	Локальное плазмохимическое травление пленок нитрида и окисла кремния до поверхности раздела "кремний--окисел кремния"
13	Химическая обработка пластин в перекисно-аммиачном растворе
14	Травление кремния в травителе $\text{HF}:\text{HNO}_3:\text{H}_2\text{O} = 1:19:8$ на глубину $(1,0 \pm 0,1)$ мкм
15	Двухстадийная диффузия бора: осаждение на поверхности пластины боросиликатного стекла из газовой фазы, содержащей BBr_3 и O_2 при 950°C с одновременным формированием диффузионного слоя с $p_s = 50 \text{ Ом/см}^2$, обработка боросиликатного стекла во влажном кислороде при 600°C в течение 30 мин., снятие боросиликатного стекла в травителе $\text{HF}:\text{H}_2\text{O} = 1:10$, отжиг структуры (разгонка бора) при 1050°C в течение 30 мин. в атмосфере кислорода до получения $p_s = (110 \pm 15) \text{ Ом/см}^2$, $x_b = (0,9 \pm 0,1)$ мкм
16	Локальное окисление канавок 1000°C в парах воды в течение 18 час. до получения толщины окисла $2,45$ мкм с точностью $\pm 5\%$
17	Удаление Si_3N_4
18	Химическая обработка пластин в перекисно-аммиачном растворе с использованием колонковых или беличьих кистей
19	Термическое окисление структур при 1050°C в сухом (10 мин.), влажном (20 мин.) и снова в сухом (10 мин.) кислороде
20	Фотолитография для вскрытия окон над коллекторными областями, совмещение фотошаблона не более 2 мкм
21	Диффузия примеси p-типа для формирования коллекторной области транзистора (первая стадия) — фосфоросиликатного стекла из газовой стадии $\text{PCl}_3 + \text{O}_2$ в течение 30 мин. при 960°C до $p_s = (10 \pm 1) \text{ Ом/см}^2$
22	Фотолитография для снятия окисла с поверхности базовых областей
23	Снятие окисла с областей базы путем травления в $\text{HF}:\text{H}_2\text{O} = 1:10$ (время травления определяется по спутнику, но не более 9 мин.)
24	Освежение поверхности базовых областей (5 с, если время хранения между операциями более 4 час.)
25	Диффузия бора в две стадии для формирования базовой области при 875°C в течение 15 мин. до $p_s = 180\text{—}190 \text{ Ом/см}^2$ и $x_b = 0,5$ мкм; обработка боросиликатного стекла во влажном кислороде при 600°C (30 мин.); снятие боросиликатного стекла; отжиг при 1050°C в сухом (15 мин.), влажном (15 мин.) и снова в сухом (5 мин.) кислороде до $p_s = 700\text{—}800 \text{ Ом/см}^2$ и $x_b = 0,5\text{—}0,6$ мкм. Одновременно формируется на базовых областях окисел толщиной $0,18\text{—}0,20$ мкм и проводится разгонка (вторая стадия) примеси в коллекторной области до $p_s = (40 \pm 4) \text{ Ом/см}^2$

Таблица 5.4 (продолжение)

№ п/п	Операция и параметры
26	Осаждение фосфоросиликатного стекла (ФСС) для уменьшения величины встроенного заряда в окисле из газовой фазы, содержащей PCl_3 и O_2 при $900^\circ C$ в течение 15 мин.
27	Пассивация структуры микросхемы: осаждение плазмохимического окисла кремния толщиной $0,1 \text{ мкм}$ при $300^\circ C$
28	Термообработка при $970^\circ C$ в течение 60 мин. в окислительной (O_2) атмосфере
29	Фотолитография под область эмиттера; время травления определяется по спутнику, но не более 4 мин., клин травления не более $0,3 \text{ мкм}$; размер эмиттера — 6 мкм , а точность совмещения фотошаблона не более $1,8 \text{ мкм}$
30	Освежение поверхности (если время хранения не более 4 час.)
31	Диффузия примеси n-типа для формирования эмиттерной области, осаждение ФСС при температуре $960^\circ C$ в течение (5 ± 1) мин.
32	Удаление ФСС
33	Фотолитография для вскрытия контактных окон в SiO_2 к областям транзистора
34	Освежение (если время хранения более 4 час.)
35	Напыление пленки сплава $Al+(1\%)Si$ толщиной $(0,60 \pm 0,1) \text{ мкм}$, температура подложки $200^\circ C$, температура отжига $250^\circ C$
36	Фотолитография по сплаву $Al+(1\%)Si$ для формирования разводки; клин травления не более $1,5 \text{ мкм}$, уход размеров не более 2 мкм
37	Контроль ВАХ по тестовым структурам $B_{con} \geq 30$, $R_g = 1,3 - 2 \text{ кОм}$, $U_{збo} \geq 5 \text{ В}$, $U_{кзo} \geq 5 \text{ В}$, $U_{кдо} \geq 10 \text{ В}$, пробивное напряжение база — подложка $U_{бп} \geq 10 \text{ В}$; отбраковка дефектных пластин
38	Отмывка с применением колонковых или беличьих кистей, сушка
39	Осаждение изолирующего слоя плазмохимического окисла при температуре $150^\circ C$ толщиной $(0,5 \pm 0,05) \text{ мкм}$
40	Фотолитография для вскрытия окон к нижнему уровню разводки
41	Осаждение второго слоя изолирующего плазмохимического окисла при $150^\circ C$ до суммарной толщины изоляции $(1,0 \pm 0,1) \text{ мкм}$
42	Фотолитография для вскрытия переходных окон между двумя уровнями разводки
43	Напыление пленки сплава $Al + (1\%)Si$ толщиной $(1,0 \pm 0,2) \text{ мкм}$ (остальные режимы аналогичны режимам операции 35) для создания второго уровня разводки
44	Фотолитография для формирования второго уровня разводки (клин травления $1,5 - 3,5 \text{ мкм}$, уход размеров $2,5 - 4,0 \text{ мкм}$)
45	Отмывка с применением колонковых или беличьих кистей, сушка
46	Осаждение защитной пленки плазмохимического окисла SiO_2 толщиной $0,30 - 0,55 \text{ мкм}$
47	Термообработка для формирования надежных контактов между уровнями разводки и вжигание алюминиевых контактов в кремний
48	Фотолитография для вскрытия окон к контактным площадкам в защитном окисле и вскрытие в окисле дорожек для скрайбирования

Таблица 5.4 (окончание)

№ п/п	Операция и параметры
49	Скрайбирование пластин для разделения их на кристаллы. Далее следуют операции контроля и разбраковки микросхем по электрическим параметрам и контроль на функционирование на еще не разделенных на кристаллы пластинах (на негодные кристаллы ставится метка краской) с использованием высокопроизводительного контрольного оборудования, в состав которого входит специализированная ЭВМ. Затем производится разделение (ломка) пластин на кристаллы без потери их взаимной ориентировки, после чего кристаллы идут на операции монтажа и сборки в корпус

МДП-интегральные схемы изготавливают, как правило, на кремниевых подложках с ориентацией рабочей поверхности по плоскости (100). Причиной тому является низкая плотность поверхностных состояний. Концентрация легирующих примесей в подложке составляет $N \sim 10^{15} \text{ см}^{-3}$.

Высокая плотность размещения элементов ИС позволяет применение изопланарного процесса (рис. 5.29).

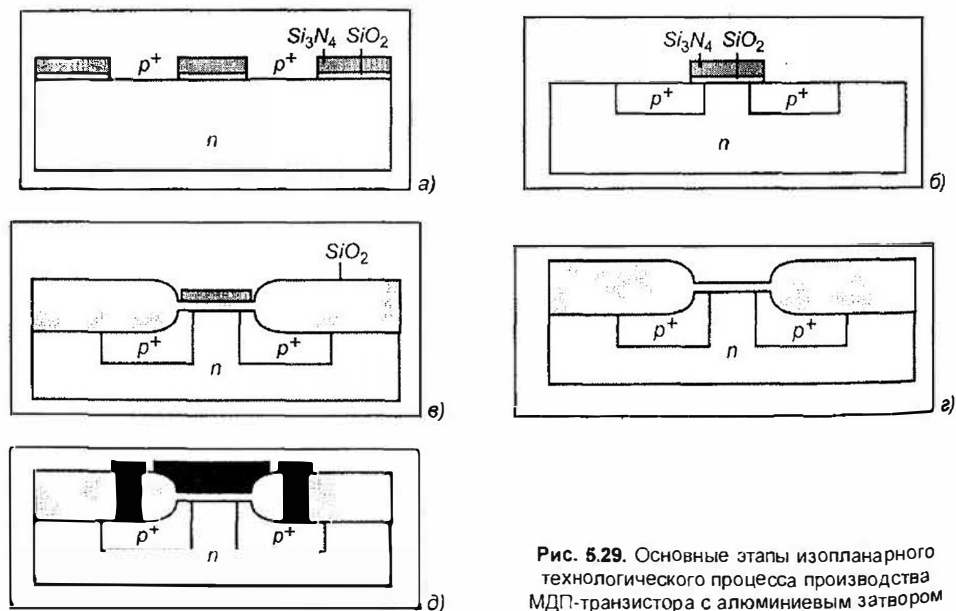


Рис. 5.29. Основные этапы изопланарного технологического процесса производства МДП-транзистора с алюминиевым затвором

На пластину кремния n -типа наносится пленка окисла и нитрида кремния для формирования окон под ионное или диффузионное легирование областей истока и стока (рис. 5.29, а). Последующие операции связаны с формированием областей истока и стока и одновременным удалением отработанных пленок окисла и нитрида кремния за исключением области канала (рис. 5.29, б).

Для защиты областей истока и стока происходит наращивание толстого окисла (рис. 5.29, в). После удаления пленки окисла и нитрида кремния, которые защищали об-

ласть канала, наносится тонкий слой подзатворного окисла. В этом случае используется метод термического окисления в сухом кислороде (рис. 5.29, з).

После вскрытия окон под контакты к областям истока и стока наносится алюминиевая металлизация к истоку и стоку, и, таким образом, формируется алюминиевый затвор (рис. 5.29, д).

Более перспективной является технология производства микросхем на МДП-транзисторах с использованием поликремниевых затворов. Этот процесс позволяет получать структуры с самосовмещенными затворами, снизить пороговое напряжение, уменьшить геометрические размеры, а следовательно, увеличить степень интеграции. Такой процесс позволяет создавать на одной подложке как МДП-, так и биполярные транзисторы. БИМОП-технология производства интегральных микросхем является весьма перспективной.

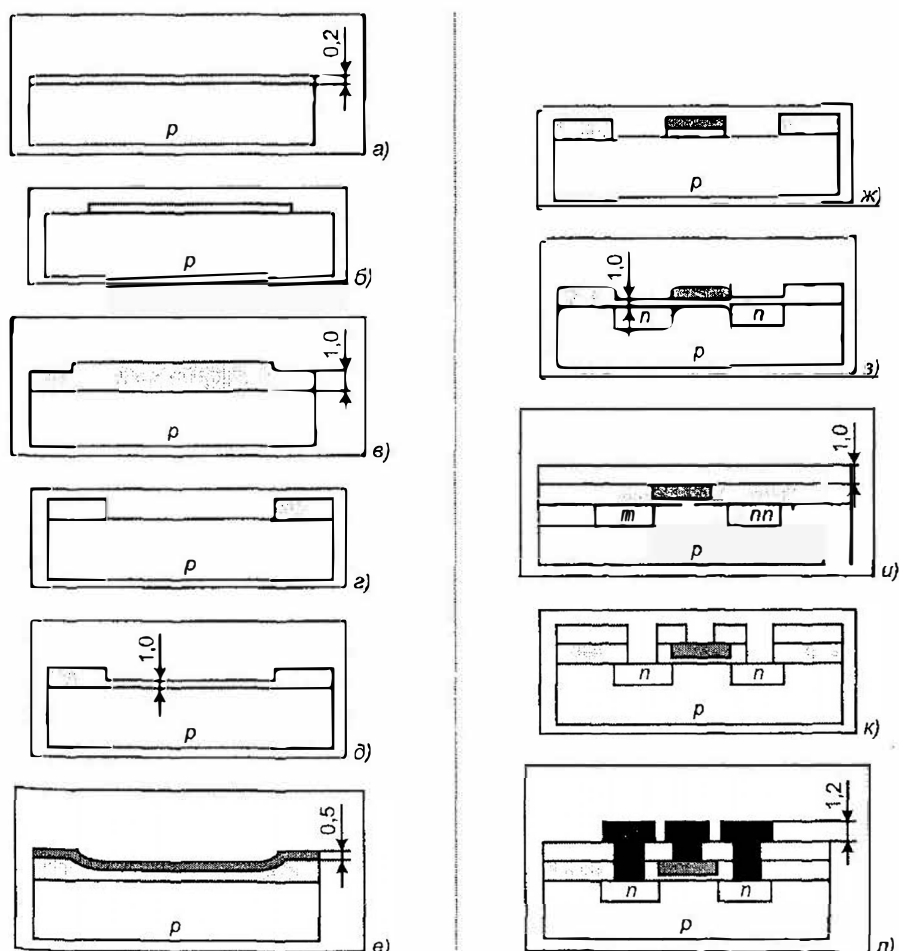


Рис. 5.30. Основные этапы технологического процесса производства МДП интегральных схем с поликремниевыми затворами

Использование в качестве затворов МДП-транзисторов поликристаллического кремния вместо алюминия позволяет использовать его в качестве маскирующих пленок и обеспечить самосовмещение затворов с областями истока и стока. На рис. 5.30 приведены этапы технологического процесса производства интегральных схем с поликремниевыми затворами.

На первом этапе происходит окисление кремниевой пластины (рис. 5.30, а). Фотолитографией формируют локальные области будущих транзисторных структур (рис. 5.30, б).

Для исключения формирования инверсионных паразитных транзисторов проводят процесс имплантации базы и снова окисляют поверхность (рис. 5.30, в).

На следующем этапе проводится фотолитография с целью вскрытия окон под области транзистора, а также диффузионных мин (рис. 5.30, г).

Подзатворный диэлектрик формируется путем окисления кремния в сухом кислороде (рис. 5.30, д). Толщина окисла составляет примерно 0,1 мкм.

На окисную пленку из газовой фазы наносится поликристаллический кремний, имеющий высокую температуру плавления. Толщина поликремниевого слоя составляет порядка 0,5 мкм (рис. 5.30, е).

Далее методом фотолитографии формируется кремниевый затвор. Одновременно методом высокотемпературной диффузии затвор легируется фосфором с целью обеспечения его электронной проводимости и снижения поверхностного сопротивления (рис. 5.30, з).

Поверхность этой структуры окисляется и сверху наносится слой фосфорно-силикатного стекла (рис. 5.30, и).

Далее производится фотолитография для вскрытия окон под контакты к областям истока и стока, поликремниевому затвору и шинам (рис. 5.30, к).

После напыления пленок алюминия производится последняя фотолитография разводки электродов транзистора и контактных площадок (рис. 5.30, л). В табл. 5.5 приведен технологический маршрут изготовления МДП-схем с поликремниевыми затворами и охранными кольцами.

Таблица 5.5

№ п/п	Операция и параметры
1	Формирование партии пластины кремния КЭФ-4,5 с ориентацией (100)
2	Кистевая мойка (0,05% раствор синтанола)
3	Химическая очистка в растворителе $\text{H}_2\text{SO}_4 + \text{H}_2\text{O}_2 + \text{NH}_4\text{OH}$
4	Термическое окисление $h(\text{SiO}_2) = 0,72 \pm 0,05$
5	Кистевая мойка в воде и инфракрасная (ИК) сушка
6	Нанесение фоторезиста и ИК-сушка
7	Первое совмещение и экспонирование
8	Проявление фоторезиста и ИК-сушка
9	Контроль копии
10	Плазмохимическое травление (защитка), 30—60 с

Таблица 5.5 (продолжение)

№ п/п	Операция и параметры
11	"Задубливание" фоторезиста
12	Травление SiO_2 (контрольный образец)
13	Травление SiO_2 (партия пластин)
14	Снятие фоторезиста с помощью растворителя $\text{H}_2\text{SO}_4 + \text{H}_2\text{O}_2$ (раствор КАРО)
15	100%-й контроль чистоты поверхности
16	10%-й контроль травления
17	Химическая очистка в растворе КАРО + H_2O_2 + NH_4OH
18	Ионное легирование бором (формирование p-кармана)
19	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
20	Разгонка бора в p-кармане
21	Стравливание SiO_2
22	Кистевая мойка (0,05% раствор синтанолоа)
23	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
24	Термическое окисление ($h(\text{SiO}_2) = 0,05$ мкм)
25	Кистевая мойка, если интервал между операциями 24 и 26 составляет более 1 часа
26	Осаждение нитрида кремния $h(\text{Si}_3\text{N}_4) = 0,2$ мкм
27	Контроль качества нитрида кремния
28	Кистевая мойка (вода)
29	100%-й контроль нитрида кремния
30	Кистевая мойка (вода) и ИК-сушка
31	Нанесение фоторезиста и ИК-сушка
32	Второе совмещение и экспонирование
33	Проявление фоторезиста и ИК-сушка
34	Контроль копии
35	Плазмохимическое травление нитрида кремния
36	100%-й контроль травления
37	Снятие фоторезиста (раствор КАРО)
38	100%-й контроль чистоты
39	10%-й контроль травления
40	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
41	Ионное легирование фосфором (формирование охранных колец p-канальных транзисторов)
42	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
43	Кистевая мойка (вода) и ИК-сушка
44	Нанесение фоторезиста и ИК-сушка

Таблица 5.5 (продолжение)

№ п/п	Операция и параметры
45	Третье совмещение и экспонирование
46	Проявление фоторезиста и ИК-сушка
47	Контроль копии
48	Ионное легирование бором (формирование охранных колец p-канальных транзисторов)
49	Снятие фоторезиста (раствор КАРО)
50	100%-й контроль чистоты
51	Освежение 1 мин. в буферном растворе
52	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
53	Локальное окисление
54	Контроль деформации (стрелы прогиба) пластины
55	Нанесение фоторезиста и ИК-сушка
56	"Задубливание" фоторезиста
57	Травление обратной стороны пластины
58	Снятие фоторезиста (раствор КАРО)
59	100%-й контроль чистоты
60	Освежение в буферном растворе 2,5 мин.
61	Стравливание нитрида кремния
62	100%-й контроль травления
63	Травление подслоя SiO_2
64	Контроль толщины толстого фонового слоя ($h \geq 1,0$ мкм)
65	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
66	Предварительное окисление $h(\text{SiO}_2) = (0,1 \pm 0,01)$ мкм
67	Травление предварительно нанесенного оксида (2 мин.)
68	Контроль толщины толстого фонового оксида ($h(\text{SiO}_2) \geq 0,9$ мкм)
69	Химическая очистка (КАРО + H_2O_2 + NH_4OH)
70	Контроль чистоты
71	Нанесение подзатворного оксида термическим окислением
72	Освежение, если интервал между операциями 71 и 73 более 7 час. (7 с. в буферном растворе $\text{HF} + \text{NH}_4\text{OH} + \text{H}_2\text{O}$)
73	Выращивание пленки поликремния (ПК) ($d_{\text{ПК}} = (0,6 \pm 0,05)$ мкм)
74	100%-й визуальный контроль ПК
75	20%-й контроль ПК под микроскопом
76	Диффузионное легирование ПК фосфором
77	Снятие фосфоросиликатного стекла, плазмохимическое травление

Таблица 5.5 (продолжение)

№ п/п	Операция и параметры
78	Кистевая мойка в воде и ИК-сушка
79	Нанесение фоторезиста и ИК-сушка
80	Четвертое совмещение и экспонирование
81	Проявление фоторезиста и ИК-сушка
82	Контроль копии
83	Плазмохимическое травление ПК
84	100%-й контроль травления
85	Снятие фоторезиста (раствор КАРО)
86	100%-й контроль чистоты
87	10%-й контроль травления
88	Контроль ОТК
89	Освежение пластин в течение 10 с. в буферном растворе
90	Химическая очистка ($\text{КАРО} + \text{H}_2\text{O}_2 + \text{NH}_4\text{OH}$)
91	Кистевая мойка в воде и ИК-сушка
92	Нанесение фоторезиста и ИК-сушка
93	Пятое совмещение и экспонирование
94	Проявление фоторезиста и ИК-сушка
95	Контроль копии
96	Ионное легирование фосфором (формирование стоков и p-канальных транзисторов)
97	Снятие фоторезиста плазмохимическим травлением
98	Снятие фоторезиста (раствор КАРО)
99	100%-й контроль чистоты
100	Освежение (10 с.) в буферном растворе
101	Химическая очистка ($\text{КАРО} + \text{H}_2\text{O}_2 + \text{NH}_4\text{OH}$)
102	Разгонка фосфора
103	Контроль параметров p^+ -слоев
104	Кистевая мойка в воде и ИК-сушка
105	Нанесение фоторезиста и ИК-сушка
106	Шестое совмещение и экспонирование
107	Проявление фоторезиста и ИК-сушка
108	Контроль копии
109	Ионное легирование бором (формирование стоков и истоков p-канальных транзисторов)
110	Снятие фоторезиста плазмохимическим травлением
111	Снятие фоторезиста (раствор КАРО)

Таблица 5.5 (продолжение)

№ п/п	Операция и параметры
112	100%-й контроль чистоты
113	Освещение (10 с.) в буферном растворе
114	Химическая очистка ($\text{KAP}O + \text{H}_2\text{O}_2 + \text{NH}_4\text{OH}$)
115	Окисление и разгонка после внедрения бора
116	Контроль результатов легирования на тестовых структурах
117	Кистевая мойка в воде и ИК-сушка
118	Нанесение фоторезиста и ИК-сушка
119	Седьмое совмещение и экспонирование
120	Проявление фоторезиста и ИК-сушка
121	Контроль копии
122	Плазмохимическое травление (зачистка)
123	"Задубливание" фоторезиста
124	Травление SiO_2 (контрольная пластина)
125	Травление SiO_2 всей партии
126	Снятие фоторезиста (раствор $\text{KAP}O$)
127	100%-й контроль чистоты
128	100%-й контроль травления
129	Химическая очистка (только состав $\text{KAP}O$)
130	Освещение в буферном растворе в течение 7 с.
131	Контроль толщины оксида ($h \approx 0,25$ мкм)
132	Контроль чистоты
133	Напыление сплава $\text{Al}+(1\%)\text{Si}$
134	Контроль качества алюминиевой пленки
135	Нанесение фоторезиста и ИК-сушка
136	Восьмое совмещение и экспонирование
137	Проявление фоторезиста и ИК-сушка
138	Контроль копии
139	Плазмохимическое травление (зачистка)
140	"Задубливание" фоторезиста
141	Травление алюминия (контрольный образец)
142	Травление алюминия в партии пластин
143	Снятие фоторезиста (только раствор MOHO)
144	100%-й контроль чистоты
145	100%-й контроль травления

Таблица 5.5 (окончание)

№ п/п	Операция и параметры
146	Осаждение фосфоросиликатного стекла
147	Контроль пленки
148	Нанесение фоторезиста и ИК-сушка
149	Девятое совмещение и экспонирование
150	Травление фоторезиста и ИК-сушка
151	Контроль копии
152	Плазмохимическое травление (зачистка)
153	"Задубливание" фоторезиста
154	Травление фосфоросиликатного стекла (контрольная пластина)
155	Травление фосфоросиликатного стекла во всей партии пластин
156	100%-й контроль травления
157	Снятие фоторезиста (только раствор MOHO)
158	100%-й контроль травления
159	Вжигание алюминиевых контактов
160	Контроль порогового напряжения и напряжения пробоя основных и паразитных транзисторов n- и p-типа
161	100%- контроль функционирования
162	Контроль внешнего вида
163	Контроль ОТК

Технологических маршрутов производства кремниевых интегральных схем достаточно много и останавливаться на каждом нет возможности.

Особый интерес вызывают интегральные схемы на основе арсенида галлия. Именно с ними связывают высокое быстродействие и радиационную стойкость. Существуют проекты различных технологических процессов:

- ионно-имплантированных полевых транзисторов с барьером Шоттки;
- полевых транзисторов с управляющим p-n-переходом;
- гетероструктурных транзисторов с высокой подвижностью электронов;
- гетепереходных биполярных транзисторов и т. п.

На рис. 5.31 приведены основные этапы изготовления полевого транзистора с барьером Шоттки на основе ионной имплантации. В одном технологическом процессе показано производство транзисторных и диодных структур.

На первом этапе производится осаждение изолятора Si_3N_4 и маскируются пластины фоторезистивной маски (рис. 5.31, а).

После формирования затвора осуществляют процесс новой имплантации p^+ -примеси с целью формирования высоколегированных областей истока и стока (рис. 5.31, б).

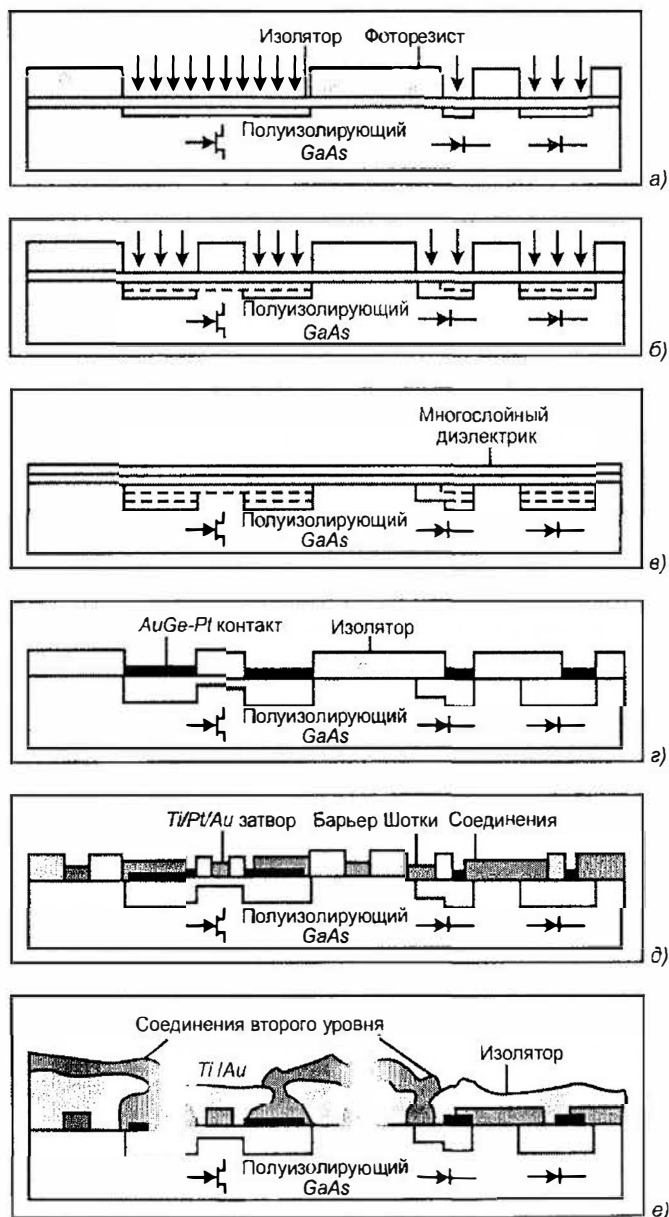


Рис. 5.31. Основные этапы изготовления ПТШ на основе ионной имплантации

На следующем этапе технологического процесса происходит осаждение дополнительного слоя защитного изолятора (рис. 5.31, в).

Металлизация омических контактов производится после очередной фотолитографии (рис. 5.31, г). Затем происходит вжигание омических контактов.

На следующем этапе происходит металлизация барьера Шоттки и остальных соединений. Одновременно производится предварительное тестирование (рис. 5.31, д).

Далее наносится второй слой изолятора, после чего посредством фотолитографии вскрываются окна, и проводится металлизация второго уровня арсенид-галлиевых интегральных микросхем (рис. 5.31, е).

Для отжига радиационных нарушений, возникающих при ионной имплантации, применяют длительный или импульсный нагрев.

Металл второго уровня обычно наносится распылением в вакууме, а рисунок на нем методом ионного травления.

5.8. Особенности производства интегральных схем

Характерной особенностью производства интегральных схем является повторяемость некоторых технологических операций: операции литографии, окисления, очистки, легирования. Поэтому подложки несколько раз проходят одни и те же технологические участки.

С повышением степени интеграции возрастает плотность упаковки элементов, а также площадь, занимаемая на кристалле отдельной интегральной схемой. Поражение дефектом хотя бы одного элемента интегральной схемы приводит к выходу из строя всей интегральной схемы.

Среди предпринимаемых мер для увеличения выхода годных одно из первых мест занимают чистые среды на участках производства.

Главным источником загрязнения атмосферы, в которой производятся интегральные схемы, является человек. С поверхности его тела ежеминутно выделяется порядок 10^6 различных частиц: ороговевшие частицы, капельки пота, бактерии и т. д.

Обычно на поверхности пластин обнаруживают кремниевую пыль, обычную атмосферную пыль, частицы абразива, волокна, бактерии. В период цветения обнаруживается пыльца растений.

На рис. 5.32 приведены относительные размеры частиц, обнаруживаемых в чистой комнате и линия топологии интегральной схемы.

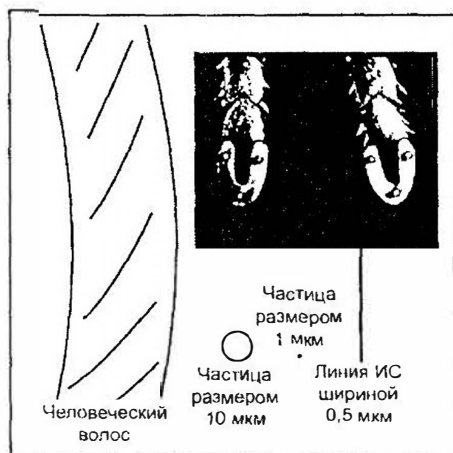


Рис. 5.32. Сравнительные размеры загрязняющих частиц, топологических линий в производстве интегральных схем и чуда микротехники — подкованной блохи. Размер шляпки гвоздика-ухналика составляет 5 мкм

В связи с этим производство интегральных схем развивалось как путем создания производственных участков в чистых комнатах, так и по пути полной автоматизации, без участия человека непосредственно на участках производства. Это исключает субъективные ошибки при проведении прецизионных технологических операций. С появлением субмикронных интегральных схем появилось так называемое кластерное производство.

Полупроводниковое производство интегральных схем отличается принципом групповой технологии. Это означает, что одновременно обрабатывается не только множество интегральных схем, расположенных на одной пластине, но и множество пластин. Это позволило резко поднять процесс выхода годных, и снизить стоимость одной интегральной схемы, существенно улучшить экономические и энергетические параметры интегральных схем.

Задачи и упражнения

5.1. Расчет профилей распределения примесей при высокотемпературной диффузии

Определить режимы окисления, загонки и разгонки примесей (бор или фосфор) при изготовлении биполярного транзистора со структурой $p^+ - p - n$ на кремнии, для которой заданы следующие параметры:

- ☐ глубина залегания эмиттерного $p^+ - p$ -перехода (d_e);
- ☐ глубина залегания коллекторного $p - n$ -перехода (d_k);
- ☐ поверхностная концентрация при эмиттерной диффузии (N_0^e);
- ☐ поверхностная концентрация при коллекторной диффузии (N_0^k);
- ☐ концентрация примеси в исходном кремнии ($N_{исх}$);
- ☐ толщина окисла для маскирования ($d_{ок}$).

Решение

Биполярные транзисторы являются наиболее распространенным типом полупроводниковых приборов. Он является основным элементом одного из типов полупроводниковых интегральных схем. На технологии изготовления его элементов базируется изготовление остальных составляющих этих интегральных схем — диодов, резисторов и конденсаторов. При этом за основу выбран именно $p^+ - p - n$ -транзистор.

Во всех перечисленных случаях независимо от топологии маршрут изготовления $p^+ - p - n$ -транзистора одинаков: сначала создается область базы и при этом формируется коллекторный $p - n$ -переход, который является $p - n$ -переходом между базой и коллектором, а затем — область эмиттера и при этом формируется эмиттерный $p - n$ -переход, который является $p - n$ -переходом между эмиттером и базой.

Области базы и эмиттера, а соответственно, и коллекторный и эмиттерный $p - n$ -переходы могут создаваться методами диффузии или ионной имплантации. Диффузией создаются более глубокие, а ионной имплантацией — более мелкие $p - n$ -переходы. Поэтому имплантационные транзисторы имеют более тонкую базу, что обеспечивает возможность их работы при больших частотах по сравнению с диффузионными аналогами.

Формирование базовой и эмиттерной областей осуществляется за счет локальной диффузии, или ионной имплантации, при которых маской, формирующей площадь легированных областей, является окисная пленка, выращенная предварительным окислением в at°

мосфере кислорода — сначала сухого, потом влажного, а затем опять сухого. Для сухого кислорода толщина окисной пленки $d_{ок}$ зависит от времени и температуры следующим образом:

$$d_{ок}^2 = 21,2t \cdot e^{-\frac{1,33(\text{эВ})}{kT}} \quad (5.1.1)$$

Для влажного кислорода эта зависимость определяется как:

$$d_{ок}^2 = 7,26t \cdot e^{-\frac{0,8(\text{эВ})}{kT}} \quad (5.1.2)$$

В выражениях (5.1.1) и (5.1.2) единицы измерения параметров следующие: $d_{ок}$ — мкм, t — минуты, постоянная Больцмана k — эВ/К, температура T — К.

При ионной имплантации может оказаться, что окисной маски недостаточно для защиты поверхности от проникающих энергетичных ионов, и тогда поверх нее наносится еще одна пленка — пленка алюминия толщиной $\sim 1 \div 2$ мкм.

Идеальные профили распределения примесных атомов в транзисторной структуре представлены на рис. 5.33, а реальные диффузионные профили — на рис. 5.34.

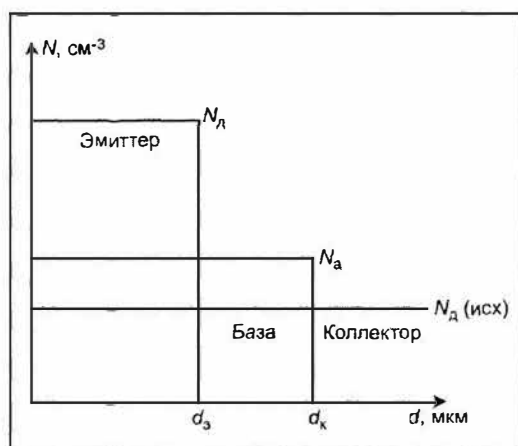


Рис. 5.33. Идеальные профили распределения донорных (N_D) и акцепторных (N_A) примесных атомов в p — n — p -биполярном транзисторе

"Затягивания" реальных профилей на рис. 5.34 относительно идеальных профилей обусловлены особенностями процесса диффузии, которые описываются законами Фика:

□ 1-й закон Фика: $J = -D \frac{\partial N}{\partial x}$,

□ 2-й закон Фика: $\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2}$,

где J — плотность потока примесных атомов, $\text{см}^{-2} \cdot \text{с}^{-1}$; D — коэффициент диффузии, $\text{см}^2/\text{с}$; N — концентрация диффундирующих атомов, см^{-3} ; x — координата вглубь кристалла.

Решение 2-го уравнения Фика позволяет получить профиль концентрации примесных атомов.

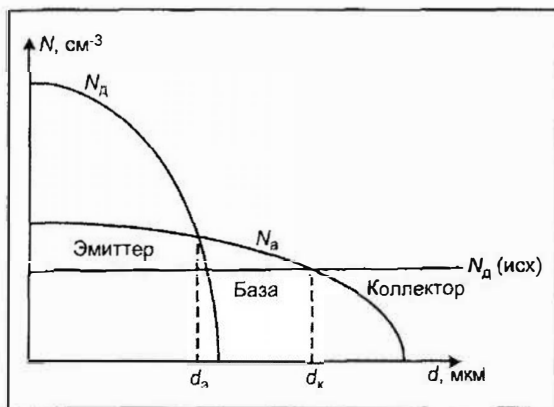


Рис. 5.34. Реальные профили распределения донорных (N_d) и акцепторных (N_a) атомов в диффузионном биполярном транзисторе

Решение может быть выполнено для двух случаев:

1. При неограниченном источнике примеси, когда поверхностная концентрация примесных атомов постоянна, т. е. $N_0 = \text{const}$ (такой случай реализуется в процессе, который называется загонкой).
2. При ограниченном источнике примеси, когда N_0 зависит от времени, т. е. $N_0(t)$ и обеспечивается только атомами, введенными при загонке (такой случай реализуется в процессе, который называется разгонкой).

При загонке профиль концентрации вводимых в кристалл атомов описывается выражением:

$$N(x) = N_0 \cdot \operatorname{erfc} \frac{x}{2\sqrt{Dt}}. \quad (5.1.3)$$

При этом доза легирования, т. е. количество примесных атомов (см^{-2}), введенных в кристалл через площадку в 1 см^2 , выражается как:

$$Q = 1,13 N_0 \sqrt{Dt}. \quad (5.1.4)$$

При разгонке профиль примесных атомов имеет следующий вид:

$$N = N_0(t) e^{-\left(\frac{x}{2\sqrt{Dt}}\right)^2} = \frac{Q}{\sqrt{\pi Dt}} e^{-\left(\frac{x}{2\sqrt{Dt}}\right)^2}. \quad (5.1.5)$$

Как видно из уравнений (5.1.3) и (5.1.5), глубина проникновения примесных атомов в кристалл при загонке и разгонке определяется фактором Dt , т. е. длительностью процесса t и температурой, поскольку коэффициент диффузии D имеет сильную экспоненциальную температурную зависимость:

$$D = D_0 \cdot e^{-\frac{\Delta E_a}{kT}},$$

где ΔE_a — энергия активации диффузии.

На практике сначала производится загонка, а потом — разгонка примесных атомов. При этом следует иметь в виду, что при многостадийном (n этапов) процессе диффузии, когда проводится последовательная загонка и разгонка нескольких видов примесных атомов, примесь, введенная на первых этапах диффузионных процессов, продолжает разгоняться

на следующих этапах. Это учитывается за счет суммирования факторов $D_i t_i$, которые характерны для рассматриваемой примеси на всех n этапах ($\sum_1^n D_i t_i$). Именно эта величина определяет результирующий профиль распределения данной примеси.

Полученные путем расчета профили распределения базовой и эмиттерной примеси позволяют определить глубину залегания коллекторного и эмиттерного р—п-переходов. При этом глубина коллекторного р—п-перехода (d_k) определяется как координата точки пересечения профиля базовой примеси — акцепторных атомов, например, бора в случае п—р—п-транзистора, с профилем примесных атомов, имеющихся в исходном кристалле (N_{acc}). Глубина эмиттерного р—п-перехода (d_e) определяется как координата точки пересечения профилей эмиттерной (донорной) и базовой примесей. При этом каждый р—п-переход — и коллекторный, и эмиттерный формируется в результате двухэтапного процесса — сначала загонки, а затем разгонки соответствующих примесных атомов.

При формировании базовой и эмиттерной областей биполярного транзистора методом ионной имплантации профиль распределения имплантированных акцепторных и донорных атомов рассчитывается по формуле:

$$N(x) = \frac{Q}{\Delta R_p \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x - R_p}{\Delta R_p} \right)^2}, \quad (5.1.6)$$

где Q — доза имплантации, R_p — средний проективный пробег иона, ΔR_p — девиация пробега.

Следует иметь в виду, что при использовании ионов с высокой энергией (выше $80 \div 100$ кэВ) у поверхности эмиттерной области может образоваться слабелегированная область, которая пагубно сказывается на характеристиках транзистора, увеличивая шумовые токи и напряжения из-за высокого контактного сопротивления, которое она образует с металлическим выводом эмиттера. В этом случае следует провести дополнительное легирование этой области ионами с меньшей энергией ($30 \div 40$ кэВ), т. е. провести легирование ионами с распределенной энергией: сначала E_1 при Q_1 , а затем E_2 при Q_2 , где $E_1 > E_2$ и $Q_1 > Q_2$.

Алгоритм выполнения задания сводится к следующему.

Сначала определяют режимы: температуру и длительность процесса окисления, обеспечивающего заданную толщину окисной пленки. При этом считают, что процесс включает три стадии:

- ☐ окисление в сухом кислороде;
- ☐ окисление во влажном кислороде;
- ☐ окисление в сухом кислороде.

При этом все три стадии проводятся при одной и той же температуре. Определяют температуру и длительность всех трех стадий, считая, что на каждой стадии сухого окисления толщина образующейся окисной пленки должна быть одинаковой и составлять ~ 15% от заданной толщины маскирующей пленки. Таким образом, за две стадии сухого окисления толщина окисной пленки составит ~ 30%, а за одну стадию влажного окисления — ~ 70% от заданной толщины. Используя выражения (5.1.1) для сухого и (5.2.2) для влажного окисления, подбирают температуру таким образом, чтобы для определенных на предыдущем этапе толщин окисла, формируемых при сухом и влажном окислениях, длитель-

ность окисления в сухом кислороде на каждой стадии была в пределах $10 \div 30$ мин., а во влажном кислороде — $1 \div 4$ часа. Подбор следует начинать с температуры 1150°C , которая обычно используется на практике. При этом величина k в выражениях (5.1.1) и (5.1.2) равна $8,07 \times 10^{-5}$ эВ/К.

Затем вычисляют эффективный фактор $D'_{p\delta}t'_{p\delta}$, определяющий результирующий профиль распределения базовой примеси — бора. Этот фактор включает в себя факторы, описывающие дополнительные разгонки бора, происходящие при последующей загонке и разгонке эмиттерной примеси — фосфора. Фактор $D'_{p\delta}t'_{p\delta}$ определяют следующим образом. Этап разгонки бора рассматривают на основе выражения (5.1.5) для двух значений x : $x_1 = 0$ и $x_2 = d_k$. При $x_1 = 0$ выражение (5.1.5) приобретает вид:

$$N = N_0^K = \frac{Q_\delta}{\sqrt{\pi D'_{p\delta}t'_{p\delta}}}, \quad (5.1.7)$$

а при $x_2 = d_k$:

$$N_{исх} = \frac{Q_\delta}{\sqrt{\pi D'_{p\delta}t'_{p\delta}}} e^{-\left(\frac{d_k}{2\sqrt{D'_{p\delta}t'_{p\delta}}}\right)^2}, \quad (5.1.8)$$

где Q_δ — доза базовой примеси, $D'_{p\delta}$ и $t'_{p\delta}$ — эффективные коэффициент диффузии базовой примеси и длительность ее разгонки. Значения $N_{исх}$, N_0^K и d_k даны в задании.

Таким образом, получаем систему уравнений (5.1.7) и (5.1.8), деление которых одно на другое приводит к выражению:

$$\frac{N_0^K}{N_{исх}} = e^{\frac{d_k^2}{4D'_{p\delta}t'_{p\delta}}},$$

откуда следует, что

$$D'_{p\delta}t'_{p\delta} = \frac{d_k^2}{4 \ln \frac{N_0^K}{N_{исх}}}. \quad (5.1.9)$$

Подставляя в формулу (5.1.9) данные в задании значения d_k , N_0^K и $N_{исх}$, можно найти величину произведения $D'_{p\delta}t'_{p\delta}$.

Далее исходя из формулы (5.1.7) определяется величина дозы Q_δ , после чего по формуле (5.1.5) рассчитывают и строят профиль распределения атомов бора. Подставив в формулу (5.1.5) $x = d_j$ (величина d_j дана в задании) находят значение $N'_{исх}$, соответствующее концентрации бора на границе эмиттерного р—п-перехода. Это значение потребуется в последующем для определения фактора Dt при эмиттерной диффузии по формуле (5.1.9) для подстановки вместо $N_{исх}$.

Режим загонки базовой примеси (бора) — температуры ($T_{ис}$) и длительности ($t_{ис}$) процесса — рассчитывается следующим образом. Задаваясь температурой загонки в пределах $800 \div 1100^\circ\text{C}$, из данных табл. 5.6 находят предел растворимости атомов бора, соответствующий выбранной температуре.

Таблица 5.6. Предельные растворимости атомов бора и фосфора в кремнии (см^{-3}) для различных температур

T, °C	800	900	1000	1050
бор В	8×10^{16}	1.2×10^{18}	10^{19}	6×10^{19}
фосфор Р	10^{18}	3×10^{18}	2×10^{19}	5×10^{20}

При этом считают, что найденный предел растворимости равен поверхностной концентрации N_0 , величину которой вместе с найденным выше значением Q_0 подставляют в формулу (5.1.4), откуда вычисляют фактор Dt при загонке бора, т. е. $D_{\text{б}} t_{\text{зб}}$. Зная температуру загонки $T_{\text{зб}}$, из данных табл. 5.6 находят соответствующее ей значение $D_{\text{б}}$, а затем вычисляют по формуле

$$t_{\text{зб}} = \frac{D_{\text{зб}} \cdot t_{\text{зб}}}{D_{\text{б}}}$$

Таким образом, режимы загонки базовой примеси (бора) определены.

Для определения режимов загонки и разгонки эмиттерной примеси (фосфора) сначала, как и в случае базовой примеси, рассматривают этап разгонки. При этом определяют фактор

$$D_{\text{р}} t_{\text{р}} = \frac{d_0^2}{4 \ln \frac{N_0}{N_{\text{псх}}}};$$

где $D_{\text{р}}$ и $t_{\text{р}}$ — коэффициент диффузии и длительность разгонки фосфора, а величина N_0^2 определена начальными условиями. Далее, задаваясь величиной $t_{\text{р}}$ (в секундах) в пределах 1 ÷ 4 часов, из значения рассчитанного фактора $D_{\text{р}} t_{\text{р}}$ находят величину $D_{\text{р}}$, на основе которой с использованием данных табл. 5.6 для фосфора определяют соответствующую ей температуру $T_{\text{р}}$.

Варьируя значениями $t_{\text{р}}$ и $D_{\text{р}}$, можно подобрать величину $T_{\text{р}}$ так, чтобы она была "округлена" до десятков градусов, например 1110°, 1120°, 1130°C. В результате получают значения $T_{\text{р}}$ и $t_{\text{р}}$, т. е. температуру и длительность стадии разгонки фосфора при формировании эмиттера. Далее, в соответствии с формулой (5.1.7), определяют дозу эмиттерной примеси Q , и по формуле (5.1.5) рассчитывают и строят профиль распределения атомов фосфора. Режимы загонки фосфора определяют в полной аналогии с тем, как это показано для загонки бора.

Определение реального режима разгонки атомов бора, который вместе с дополнительной разгонкой, обусловленной последующими нагревами при загонке и разгонке атомов фосфора, формирует уже вычисленный результирующий профиль распределения атомов бора.

Для этого из данных табл. 5.7 находят значения коэффициентов диффузии бора при температурах загонки ($T_{\text{зб}}$) и разгонки ($T_{\text{рб}}$) фосфора: $D_{\text{б}}^{\text{зб}}$ и $D_{\text{б}}^{\text{рб}}$, и далее рассчитывают соответствующие им факторы: $D_{\text{зб}}^{\text{бп}} \cdot t_{\text{зб}}$ и $D_{\text{рб}}^{\text{бп}} \cdot t_{\text{рб}}$.

После этого определяется реальный фактор при разгонке бора ($D_{\text{рб}} \cdot t_{\text{рб}}$), который получается вычитанием только что найденных факторов из эффективного, ранее определенного:

$$D_{\text{рб}} \cdot t_{\text{рб}} = D_{\text{рб}}' \cdot t_{\text{рб}}' - D_{\text{зб}}^{\text{бп}} \cdot t_{\text{зб}} - D_{\text{рз}}^{\text{бп}} \cdot t_{\text{рз}}.$$

Таблица 5.7. Значения коэффициентов диффузии D ($\text{см}^2/\text{с}$) атомов бора и фосфора для различных температур

Т, °C	1000	1050	1100	1150	1200
бор В	10^{-14}	5×10^{-14}	10^{-13}	3×10^{-13}	10^{-12}
фосфор Р	10^{-13}	3×10^{-13}	10^{-12}	3×10^{-12}	5×10^{-12}

Далее аналогично тому, как это делалось при рассмотрении разгонки эмиттерной примеси (фосфора), из величины фактора $D_{\text{рб}} \cdot t_{\text{рб}}$ определяются режимы разгонки базовой примеси (бора) $T_{\text{рб}}$ и $t_{\text{рб}}$.

ПРИМЕЧАНИЯ:

1. Профили распределения атомов бора и фосфора строят на одном рисунке: по оси ординат обозначают концентрацию в логарифмическом масштабе, где $20 \div 30$ мм соответствуют изменению концентрации на порядок величины, а по оси абсцисс — глубину в линейном масштабе. Концентрацию обозначают в пределах $10^{14} \div 10^{21} \text{ см}^{-3}$; расчет профили производят также в этом пределе концентраций. На графике указывают координаты эмиттерного (d_j) и коллекторного (d_k) р—п-переходов, как это показано на рис. 5.34.
2. Все найденные режимы выписывают в конце работы в том порядке, в каком производятся соответствующие процессы.

Расчет необходимо произвести в программе Mathcad 2001.

5.2. Расчет профилей распределения примесей при ионной имплантации

Определить режимы имплантации: энергии ионов E , дозы имплантации Φ и длительности процессов t для ионов бора и фосфора при формировании биполярного транзистора со структурой $\text{p}^+-\text{p}-\text{n}$ на кремнии, для которой заданы следующие параметры:

- ☐ глубина залегания эмиттера (d_j);
- ☐ глубина залегания коллектора (d_k);
- ☐ средняя концентрация в эмиттере ($N_{\text{ср}}^j$);
- ☐ средняя концентрация в базе ($N_{\text{ср}}^b$);
- ☐ концентрация примеси в исходном материале ($N_{\text{исх}}$).

Примите, что коэффициент использования примесных атомов после отжига составляет для бора $k_{\text{бора}} = 0,8$, а для фосфора $k_{\text{фосфора}} = 0,98$. Плотность ионного тока во всех случаях составляет 1 мкА/см^2 .

Решение

Для выполнения этой задачи — нахождения энергий и доз имплантации ионов бора и фосфора — в первую очередь рассчитывают дозы легирования базовой и эмиттерной областей исходя из заданных глубин залегания коллектора и эмиттера, и средних концентраций бора и фосфора в этих областях. Расчет производят с учетом величин коэффициентов использования при последующем отжиге по формулам:

$$Q_{\text{бора}} = N_{\text{ср}}^b \cdot \frac{d_k}{k_{\text{бора}}} \quad (\text{см}^{-2}),$$

$$Q_{\text{фосфор}} = N_{\text{ср}} \cdot \frac{d_2}{k_{\text{фосфор}}} (\text{см}^{-2}).$$

Длительность процессов имплантации бора (t_b) и фосфора ($t_{\text{ф}}$) определяют по формулам:

$$t_b = \frac{Q_{\text{бор}}}{j_{\text{бор}}},$$

$$t_{\text{ф}} = \frac{Q_{\text{фосфор}}}{j_{\text{фосфор}}},$$

где j — плотность тока соответствующих ионов.

Примем $j_{\text{бор}} = j_{\text{фосфор}} = 1 \text{ мкА/см}^2 = 6,2 \times 10^{12} \text{ ион/см}^2 \cdot \text{с}$ или $6,2 \times 10^{12} \text{ см}^{-2} \cdot \text{с}^{-1}$. Тогда, например, при дозе $1,24 \times 10^{15} \text{ см}^{-2}$ длительность процесса будет составлять

$$t = \frac{1,24 \cdot 10^{15}}{6,2 \cdot 10^{12}} = 200 \text{ с.}$$

Таким образом, дозы и длительности процессов определены.

Определение энергий ионов бора и фосфора производят с использованием данных табл. 5.8.

Таблица 5.8. Значения пробегов (R_p) и дисперсии пробегов (ΔR_p) в микронах для ионов бора и фосфора в кремнии

E, кэВ		30	100	300
бор B ⁺	R_p	0,187	0,527	1,190
	ΔR_p	0,045	0,087	0,122
фосфор P ⁺	R_p	0,050	0,155	0,454
	ΔR_p	0,012	0,038	0,075

Сначала по этим данным выбирают такую энергию ионов бора, которая соответствует значению $R_p + 5\Delta R_p \cong d_k$, а затем энергия уточняется подбором таким образом, чтобы при подстановке заданного значения d_k вместо x получалось значение $N(x) = N_{\text{исх}}$ с точностью $\pm 20\%$. После того как подобрана энергия, строится профиль распределения атомов бора.

Подбор энергии ионов фосфора производят аналогично, только за величину $N_{\text{исх}}$ считают концентрацию атомов бора на уже построенном профиле в точке, координата которой $x = d_k$.

ПРИМЕЧАНИЕ:

1. Как и в случае диффузионного транзистора, построение обоих профилей бора и фосфора производят на одном рисунке с обозначением координат эмиттерного и коллекторного p—n-переходов.
2. В работе следует представить данные по расчету профилей концентрации после диффузии и имплантации в виде таблиц. Количество расчетных точек на каждом участке профиля должно быть не менее 20 с равномерной разбивкой по координате.

3. Если при эмиттерной имплантации с энергией выше $80 \div 100$ кэВ у поверхности образуется область с концентрацией на $2 \div 3$ порядка величины ниже, чем в максимуме, следует подобрать режим еще одной имплантации с энергией $20 \div 40$ кэВ и такой дозой, чтобы концентрация фосфора на поверхности была $10^{19} \div 10^{20} \text{ см}^{-3}$.

Расчет необходимо произвести в программе Mathcad 2001.

Контрольные вопросы

1. Перечислите основные технологические процессы первичной обработки полупроводниковых материалов.
2. Что такие процессы литографии?
3. Что такое фотолитография? Какие у нее преимущества и недостатки?
4. Что такое электролитография? Какие у нее преимущества и недостатки?
5. Что такое ионолитография? Какие у нее преимущества и недостатки?
6. Что такое фотонаблон? Что такое рентгенонаблон?
7. Что такое рентгенолитография? Какие у нее преимущества и недостатки?
8. Что такое эпитаксия? Какие виды эпитаксии вы знаете?
9. Расскажите о методе молекулярно-лучевой эпитаксии.
10. Что такое процесс легирования полупроводников?
11. Какие законы, отражающие процесс высокотемпературной диффузии, вы знаете?
12. Что такое процесс ионной имплантации?
13. Какие функции выполняет пленка диоксида кремния на кремнии?
14. Что такое процесс травления, и какие функции он выполняет?
15. Как классифицируют процесс плазменного травления?
16. Какие процессы металлизации вы знаете? Охарактеризуйте каждый.
17. Перечислите основные технологические процессы сборки ИС. Охарактеризуйте каждый из них.
18. Охарактеризуйте основные этапы изопланарного технологического процесса МДП-транзисторов.
19. Охарактеризуйте основные этапы производства ПТНП-транзисторов на основе ионной имплантации.

Рекомендуемая литература

1. Ефимов И. Е., Козырь И. Я., Горбунов Ю. Я. Микроэлектроника. Физические и технологические основы, надежность: Учебное пособие. — 2-е изд. — М.: Высшая школа, 1986.
2. Коледов Л. А. Технология и конструкции микросхем, микропроцессоров и микросборок. Учебник для вузов. — М.: Радио и связь, 1989.
3. Конструирование и технология микросхем. Курсовое проектирование: Учебное пособие для вузов под ред. Л. А. Коледова. — М.: Высшая школа, 1984.
4. Березин А. С., Мочалкина О. Р. Технология и конструирование интегральных микросхем: Учебное пособие для вузов под ред. И. П. Степаненко. — М.: Радио и связь, 1983.
5. Лилич П., Николайдеас А. Задачи по физической электронике под ред. проф. Г. В. Скопского. — М.: Мир, 1975.

6. Интегральные схемы и их классификация

6.1. Классификация интегральных схем

Интегральная схема представляет собой конструктивно законченное изделие электронной техники, содержащее совокупность электрически связанных в функциональную схему транзисторов, диодов, конденсаторов, резисторов и других электрорадиоэлементов, изготовленных в едином технологическом цикле.

Интегральные схемы (ИС) являются основным продуктом микроэлектронного производства. ИС является элементной базой средств электронной техники, предназначенной для преобразования, обработки и хранения информации.

Условное графическое изображение интегральной схемы зависит от стандарта отдельных выпускающих ИС стран (рис. 6.1).

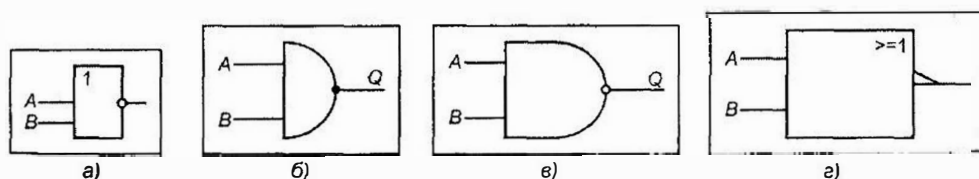


Рис. 6.1. Условное обозначение интегрального логического элемента типа ИЛИ — НЕ (NAND):
а — СССР—Россия; б — Германия (старое) DIN 40700old; в — США, US-ASA;
г — Германия (новое) DIN 40700new \ IEEE Std91

На рис. 6.2 показан общий вид отечественных интегральных микросхем.

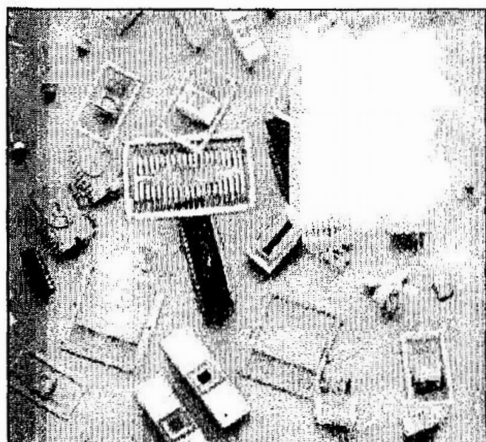


Рис. 6.2. Общий вид отечественных интегральных схем

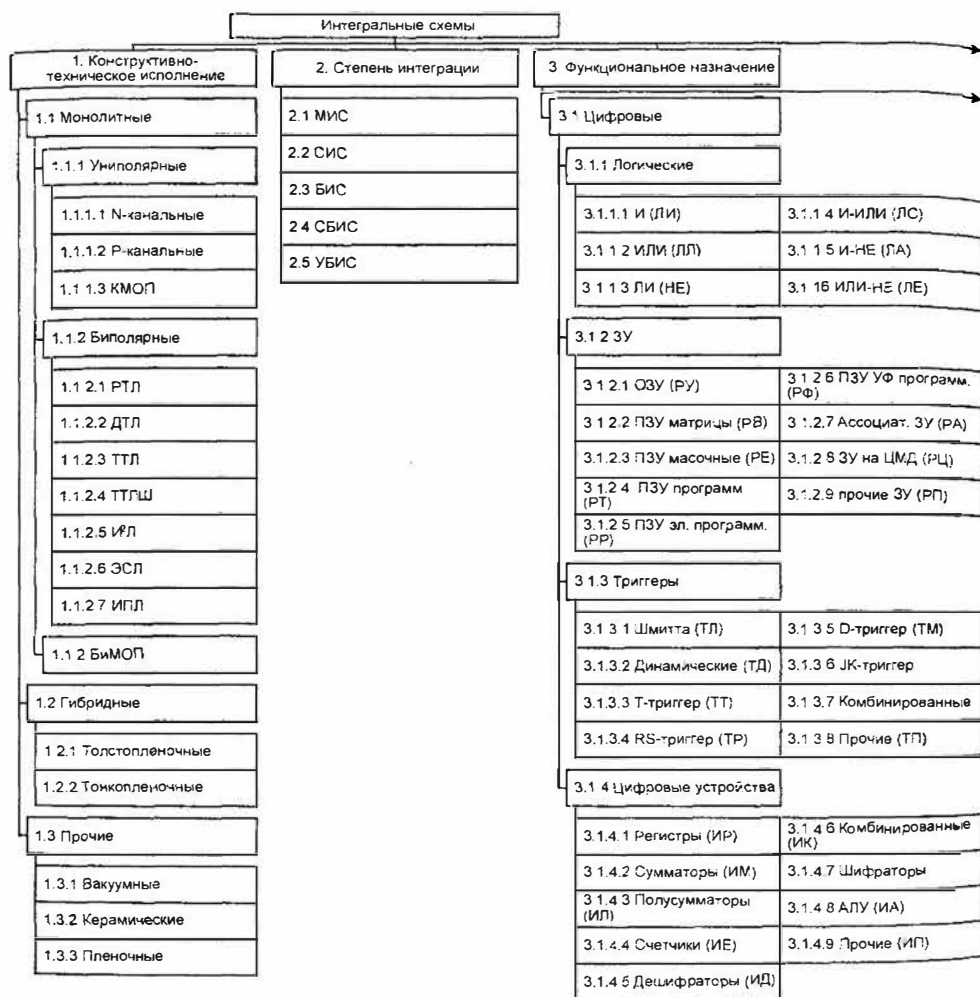


Рис. 6.3. Классификация интегральных схем (см. продолжение)

Интегральные схемы можно классифицировать по ряду независимых параметров. Общий подход к классификации интегральных схем представлен на рис. 6.3.

По конструктивно-технологическому исполнению различают три группы ИС:

- ☐ 1.1 — монолитные (полупроводниковые);
- ☐ 1.2 — гибридные;
- ☐ 1.3 — прочие (пленочные, вакуумные, керамические и т. д.).

3.2 Аналоговые

3.2.1 Генераторы

3.2.1.1 Гармонических (ГС)	3.2.1.3 Линейноменяющихся (ГЛ)	3.2.1.5 Спецформы (ГФ)
3.2.1.2 Прямоугольных (ГГ)	3.2.1.4 Шума (ГМ)	3.2.1.6 Прочие (ГП)

3.2.2 Усилители

3.2.2.1 Постоянного тока (УТ)	3.2.2.5 Промеж. частоты (УР)	3.2.2.9 Индикации (УМ)
3.2.2.2 Импульсные (УИ)	3.2.2.6 Низкочастотные (УН)	3.2.2.10 Операционные (УД)
3.2.2.3 Повторители (УЕ)	3.2.2.7 Широкополосные (УК)	3.2.2.11 Дифференциальные (УС)
3.2.2.4 Высокочастотные (УВ)	3.2.2.8 Считыв. и воспроизв. (ИЛ)	3.2.2.12 Прочие

3.2.3 Детекторы

3.2.3.1 Амплитудные (ДА)	3.2.3.3 Частотные (ДС)	3.2.3.5 Прочие (ДП)
3.2.3.2 Импульсные (ДИ)	3.2.3.4 Фазовые (ДФ)	

3.2.4 Задержки

3.2.4.1 Пассивные (БМ)	3.2.4.2 Активные (БР)	3.2.4.3 Прочие (БП)
------------------------	-----------------------	---------------------

3.2.4 Задержки

3.2.4.1 Пассивные (БМ)	3.2.4.2 Активные (БР)	3.2.4.3 Прочие (БП)
------------------------	-----------------------	---------------------

3.2.5 Селекции

3.2.5.1 Амплитудные (СА)	3.2.5.3 Частотные (СС)	3.2.5.5 Прочие (СП)
3.2.5.2 Временные (СВ)	3.2.5.4 Фазовые (СФ)	

3.2.6 Фильтры

3.2.6.1 Верхних частот (ФВ)	3.2.6.3 Полосовые (ФЕ)	3.2.6.5 Прочие (ФП)
3.2.6.2 Нижних частот (ФН)	3.2.6.4 Режекторные (ФР)	

3.2.7 Формирователи

3.2.7.1 Прямоугольной формы (АГ)	3.2.7.3 Адресных токов (АА)	3.2.7.5 Прочие (АП)
3.2.7.2 Специальной формы (АФ)	3.2.7.4 Разрядных токов (АР)	

3.2.8 Преобразователи

3.2.8.1 Частот (ПС)	3.2.8.5 Мощности (ПМ)	3.2.8.9 ЦАП (ПА)
3.2.8.2 Фазы (ПФ)	3.2.8.6 Уровня (ПУ)	3.2.8.10 АЦП (ПВ)
3.2.8.3 Длительности (ПД)	3.2.8.7 Синтезаторы частот (ПЛ)	3.2.8.11 КОД-КОД (ПР)
3.2.8.4 Напряжения (ПН)	3.2.8.8 Делители частот (ПЕ)	3.2.8.12 Прочие (ПП)

3.2.9 Модуляторы

3.2.9.1 Амплитудные (МА)	3.2.9.3 Фазовые (МФ)
3.2.9.2 Импульсные (МИ)	3.2.9.4 Прочие (МЛ)

Рис. 6.3. Продолжение

3.3 Вычислительные устройства	3.4 Источники вторичного питания	3.5 Многофункциональные устройства	3 / Базовые матричные кристаллы
3.3.1. Устройства обработки	3.4.1 Преобразователи (ЕМ)	3.5.1 Матрицы	3.6 Фоточувствительные ПЗС-устройства
3.3.1.1 Микро-ЭВМ (ВЕ)	3.4.2 Выпрямители (ЕВ)	3.5.1.1 Аналоговые (ХА)	3.6.1 Матричные (ЦМ)
3.3.1.2 Микропроцессоры (ВМ)	3.4.3 Стабилизаторы напряжения (ЕН)	3.5.1.2 Цифровые (ХП)	3.6.2 Линейные (ЦЛ)
3.3.1.3 Микрокалькуляторы (ВХ)	3.4.4 Стабилизаторы тока (ЕТ)	3.5.1.3 Комбинированные (ХК)	3.6.3 Прочие (ЦП)
3.3.1.4 Контроллеры (ВГ)	3.4.5 Устройства управления напряжением (ЕУ)	3.5.1.4 Цифровые матрицы (ХМ)	
3.3.1.5 Комбинированные устройства (ВК)	3.4.6 Источники вторичного питания (ЕС)	3.5.1.5 Аналоговые матрицы (ХИ)	
3.3.1.6 Специальные устройства (ВЖ)	3.4.7 Прочие (ЕП)	3.5.1.6 Комбинированные матрицы (ХТ)	
3.3.1.7 Прочие (ВП)		3.5.1.7 Прочие (ХИ)	
3.3.2 Микропроцессорные комплекты		3.5.2 Коммутаторы	
3.3.2.1 Микропроцессорные секции (ВС)		3.5.2.1 Тока (КТ)	
3.3.2.2 Устройство микропрограммного управления (ВУ)		3.5.2.2 Напряжения (КН)	
3.3.2.3 Функциональные расширители (ВР)		3.5.2.3 Прочие (КП)	
3.3.2.4 Устройство синхронизации (ВР)		3.5.3 Наборы элементов	
3.3.2.5 Устройство управления прерываниями (ВН)		3.5.3.1 Диодов (НД)	
3.3.2.6 Устройство ввода/вывода (ВВ)		3.5.3.2 Транзисторов (НТ)	
3.3.2.7 Устройство управления ЗУ (ВТ)		3.5.3.3 Резисторов (НР)	
3.3.2.8 Функция преобразования информации (ВФ)		3.5.3.4 Конденсаторов (НЕ)	
		3.5.3.5 Комбинированные (НК)	
		3.5.3.6 Функциональные (НФ)	
		3.5.3.7 Прочие (НП)	

Рис. 6.3. Продолжение

По отечественному стандарту ГОСТ 18682-73 им присвоены следующие индексы:

- 1, 5, 7 — монолитные;
- 2, 4, 6, 8 — гибридные;
- 3 — прочие.

Полупроводниковой (монолитной) интегральной схемой называют ИС, элементы которой выполнены в объеме и (или) на поверхности полупроводниковой подложки. Элементом интегральной схемы считается неотделимая составная часть ИС, выполняющая функцию какого-либо электрорадиоэлемента.

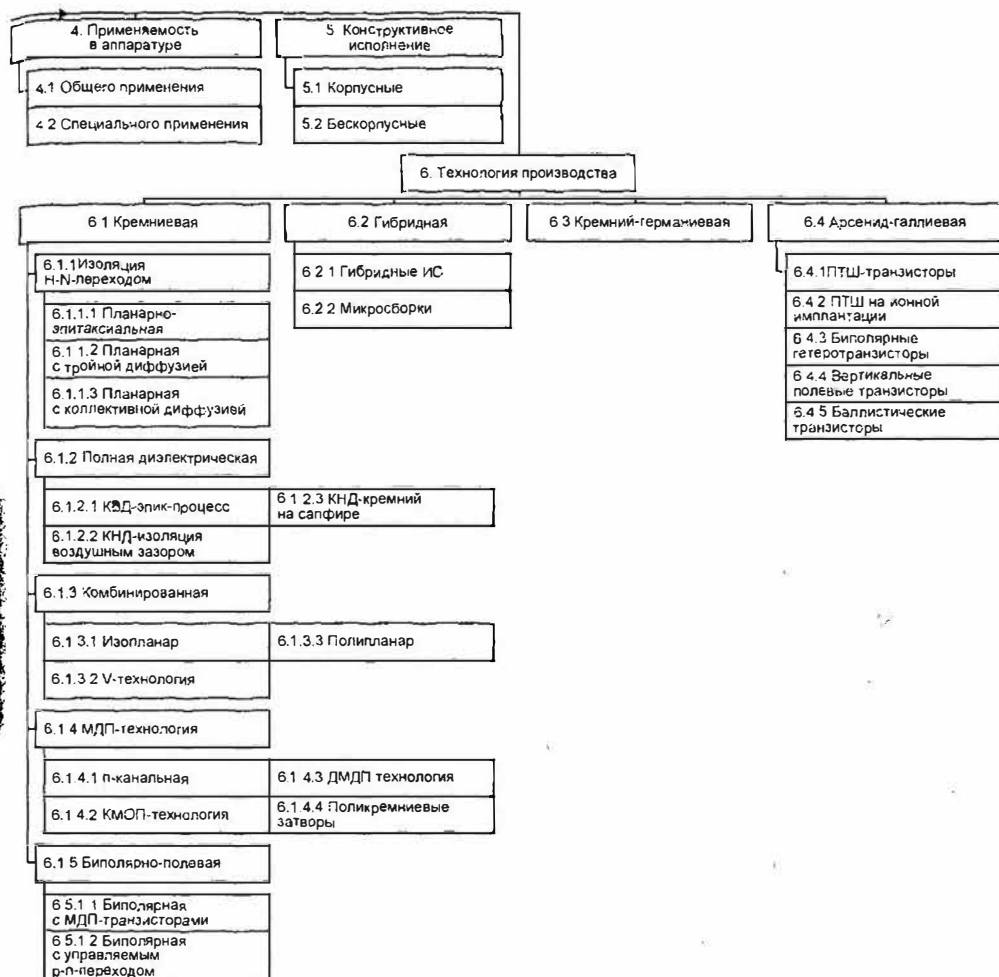


Рис. 6.3. Окончание

По типу используемых активных элементов монолитные ИС подразделяются на:

- ☐ 1.1.1 — униполярные (полевые, МДП или МОП);
- ☐ 1.1.2 — биполярные;
- ☐ 1.1.3 — комбинированные (биполярно-полевые).

МДП ИС в зависимости от типа интегральных структур подразделяются на:

- ☐ n-канальные (1.1.1.1);
- ☐ p-канальные (1.1.1.2);
- ☐ комплементарные или КМОП (1.1.1.3).

3 свою очередь, в зависимости от используемого типа логических структур и конструктивно-технологических решений *биполярные ИС* делят на:

- 1 резисторно-транзисторную логику (РТЛ) — 1.1.2.1;
- 2 диодно-транзисторную логику (ДТЛ) — 1.1.2.2;
- 3 транзисторно-транзисторную логику, в том числе с диодами Шоттки (ТТЛ и ТТЛШ) — 1.1.2.3 и 1.1.2.4, соответственно;
- 4 интегрально-инжекционную логику (I^2I) — 1.1.2.5;
- 5 эмиттерно-связанную логику (ЭСЛ) и ее разновидности — 1.1.2.6.

3 *комбинированных ИС* одновременно используются биполярные и полевые транзисторы, эта технология получила название *БиМОП* (1.1.3.1).

Гибридной интегральной схемой (1.2) называют ИС, в которой элементы и компоненты выполнены на диэлектрической подложке. Под компонентами понимаются миниатюрные навесные дискретные электрорадиоэлементы.

Пассивные элементы гибридных ИС могут быть выполнены на основе толстопленочной (1.2.1) (с толщиной пленки больше 1 мкм) или тонкопленочной (1.2.2) технологий, а активные компоненты выполняются по полупроводниковой технологии и монтируются на поверхность диэлектрической подложки.

3 прочих ИС (1.3) могут быть использованы:

- 1 вакуумные микродиодные и триодные структуры (ВИС — 1.3.1);
- 2 керамические элементы (1.3.2);
- 3 пленочные активные и пассивные элементы (1.3.3).

ИС этой группы промышленных образцов не имеют.

Следующим независимым признаком классификации является степень интеграции K . Значение K определяется как показатель степени числа элементов N в ИС:

$$K = \log N.$$

Малые интегральные схемы (МИС — 2.1) содержат до 100 элементов и компонентов на кристалле ($N \leq 2$).

Средние ИС (СИС — 2.2) содержат до 1000 элементов на кристалле ($N \leq 3$).

Большие интегральные схемы (БИС — 2.3) содержат до 10 000 элементов на кристалле ($N \leq 4$).

Верхнебольшие ИС (СВИС — 2.4) представляют собой законченное микроэлектронное устройство, способное выполнять функции аппаратуры и содержат до 1 000 000 элементов на кристалле ($N \leq 6$).

С *ультрабольшими ИС* (УБИС — 2.5) относят интегральные схемы со степенью интеграции $N \geq 6$.

По функциональному назначению интегральные схемы делятся на шесть основных классов.

Цифровые ИС (3.1) предназначены для обработки сигналов, заданных в виде дискретных функций. В свою очередь цифровые ИС делятся на:

- 1 логические ИС (3.1.1);
- 2 запоминающие устройства ЗУ (3.1.2);
- 3 триггеры (3.1.3);
- 4 устройства для обработки цифровой информации (3.1.4).

На рис. 6.3 приведена классификация ИС, в которой две буквы обозначают функциональную подгруппу и вид микросхемы в системе условных обозначений отечественных интегральных схем.

Аналоговые интегральные схемы (3.2) предназначены для обработки сигналов, заданных в виде непрерывной функции. Этот класс ИС подразделяется на:

- ☐ генераторы (3.2.1);
- ☐ усилители (3.2.2);
- ☐ детекторы (3.2.3);
- ☐ устройства задержки сигналов (3.2.4);
- ☐ устройства селекции (3.2.5);
- ☐ фильтры частот (3.2.6);
- ☐ формователи (3.2.7);
- ☐ преобразователи (3.2.8);
- ☐ модуляторы (3.2.9).

С увеличением степени интеграции ИС увеличилась функциональная сложность микросхем и одновременно уменьшилась универсальность ИС. Появились БИС, функции которой позволяли программировать и хранить цифровые массивы. Класс ИС для вычислительных устройств получил название *микропроцессоров* (3.3.1), серии ИС программно и технологически совместимые получили название *микропроцессорных комплектов* (3.3.2).

Отдельный класс ИС составляют источники питания (3.4), многофункциональные устройства (3.5), фоточувствительные приборы с зарядовой связью (3.6).

Сокращение номенклатуры БИС возможно путем создания базового кристалла, представляющего собой матрицу из соединенных между собой элементов, электрические связи между которыми формируются в соответствии с функциональным назначением ИС. Такой класс ИС получил название *базового матричного кристалла* (БМК — 3.7).

По применяемости в аппаратуре интегральные схемы делятся на схемы общего применения (4.1) и схемы специального назначения (4.2).

По конструктивному оформлению ИС подразделяются на корпусные (5.1) и бескорпусные (5.2). Существует пять типов корпусов, отличающиеся как формой (прямоугольная, круглая, овальная), так и материалом (пластмассовые, керамические, металлотеклянные, металлокерамические, металлополимерные).

По применяемости в аппаратуре интегральные схемы подразделяются на схемы общего применения и схемы специального применения.

По технологии производства монолитных интегральных схем различают несколько типов:

- ☐ 6.1 — кремниевая технология;
- ☐ 6.2 — арсенид-галлиевая;
- ☐ 6.3 — кремний-германиевая.

6.2. Условные обозначения

Система условных обозначений типов отечественных интегральных схем и их маркировка установлены стандартом ГОСТ 1107 395-80. В основе системы условных обозначений лежит буквенно-цифровой код (рис. 6.4).

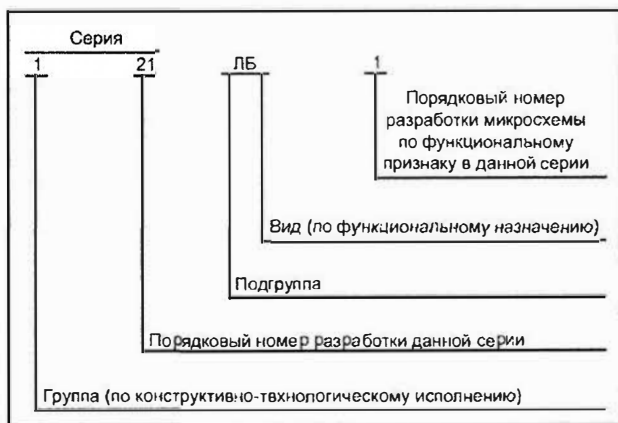


Рис. 6.4. Условные обозначения отечественных интегральных схем

Первый элемент — цифра, обозначающая группу интегральной микросхемы по конструктивно-технологическому исполнению: 1, 5, 6, 7 — полупроводниковые ИМС; 2, 4, 8 — гибридные; 3 — прочие (пленочные, вакуумные, керамические).

Второй элемент — две или три цифры (от 01 до 99 или от 001 до 999), указывающие на порядковый номер разработки данной серии ИМС. Первый и второй элемент образуют серию микросхем.

Третий элемент — две буквы, обозначающие функциональную подгруппу и вид микросхемы.

Четвертый элемент — число, обозначающее порядковый номер разработки микросхемы в серии.

Пример условного обозначения интегральной полупроводниковой логической микросхемы "И — НЕ / ИЛИ — НЕ" с порядковым номером разработки серии — 21, порядковым номером разработки данной схемы в серии по функциональному признаку — 1.

В обозначение также могут быть введены дополнительные символы (от А до Я), определяющие допуски на разброс параметров микросхем и т. п.

Перед первым элементом обозначения могут стоять следующие буквы:

- К — для аппаратуры широкого применения;
- Э — на экспорт (шаг выводов 2,54 и 1,27 мм);
- Р — пластмассовый корпус второго типа;
- М — керамический, металло- или стеклокерамический корпус второго типа;
- Е — металлополимерный корпус второго типа;
- А — пластмассовый корпус четвертого типа;
- И — стеклокерамический корпус четвертого типа;
- Н — кристаллоноситель.

Для бескорпусных интегральных микросхем перед номером серии может добавляться буква Б, а после нее или после дополнительного буквенного обозначения через дефис указывается цифра, характеризующая модификацию конструктивного исполнения:

- ☐ 1 — с гибкими выводами;
- ☐ 2 — с ленточными выводами;
- ☐ 3 — с жесткими выводами;
- ☐ 4 — на общей пластине (неразделенные);
- ☐ 5 — разделенные без потери ориентировки (например, наклеенные на пленку);
- ☐ 6 — с контактными площадками без выводов (кристалл).

Интегральные схемы зарубежного производства не имеют единой системы.

Так обозначения интегральных микросхем PRO ELEKTRON получили распространение в большинстве фирм-производителей интегральных микросхем европейских стран (Англии, Италии, Испании, Швеции, Франции, ФРГ и т. д.).

Код состоит из трех букв, за ними следует серийный номер (например, FYN121).

Для одиночных микросхем маркировка будет следующей.

Первый элемент — буква, отражающая принцип преобразования сигнала:

- ☐ S — цифровой;
- ☐ T — аналоговый;
- ☐ U — аналогово-цифровой.

Второй элемент — буква, выбираемая фирмой-изготовителем и не имеющая специального назначения. Исключение составляет буква H, которой обозначаются гибридные микросхемы.

Для семейств (серий) цифровых микросхем первый и второй элемент — это буквы, отражающие схемотехнологические особенности:

- ☐ FD — МОП-схемы;
- ☐ FL — стандартные ТТЛ-схемы;
- ☐ FQ — ДТЛ-схемы;
- ☐ FY — ЭСЛ-серия;
- ☐ GA — маломощные ТТЛ-схемы;
- ☐ GD — МОП-схемы;
- ☐ GF — стандартные ТТЛ-схемы;
- ☐ GJ — быстродействующие ТТЛ-схемы;
- ☐ GM — маломощные ТТЛ-схемы с диодами Шоттки;
- ☐ HB — комплементарные МОП-схемы серии 4000 А;
- ☐ HC — комплементарные МОП-схемы серии 4500 В.

Третий элемент — буква, обозначающая диапазон рабочих температур или другую важную характеристику:

- ☐ A — температурный диапазон не нормирован;
- ☐ B — от 0 до +70 °C;
- ☐ C — от -55 до +125 °C;
- ☐ D — от -25 до +70 °C;
- ☐ E — от -25 до +85 °C;

□ F — от -40 до $+85$ °C;

□ G — от -55 до $+85$ °C.

Четвертый элемент — четыре цифры, обозначающие серийный номер.

Кроме того, за цифрами может следовать буква для обозначения разновидности основного типа. Типы корпусов могут обозначаться одной или двумя буквами.

При двухбуквенном обозначении вариантов корпусов (после серийного номера) первая буква обозначает конструкцию:

□ C — цилиндрический корпус;

□ D — с двухрядным параллельным расположением выводов (DIP);

□ E — мощный с двухрядным расположением выводов (с внешним теплоотводом);

□ F — плоский с двусторонним расположением выводов;

□ G — плоский с четырехсторонним расположением выводов;

□ K — корпус типа ТО-3;

□ M — многорядный (больше четырех рядов);

□ Q — с четырехрядным параллельным расположением выводов;

□ R — мощный с четырехрядным расположением выводов (с внешним теплоотводом);

□ S — с одnorядным расположением выводов;

□ T — с трехрядным расположением выводов.

Вторая буква показывает материал корпуса:

□ G — стеклокерамика;

□ M — металл;

□ P — пластмасса;

□ X — прочие.

Обозначения корпусов с одной буквой следующие:

□ C — цилиндрический;

□ D — керамический;

□ F — плоский;

□ L — ленточный кристаллодержатель;

□ P — пластмассовый DIP;

□ Q — с четырехрядным расположением выводов;

□ T — миниатюрный пластмассовый;

□ U — бескорпусная ИС.

6.3. Основные параметры интегральных схем

Для сравнения различных типов микросхем используют такой параметр, как произведение задержки переключения на мощность. Чем меньше эта величина, тем выше качество интегральной схемы. Чем меньше этот параметр, тем более чувствительна интегральная схема к выходной нагрузке.

К основным параметрам интегральных схем относятся следующие:

- ☐ *Максимальное входное напряжение $U_{вх\ макс}$* — наибольшее значение входного напряжения интегральной схемы, при котором выходное напряжение соответствует заданному значению.
- ☐ *Минимальное входное напряжение $U_{вх\ мин}$* — наименьшее значение входного напряжения интегральной схемы, при котором выходное напряжение соответствует заданному значению.
- ☐ *Чувствительность S* — наименьшее значение входного напряжения, при котором электрические параметры интегральной схемы соответствуют заданным значениям.
- ☐ *Диапазон входных напряжений $\Delta U_{вх}$* — интервал значений напряжений от минимального входного напряжения до максимального.
- ☐ *Входное напряжение $U_{вх}$* — значение напряжения на входе интегральной схемы в заданном режиме.
- ☐ *Напряжение смещения $U_{сш}$* — значение напряжения постоянного тока на входе интегральной схемы, при котором выходное напряжение равно нулю.
- ☐ *Максимальное выходное напряжение $U_{вых\ макс}$* — наибольшее значение выходного напряжения, при котором изменения параметров интегральной схемы соответствуют заданным значениям.
- ☐ *Минимальное выходное напряжение $U_{вых\ мин}$* — наименьшее значение выходного напряжения, при котором изменения параметров интегральной схемы соответствуют заданным значениям.
- ☐ *Выходное напряжение баланса $U_{вых\ б\с}$* — значение напряжения постоянного тока на каждом выходе интегральной схемы относительно общего вывода, когда напряжение между выходами равно нулю.
- ☐ *Напряжение срабатывания $U_{срб}$* — наименьшее значение напряжения постоянного тока на входе, при котором происходит переход интегральной схемы из одного устойчивого состояния в другое.
- ☐ *Напряжение отпускания $U_{отп}$* — наибольшее значение напряжения постоянного тока на входе, при котором происходит переход интегральной схемы из одного устойчивого состояния в другое.
- ☐ *Минимальное прямое напряжение на переходах $U_{пр\ мин}$* — наименьшее значение падения напряжения на переходах интегральной схемы, при котором обеспечивается заданное значение электрических параметров интегральной микросхемы.
- ☐ *Максимальное обратное напряжение на переходах $U_{обр\ макс}$* — наибольшее значение падения напряжения на переходах интегральной схемы при протекании обратного тока.
- ☐ *Напряжение логической единицы U^1* — значение высокого уровня напряжения для "положительной" логики и значение низкого уровня напряжения для "отрицательной" логики.
- ☐ *Напряжение логического нуля U^0* — значение низкого уровня напряжения для "положительной" логики и значение высокого уровня напряжения для "отрицательной" логики.
- ☐ *Пороговое напряжение логической единицы $U_{пор}^1$* — наименьшее значение высокого уровня напряжения для "положительной" логики или наибольшее значение низкого уровня напряжения для "отрицательной" логики на входе интегральной схемы, при

котором происходит переход интегральной схемы из одного устойчивого состояния в другое.

- Пороговое напряжение логического нуля $U_{\text{нп}}^0$ — наибольшее значение низкого уровня напряжения для "положительной" логики или наименьшее значение высокого уровня напряжения для "отрицательной" логики на входе интегральной схемы, при котором происходит переход интегральной схемы из одного устойчивого состояния в другое.
- Входной ток $I_{\text{вх}}$ — значение тока, протекающего во входной цепи интегральной микросхемы в заданном режиме.
- Выходной ток $I_{\text{вых}}$ — значение тока, протекающего в цепи нагрузки интегральной микросхемы в заданном режиме.
- Максимальный выходной ток $I_{\text{вых макс}}$ — наибольшее значение выходного тока, при котором обеспечиваются заданные параметры интегральной схемы.
- Минимальный выходной ток $I_{\text{вых мин}}$ — наименьшее значение выходного тока, при котором обеспечиваются заданные параметры интегральной схемы.
- Входной ток логической единицы I^1 — входной ток, обеспечивающий формирование логической единицы.
- Входной ток логического нуля I^0 — входной ток, обеспечивающий формирование логического нуля.
- Выходной ток логической единицы $I_{\text{вых}}^1$ — выходной ток, обеспечивающий формирование логической единицы.
- Выходной ток логического нуля $I_{\text{вых}}^0$ — выходной ток, обеспечивающий формирование логического нуля.
- Ток утечки на входе $I_{\text{ут вх}}$ — значение тока во входной цепи интегральной схемы при закрытом состоянии входа и заданных режимах на остальных выводах.
- Ток утечки на выходе $I_{\text{ут вых}}$ — значение тока в выходной цепи интегральной схемы при закрытом состоянии выхода и заданных режимах на остальных выводах.
- Ток потребления $I_{\text{пот}}$ — значение тока, потребляемого интегральной схемой, от источников питания в заданном режиме.
- Ток короткого замыкания $I_{\text{кз}}$ — значение тока, потребляемого интегральной схемой при закороченном выходе.
- Ток холостого хода $I_{\text{хх}}$ — значение тока, потребляемого интегральной схемой при отключенной нагрузке.
- Потребляемая мощность $P_{\text{пот}}$ — значение мощности, потребляемой интегральной схемой, работающей в заданном режиме, от источников питания.
- Максимальная потребляемая мощность $P_{\text{пот макс}}$ — значение потребляемой мощности интегральной схемы в предельном режиме потребления.
- Потребляемая мощность в состоянии логической единицы $P_{\text{пот}}^1$ — потребляемая мощность, обеспечивающая формирование логической единицы.
- Потребляемая мощность в состоянии логического нуля $P_{\text{пот}}^0$ — потребляемая мощность, обеспечивающая формирование логического нуля.
- Средняя потребляемая мощность $P_{\text{пот ср}}$ — значение мощности, равное полусумме мощностей, потребляемых логической интегральной схемой от источников питания в двух различных устойчивых состояниях.

- Выходная мощность $P_{вых}$ — значение мощности сигнала, выделяемой на нагрузке интегральной схемы в заданном режиме.
- Рассеиваемая мощность $P_{расс}$ — значение мощности, рассеиваемой интегральной схемой, работающей в заданном режиме.
- Нижняя граничная частота полосы пропускания f_n — наименьшее значение частоты, на которой коэффициент усиления интегральной схемы уменьшается на 3 дБ при заданной частоте.
- Верхняя граничная частота полосы пропускания f_v — наибольшее значение частоты, на которой коэффициент усиления интегральной схемы уменьшается на 3 дБ от значения на заданной частоте.
- Полоса пропускания Δf — диапазон частот между верхней и нижней граничными частотами полосы пропускания.
- Время задержки импульса t_z — интервал времени между фронтами входного и выходного импульсов интегральной схемы, измеренный на заданном уровне напряжения или тока.
- Время нарастания выходного напряжения $t_{нр}$ — интервал времени, в течение которого выходное напряжение интегральной схемы изменяется с первого достижения уровня 0,1 до первого достижения уровня 0,9 установившегося значения.
- Время установления выходного напряжения $t_{уст}$ — интервал времени, в течение которого выходное напряжение интегральной микросхемы изменяется с первого достижения уровня 0,1 до последнего достижения уровня 0,9 установившегося значения.
- Время перехода из состояния логической единицы в состояние логического нуля $t^{1,0}$ — интервал времени, в течение которого напряжение на выходе интегральной схемы переходит от напряжения логической единицы к напряжению логического нуля, измеренный на уровнях 0,1 и 0,9 или на заданных значениях напряжения.
- Время перехода из состояния логического нуля в состояние логической единицы $t^{0,1}$ — интервал времени, в течение которого напряжение на выходе интегральной схемы переходит от напряжения логического нуля к напряжению логической единицы, измеренный на уровнях 0,1 и 0,9 или на заданных значениях напряжения.
- Среднее время задержки распространения сигнала логической интегральной микросхемы $t_{зд.р.ср}$ — интервал времени, равный полусумме времен задержки распространения сигнала при включении и выключении логической интегральной микросхемы.
- Время хранения информации при отключении напряжения питания $t_{хр}$ — интервал времени с момента отключения источника питания интегральной схемы, в течение которого записанная информация сохраняется с заданными параметрами.
- Время считывания информации $t_{сч}$ — интервал времени между фронтами адресного и считанного сигналов интегральной схемы, измеренный на заданных уровнях в заданном режиме.
- Время записи информации $t_{зп}$ — интервал времени между началом адресного сигнала и появлением записанной информации на выходе интегральной схемы, измеренный на заданных уровнях.
- Коэффициент усиления напряжения $K_{у,н}$ — отношение выходного напряжения интегральной схемы к входному напряжению.
- Коэффициент усиления тока $K_{у,т}$ — отношение выходного тока интегральной схемы к входному току.

- Коэффициент усиления мощности K_y — отношение выходной мощности интегральной схемы к входной мощности.
- Коэффициент нелинейности амплитудной характеристики $K_{нл}$ — наибольшее отклонение значения крутизны амплитудной характеристики интегральной схемы относительно значения крутизны амплитудной характеристики, изменяющейся по линейному закону.
- Коэффициент объединения по входу $K_{об}$ — число входов интегральной схемы, по которым реализуется логическая функция.
- Коэффициент разветвления по выходу $K_{раз}$ — число единичных нагрузок, которое можно одновременно подключить к выходу интегральной схемы.
- Входное сопротивление $R_{вх}$ — величина, равная отношению приращения входного напряжения интегральной схемы к приращению активной составляющей входного тока при заданном значении частоты сигнала.
- Выходное сопротивление $R_{вых}$ — величина, равная отношению приращения выходного напряжения интегральной схемы к вызвавшему его активной составляющей выходного постоянного или синусоидального тока при заданном значении частоты сигнала.
- Входная емкость $C_{вх}$ — величина, равная отношению емкостной реактивной составляющей входного тока интегральной схемы к произведению круговой частоты на синусоидальное входное напряжение схемы при заданном значении частоты сигнала.
- Выходная емкость $C_{вых}$ — величина, равная отношению емкостной реактивной составляющей выходного тока интегральной схемы к произведению круговой частоты на вызванное им выходное напряжение при заданном значении частоты сигнала.

6.4. Применение и эксплуатация интегральных схем

Интегральные схемы должны быть изготовлены в соответствии с требованиями стандарта отрасли ОСТ В 11 0398-2000 и техническими условиями на микросхемы конкретных типов по рабочей конструкторской и технологической документации, утвержденной в установленном порядке.

Применение и эксплуатация интегральных схем производится в соответствии с указаниями, изложенными в ОСТ В 11 0398-2000.

Технологический процесс изготовления аппаратуры на основе интегральных схем должен соответствовать типовому технологическому процессу с учетом требований стандарта ОСТ 11 073.063.

ИС могут применяться в газовых смесях следующего состава:

- смеси воздуха, азота, инертных газов в любых соотношениях;
- газовых смесях, содержащих кислород до 50%, углекислый газ до 3%, остальное — азот или инертные газы;
- озона с концентрацией $0,1 \text{ мг·см}^{-3}$.

При этом давление газовых смесей не должно превосходить 3 атм. Время воздействия не ограничивается.

ИС могут применяться в условиях невесомости и любых значениях атмосферного пониженного давления при обеспечении конструктивных мер, обеспечивающих температур-

ные режимы, установленные в нормативной документации (НД). ИС могут применяться в условиях воздействия акустических шумов с нижней границей частотного диапазона менее 50 Гц, при этом уровень звукового давления в диапазоне 10 000 Гц не должен превосходить указанных в технических условиях (ТУ) значений. Конструкция микросхем обеспечивает отсутствие резонансных частот в диапазоне частот до 100 Гц.

ИС могут применяться в условиях:

- ☐ постоянных и медленно меняющихся полей с частотой до 0,034 Гц и напряженностью магнитного поля до 8000 А/м;
- ☐ переменных и импульсных магнитных полей с частотой до 500 Гц и напряженностью 80 А/м;
- ☐ электромагнитного импульса с напряженностью электрического поля до 100 кВ/м, магнитного поля до 300 А/м и длительностью до 1 мкс (при условии, что наведенные сигналы на подводящих проводах в момент воздействия электромагнитного излучения не превосходят допустимые по ТУ электрические режимы микросхем);
- ☐ солнечного излучения — согласно ГОСТ В 20.39.404. Микросхемы могут применяться при длительностях действия ударного ускорения. Допустимое значение длительности действия ударного ускорения устанавливают согласно рекомендации.

При проверке электрических цепей РЭА, содержащих интегральные схемы, напряжения, прикладываемые между двумя выводами микросхем, не должны превышать 0,3 В, и ток по любому выводу микросхем не должен превышать 1 мА, если иное не оговорено в ТУ или руководствах по применению конкретных типов микросхем.

Требования для ИС со спецприемкой значительно выше, чем для схем для народнохозяйственного применения.

Рекомендуемая литература

1. ОСТ В 11 0398-2000.
2. Пляц О. М. Справочник по электровакуумным, полупроводниковым приборам и интегральным схемам. — Минск: Вышэйш. школа, 1976.

7. Логические элементы интегральных схем

7.1. Классификация логических элементов

Логическими элементами ИС называются электронные схемы, выполняющие простейшие логические операции. Логические элементы используются в цифровых схемах в качестве основных элементов и определяют параметры микросхемы.

Логические элементы представляют собой техническую модель логических выражений булевой алгебры. Ее создателем был английский математик и логик Джордж Буль, который в 1848 году разработал алгебру логики.

Одна из первых моделей была предложена в 1910 году физиком П. Эренфестом, который на основе телефонных реле создал вычислительное устройство. В нем использовалась аналогия между высказываниями в булевой алгебре и электрическими переключателями. Физическая природа переключателей и математическая природа высказываний в булевой алгебре идентичны.

На рис. 7.1 приведены основные электрические модели для констант 0 и 1 и базовых функций И, ИЛИ, НЕ.

В алгебре логики различные логические выражения могут принимать только два значения: "истинно" или "ложно". Для обозначения истинности используется цифра 1, ложности — 0. Можно и наоборот, но в этом случае имеет место понятие обратной логики.

Исчисление, в котором используются только две цифры, называется *двоичным исчислением*.

В алгебре логики используются три основные операции:

- ☐ Логическое отрицание или инверсия называется операцией НЕ и обозначается штрихом "—" переменной.
- ☐ Логическая операция ИЛИ для двух переменных x_1 и x_2 записывается как $y = x_1 + x_2 = x_1 \vee x_2$. Знак "+" может быть заменен знаком " \vee ", обозначающим логическое сложение. Логическое ИЛИ определяется как $y = 1$, если $x_1 = 1$ или $x_2 = 1$, или $x_1 = x_2 = 1$.
- ☐ Логическая операция И для двух переменных записывается в виде $y = x_1 * x_2 = x_1 \wedge x_2$. Знак "*" может быть заменен знаком " \wedge ", обозначающим логическое умножение. Логическое И для двух переменных представляется как $y = 1$ только в том случае, если $x_1 = 1$ и $x_2 = 1$.

Существует множество комбинаций логических операций как для двух, так и для многих переменных.

Например, комбинация операций ИЛИ и НЕ формирует функцию ИЛИ --- НЕ:

$$y = x_1 \vee x_2,$$

а комбинация И и НЕ функцию И — НЕ:

$$y = x_1 \wedge x_2$$

В двоичной логике число разложенных сочетаний из n аргументов равно 2^n , а число логических функций — 2^{2^n} .

Логические элементы могут быть реализованы на различных физико-технических принципах: электромеханическом, пневматическом, оптическом и т. п.

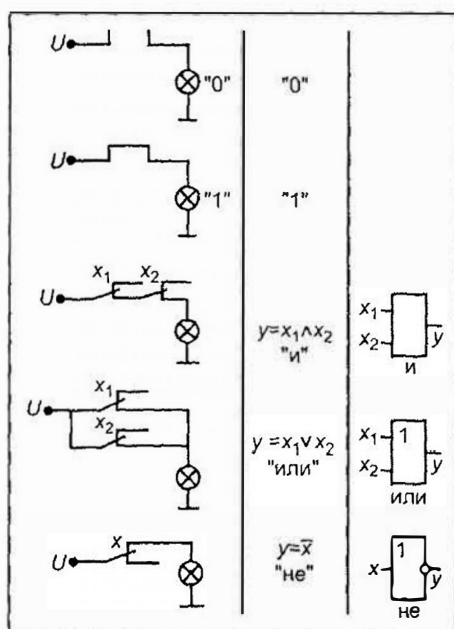


Рис. 7.1. Переключательные модели констант 0 и 1, функции И, ИЛИ, НЕ и их условное обозначение

Совокупность требований по быстродействию, массогабаритным размерам, надежности, энергопотребления лучше всего реализуется в цифровых интегрированных схемах.

В основе цифровых схем лежат *транзисторные ключи* — аналоги металлических контактов (рис. 7.2).

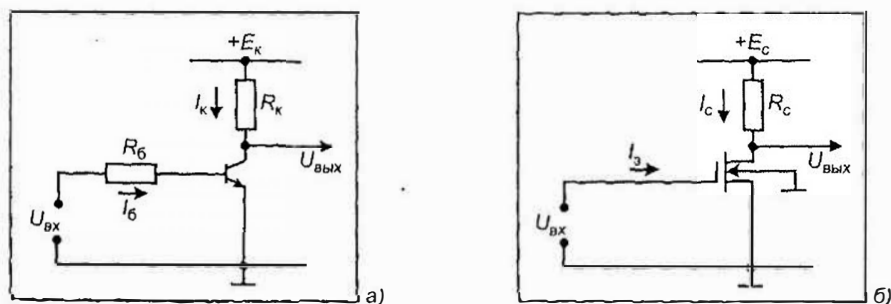


Рис. 7.2. Транзисторный ключ, реализованный на биполярном (а) и МДП (б) транзисторах: E_k — напряжение коллектора; E_c — напряжение стока; I_k — ток коллектора; I_c — ток стока; R_k, R_a, R_c — резисторы

Ключи характеризуются двумя устойчивыми состояниями — разомкнутым и замкнутым. Формируя алгебраические логические функции в нужную комбинацию, можно получить необходимый результат. Практическое аппаратное решение можно получить, соединяя соответствующим образом логические элементы.

Логические элементы по режиму работы подразделяются на *статические* и *динамические*.

Статические логические элементы могут работать как в статическом, так и в импульсном или динамическом режимах.

Динамические логические элементы могут работать только в импульсном режиме.

Логические элементы в микроэлектронике подразделяются на комбинационные и последовательностные схемы.

Комбинационные логические схемы представляют собой схемы без запоминания переменных. Они состоят из логических либо операционных элементов и собраны из логических элементов для выполнения заданных операций над входными сигналами.

Наиболее распространенными являются следующие типы комбинационных схем.

- ❑ *Шифратор* (кодировщик) — операционный элемент, преобразующий единичный сигнал на одном из n входов в m -разрядный выходной код. Конструктивно он может быть реализован из четырех элементов ИЛИ: одного пятиходового, двух четырехходовых и одного двухходового.
- ❑ *Дешифратор* (декодировщик) — операционный элемент, преобразующий n -разрядный входной код в сигнал только на одном из своих m выходов. Логические функции, а затем и схема дешифратора составляются по таблицам истинности.
- ❑ *Мультиплексор* — операционный элемент, осуществляющий адресное переключение заданного числа входных сигналов на один выход.
- ❑ *Демultipлексор* — операционный элемент, осуществляющий адресное подключение одного входного сигнала к одному из множества выходов.
- ❑ *Компаратор* — операционный элемент, производящий сравнение двух чисел x_1 и x_2 . Результат сравнения отображается единичным логическим уравнением.
- ❑ *Сумматор* — операционный элемент, выполняющий операцию сложения нескольких чисел.

В классификации интегральных схем эти устройства вполне логично отнесены к цифровым устройствам, потому как используются для преобразования информации.

Последовательностными логическими схемами называют схемы, состояние выходов которых зависит от последовательности смены состояний на их входах. Они могут запоминать переменные, выходные сигналы которых зависят не только от значения входных сигналов в данный момент времени, но и от последовательности значений входных сигналов в предшествующие моменты времени. Последовательностные схемы собираются из комбинационных путем введения в них обратных связей.

К последовательностным логическим схемам относятся:

- ❑ *триггеры* — последовательные элементы с двумя устойчивыми выходными состояниями;
- ❑ *регистры* — последовательный операционный элемент, предназначенный для хранения и (или) преобразования многоразрядных двоичных чисел. Регистр состоит из набора триггеров, число которых равно разрядности хранимых чисел;

□ **счетчики** — последовательный операционный элемент, предназначенный для счета импульсов, поступающих на вход. Конструктивно счетчик состоит из цепочки триггеров, число которых определяет его разрядность.

Перечисленные устройства относятся к времязадающим.

К последовательностным схемам относятся также и *запоминающие устройства*.

7.2. Основные характеристики логических элементов

Основной характеристикой логического элемента, в том числе транзисторного ключа, является *передаточная характеристика*. Эта характеристика представляет собой зависимость выходного напряжения $U_{\text{вых}}$ от напряжения на одном из входов при условии сохранения постоянных напряжений на остальных входах.

Для транзисторных ключей передаточная характеристика определяется выражением $U_{\text{вых}} = f(U_{\text{вх}})$ (рис. 7.3).

По виду передаточной характеристики различают *инвертирующие* и *неинвертирующие* логические элементы.

К инвертирующим элементам относятся схемы типа НЕ, И — НЕ, ИЛИ — НЕ и др., на выходе которых получают инверсные по отношению к выходным логические сигналы.

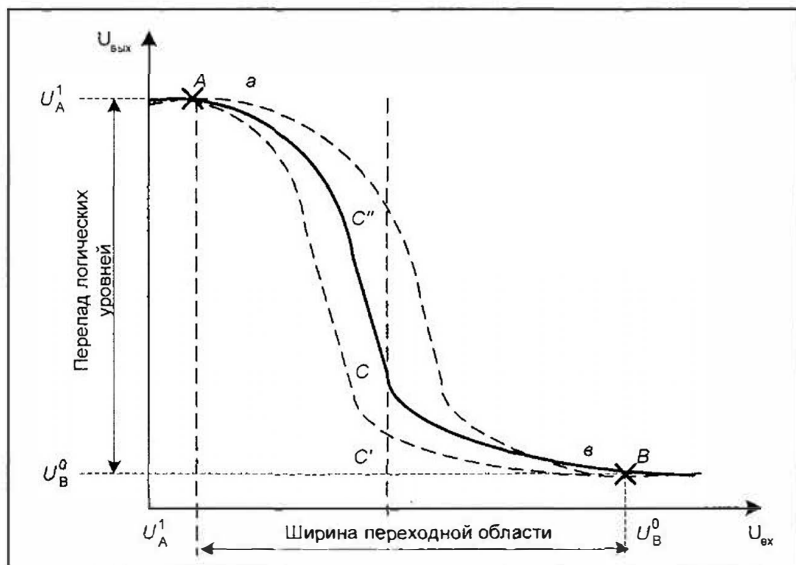


Рис. 7.3. Инвертирующая передаточная характеристика

К неинвертирующим логическим элементам относят схемы типа И, ИЛИ и др., на выходе которых получают прямые по отношению к входным логические сигналы.

Транзисторные ключи относятся к инвертирующим логическим элементам. Для них различают два устойчивых состояния: разомкнутое, соответствующее точке А, и замкну-

тос — точке B (рис. 7.3). В точке A ключ разомкнут и на него подают большое напряжение (при малом входном). С увеличением входного напряжения ключ срабатывает, транзистор открывается, и все напряжение падает на нагрузочном сопротивлении (R_c или R_e). Напряжение на выходе близко к нулю.

Входные и выходные сигналы имеют дискретные уровни, которые и определяются передаточной характеристикой.

Следует заметить, что форма передаточной характеристики между точками A и B несущественна. Штриховой линией обозначен возможный разброс значений $C' < C < C''$ на интервале $]U_A^{ex}, U_B^{ex}[$, который не влияет на точки A и B . Ясно, что ключи, а значит и цифровые схемы, малочувствительны к разбросу параметров, температурному дрейфу, временному изменению параметров, внешним электромагнитным наводкам, собственным шумам.

Входные напряжения, определяющие границы участков, называются *порогами переключения* U_A^1 и U_B^0 , которые определяют ширину переходной области. Разность напряжений логической 1 и 0 называется *логическим перепадом* и определяется как

$$U_{\Delta} = U_A^1 - U_B^0.$$

Помимо логических сигналов на входах может появиться напряжение помех, которое либо повышает, либо понижает входное напряжение. Это может привести к сбоям в работе логических элементов. Для повышения помехоустойчивости логических элементов необходимо увеличивать логический перепад и уменьшать ширину переходной области.

Входная характеристика представляет собой зависимость входного тока от напряжения на данном входе при постоянных напряжениях на остальных входах.

Для биполярного ключа это будет характеристика $I_b = f(U_{вх})$, а для униполярного — $I_z = f(U_{вх})$.

Выходная характеристика представляет собой зависимость выходного напряжения от выходного тока. Для транзисторного биполярного ключа это $I_k = f(U_{вых})$, а для транзисторного униполярного ключа — $I_c = f(U_{вых})$. Эти характеристики могут строиться для напряжений низкого и высокого уровней на выходе.

Нагрузочная способность n или *коэффициент разветвления* на выходе характеризует максимальное число логических элементов, которые можно одновременно подключить к его выходу.

Коэффициент объединения по выходу m характеризует расширение логических возможностей элемента за счет выполнения функций над большим числом логических переменных.

Быстродействие логического элемента оценивается средним временем задержки распространения сигнала, по существу определяющее среднее время выполнения логической операции.

В схемах, реализующих логические функции, логические нули и логические единицы представлены различными значениями напряжения — уровнем нуля U^0 и уровнем единицы U^1 . Если $U^1 - U^0 > 0$, то схема работает в "положительной логике", если $U^0 - U^1 > 0$, то схема работает в "отрицательной логике". Между "положительной" и "отрицательной" логикой принципиальной разницы нет.

7.3. Логические ИС на биполярных транзисторах

Различают три основные группы логических элементов интегральных схем, реализованных на биполярных транзисторах.

К первой группе относятся логические схемы с передачей выходного тока или напряжения на вход нагрузочного элемента. В эту группу входят логические элементы транзисторной логики с непосредственной связью, транзисторной логики с резистивной связью (РТЛ), транзисторной логики с резистивно-емкостной связью (РЕТЛ) и интегральной инжекционной логики (I^2L).

Ко второй группе относятся ИС с логикой на входе (конъюнкция и дизъюнкция) и с передачей входного тока на выход управляющего элемента. В эту группу могут быть включены логические элементы диодно-транзисторной логики (ДТЛ), транзисторно-транзисторной логики (ТТЛ), в том числе ТТЛШ, диодно-транзисторной логики с дополнительной симметрией (ДСДТЛ), модифицированной диодно-транзисторной логики (МДТЛ) и транзисторной логики с переменным порогом (ППТЛ).

Третью группу образуют логические ИС с эмиттерной связью и токовым переключением. К ним относятся элементы эмиттерно-связанной логики (ЭСЛ), эмиттерно-связанной логики с эмиттерным повторителем (ЭСЭЛ), эмиттерно-связанной логики с дополнительной симметрией (ДСЭЛ).

Рассмотрим каждую из групп и приведем некоторые характеристики логических схем.

7.3.1. Логические элементы с передачей тока или напряжения

Транзисторная логика с непосредственными связями

Транзисторная логика с непосредственными связями (ТЛНС) основана на параллельном (или последовательном) соединении транзисторных ключей и использовании общей коллекторной нагрузки (рис. 7.4). Конструктивно транзисторы T_1 , T_2 , T_3 объединены по коллектору и подключены через резистор R_k к шине E_k . Входные сигналы подаются на базы транзисторов. Выходы подключены на входы таких же элементов. Схема рассчитывается так, чтобы при подаче высокого напряжения (соответствующей логической 1) на базу одного или нескольких транзисторов происходит насыщение транзисторов и выходной потенциал снижается до низкого уровня — остаточного потенциала $U^0 = U_{ост}$.

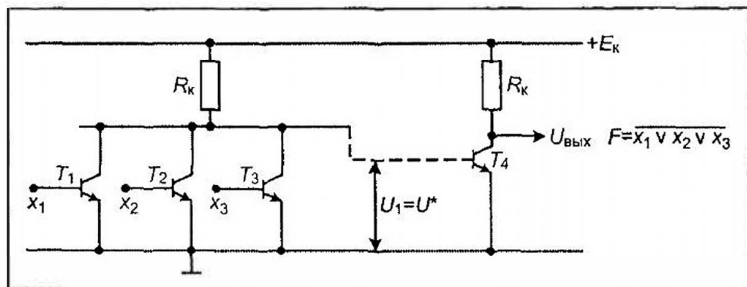


Рис. 7.4. Основная логическая схема ТЛНС

Выходной потенциал соответствует логическому нулю, и все нагрузочные транзисторы оказываются запертыми. При подаче на входы транзисторов T_1 , T_2 , T_3 низкого потенциала,

соответствующего логическому 0, все транзисторы запираются, потенциал их коллектора повышается, стремясь достигнуть уровня E_k . В этом случае последующие транзисторы отпираются, потому что потенциал выходного напряжения соответствует потенциалу базы насыщенного транзистора. Этот потенциал также соответствует логической 1. Таким образом, в положительной логике ТЛНС выполняет операцию дизъюнкции, реализуя на выходе функцию

$$F = x_1 \vee x_2 \vee x_3 \vee \dots \vee x_l.$$

Эта соответствует логической схеме ИЛИ — НЕ. Основным преимуществом элемента ТЛНС является его простота. Принципиальным недостатком является сильная зависимость процессов от характеристик транзисторов. Это проявляется, прежде всего, в изменении уровней сигналов с изменением числа входов и нагрузки.

Базовые токи коллекторов распределяются неравномерно вследствие разброса параметров транзисторов. Происходит перехват токов транзисторов. При этом большая часть тока источника E_k будет поступать в базу одного из транзисторов, а остальные режимы насыщения не достигнут. Работа схем ТЛНС становится ненадежной.

Резисторно-транзисторная логика

Ненадежность работы ТЛНС послужила причиной поиска более совершенных вариантов. Появилась *транзисторная логика с резистивной связью* (ТЛРС и РТЛ). Принципиальным отличием является включение в базовые цепи транзисторов и резисторов с сопротивлением порядка 10^2 Ом (рис. 7.5). Наличие резисторов позволяет выравнивать входные токи в базовые цепи. Сопротивления резисторов R_{B_i} должны быть большим для выравнивания входных характеристик и одновременно достаточно малым, чтобы не препятствовать насыщению транзисторов вследствие уменьшения тока базы. Этим противоречивым требованиям удовлетворяет некоторое оптимальное отношение R_B / R_k , при котором нагрузочная способность становится максимальной.

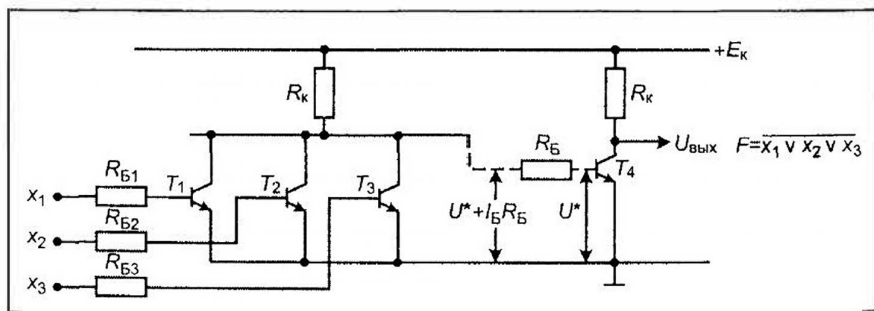


Рис. 7.5. Логическая схема РТЛ

Следует заметить также, что введение резисторов R_{B_i} уменьшает быстродействие схемы вследствие возрастания длительности фронта в ключе.

Схема реализует функцию

$$F = x_1 \vee x_2 \vee x_3.$$

резисторно-емкостная транзисторная логика

Для ослабления влияния сопротивления на быстродействие элемента целесообразно его шунтировать конденсатором небольшой емкости. Логика, в которой реализован этот принцип, получила название *резистивно-емкостной транзисторной логики* (РЕТЛ) (рис. 7.6).

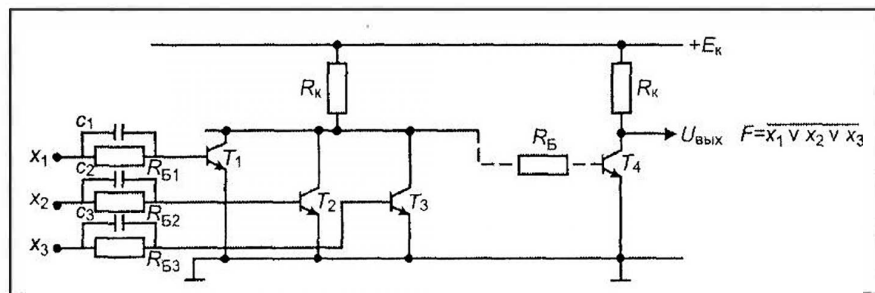


Рис. 7.6. Логическая схема РЕТЛ

Конструктивно такой конденсатор получают в виде обратносмещенного р—п-перехода. Во время переключения конденсатор закорачивает высокоомный резистор $R_{Б}$, что приводит к образованию всплеска тока базы. Этот прием обеспечивает быстрое отпирание и запираание транзисторов. Оптимальное значение составляет при выполнении условия $R_{Б} / R_K \geq 10$, что требует создания высокоомного резистора с большой площадью. Такие схемы использовались на первом этапе развития микроэлектроники. В итоге они оказались бесперспективными вследствие большого числа резисторов и емкостей, которые занимали большую площадь.

В результате было предложено все элементы, формирующие ток в базу, заменить одним инжектором (рис. 7.7).

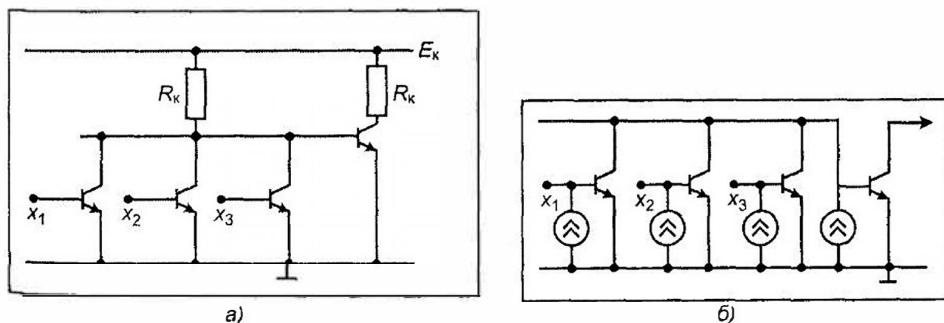


Рис. 7.7. Эволюция ТЛНС (а) в И²Л путем подключения генераторов тока в базы (б)

Интегральная инжекционная логика

В ходе развития дискретной полупроводниковой электроники возникла принципиально новая ранее не известная логика — *интегральная инжекционная логика* (И²Л).

В основе интегральной инжекционной логики лежат функционально интегрированные транзисторные структуры (рис. 7.8, а). Транзистор T_n называют токозадающим. Он состо-

ит из инжектора I , который эмитирует носители заряда-дырки в эмиттерную область многоколлекторного транзистора T . Транзистор T_n представляет собой $p_1-n_1-p_2$ и расположен горизонтально. Многоколлекторный инвертирующий транзистор $T_{p_2-p_2-n_1}$ -типа расположен вертикально и имеет общий эмиттер \mathcal{O} . Эмиттерная область представляет собой сильно легированное основание подложки. Эмиттерная область транзистора T одновременно служит базой токозадающего транзистора T_n .

Инвертор включается тогда, когда ток инжектора T_n отбирается из базы многоколлекторного транзистора T в другую цепь, например предшествующей структуры в схеме. Такое включение может быть обеспечено за счет соответствующего уменьшения входного напряжения $U_{вх}$. Это напряжение одновременно управляет смещением на эмиттерном переходе инвертора T .

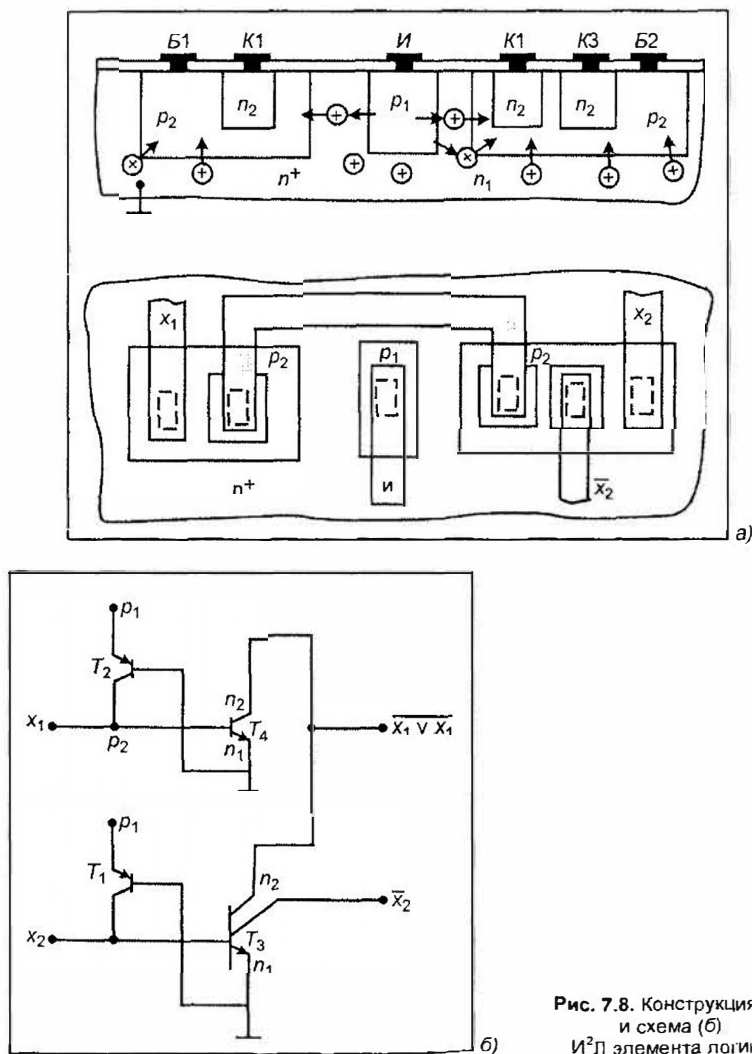


Рис. 7.8. Конструкция (а) и схема (б) И²Л элемента логики

Элемент I^2L обычно реализует функции ИЛИ — НЕ. Функция И — НЕ может быть реализована при условии использования коллекторных выходов в качестве независимых входов И для последующих логических элементов. Наличие многоколлекторного инвертора позволяет осуществить логическую развязку без дополнительных схемных элементов.

Логические уровни и логический перепад в схеме I^2L описываются как в ТЛНС и имеют подобные характеристики. Оригинальность схемотехнического решения сочетается с оригинальностью технологического решения.

Инжектор реализуется в виде длинной р-полоски, выполненной на этапе базовой диффузии. Базой р—п—р-транзистора является эпитаксиальный п-слой, а коллекторами — базовые слои п—р—п-транзисторов. Расположение р—п—р-транзисторов относительно инжектора может быть как перпендикулярным, так и параллельным.

Преимуществом I^2L является отсутствие изолирующих карманов и резисторов, приводящих к экономии площади, уменьшение напряжения питания, мощности и времени задержки. Малая емкость коллектора, малое остаточное напряжение на насыщенных транзисторах обусловлено низкоомным слоем n^+ -коллектора. Структуры с инжекционным питанием достаточно универсальны. Они могут использоваться для построения арифметических устройств, устройств памяти, логики. I^2L -схемы хорошо согласуются с ТТЛ и ДТЛ-схемами.

7.3.2. Логические элементы с логикой на входе

Диодно-транзисторная логика

Диодно-транзисторная логика (ДТЛ) отличается от предыдущих схем тем, что количество логических входов не связано с количеством транзисторов. Логическая функция в этом случае осуществляется диодами D_1 , D_2 и D_3 , а транзистор T выполняет функцию инверсии. Таким образом, резко сокращается количество транзисторов. В этой группе схем с логикой на входе и передачей входного тока на выход управляющей ИС не возникает эффекта перехвата тока из-за неравномерного распределения его между входными цепями. Входные диоды обеспечивают развязку цепей друг от друга (рис. 7.9). Диоды D_i и D_j выполняют задачу обеспечения сдвига уровня постоянного напряжения между точками a и b . Они называются диодами смещения. Для того чтобы работа диодов не зависела от состояния транзистора (наличия или отсутствия в нем тока), существует цепь смещения ($-E$ и R_1), через которую протекает ток. Этот ток обеспечивает работу диодов D_i и D_j в прямом направлении и создает смещение $2U^*$.

Пусть $U_1 = U_2 = \dots = U_n = U_0 = 0$, тогда через диоды D_1 , D_2 будет протекать ток I_0 . Тогда напряжение в точке a будет равно прямому напряжению на диоде $U_a = U^*$, а ток — $I = E_k - U^* / R_0$. В точке a $U_b = -U^*$. В этом случае эмиттерный переход находится под обратным смещением и транзистор T заперт. Выходное напряжение $U_{вых}$ имеет максимальную величину $U_{вых} = E_k = U^1$, другими словами, $U_{вых}$ имеет уровень логической единицы. Тогда на вход подадим напряжение $U_{вх} = U^1$. Напряжение на остальных входах осталось равным нулю и диод D_2 , в частности, по-прежнему открыт. Напряжение в точке a по-прежнему равно U^* . Следовательно, диод D_1 оказывается под большим обратным напряжением, запирается и ток через него делается близким к нулю. Никаких других изменений в схеме не происходит и на выходе сохраняется равенство $U_{вых}^1 = E_k$. То же самое получится, если подать на вход 2 напряжение U^1 . Если подать напряжение U^1 на все входы,

то все диоды запираются. Ток I_0 будет протекать через диоды смещения в базу транзистора. Транзистор отпирается, потенциал базы делается равным U^* и при условии насыщения потенциал коллектора становится равным $U_{\text{сст}}$. Это и есть значение уровня логического нуля $U_{\text{сст}} = U^0 \cong 0,1$ В. Таким образом, уровень выходного напряжения меняется с U^1 до U^0 только при подаче на все входы уровня напряжения U^1 . В положительной логике схема выполняет функцию

$$F = \overline{x_1 \wedge x_2 \wedge x_3}.$$

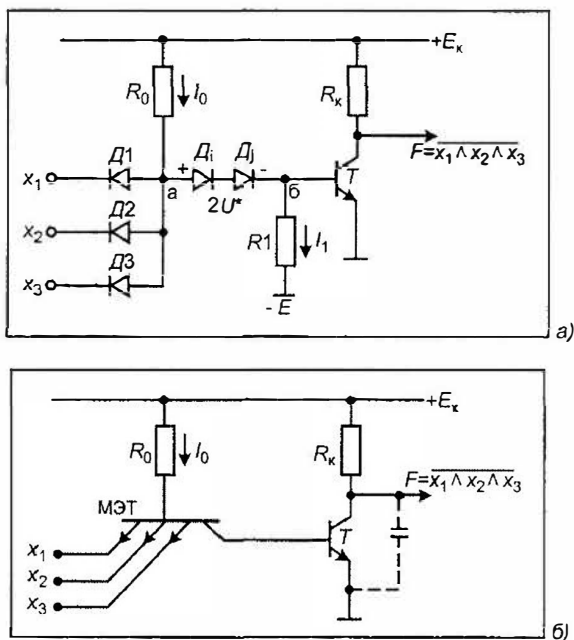


Рис. 7.9. Логический элемент ДТЛ (а) и его эволюция к элементу ТТЛ (б)

Преимуществом схем ДТЛ является надежное запираение транзистора путем подачи на его эмиттерный переход обратного смещения. Кроме того, этим схемам свойственен большой логический перепад $U^1 - U^0 \cong E_k$. К разновидностям схем ДТЛ относятся *модифицированная диодно-транзисторная логика* (МДТЛ), входом которой служат эмиттерные повторители.

Эмиттерные повторители способны усиливать входной ток и, следовательно, улучшить параметры схем ДТЛ. Разновидностью элемента ДТЛ является *диодно-транзисторная логика с дополнительной симметрией* (ДСДТЛ), а также логики с *переменным порогом* (ПНТЛ), на входе которых расположены инверторы. Эти схемы отличаются высокой помехоустойчивостью. Однако большое количество диодов требует создания изолирующего кармана. Площадь под диоды существенно повышается, падает степень интеграции. В схеме ДТЛ можно заметить, что совокупность логических диодов и диодов смещения соответствует структуре транзистора — два встречно включенных р—п-перехода.

Транзисторно-транзисторная логика

Различие транзисторно-транзисторной логики (ТТЛ) схемы от ДТЛ сводится к двум обстоятельствам (рис. 7.9, б). Во-первых, вместо диодов смещения имеется один коллекторный переход многоэмиттерного транзистора (МЭТ). В этом случае при нулевом входном напряжении потенциал на базе транзистора будет не отрицательным, а близким к нулю. Транзистор будет в этом случае заперт, но помехоустойчивость схемы снижается. Это обусловлено отсутствием источника смещения ($-E$) и экономией площади под диоды и резистор R_1 .

Во-вторых, возможно взаимодействие между эмиттерами МЭТ, в отличие от изолированных диодов. В результате горизонтального транзисторного эффекта в эмиттере, на которое подано запирающее напряжение U^1 , может протекать обратный паразитный ток. Этот ток обязан своим появлением инжекцией электронов из смежного открытого эмиттера. Чтобы избежать горизонтального транзисторного эффекта, необходимо увеличить расстояние между эмиттерами l_0 так, чтобы превысить диффузионную длину носителей в базовом слое. Схема выполняет логическую функцию И — НЕ:

$$F = x_1 \wedge x_2 \wedge x_3.$$

Одним из недостатков схем ТТЛ является ее малая нагрузочная способность. Причиной этого является насыщение транзисторов. Для преобразования эффекта насыщения транзисторов в области базы используется нелинейная обратная связь.

Транзисторно-транзисторная логика с диодами Шоттки

Введение нелинейной обратной связи состоит в том, что между коллектором и базой транзистора включается диод Шоттки. Это привело к созданию транзисторно-транзисторной логики с диодами Шоттки (рис. 7.10). Шунтирование диодом Шоттки перехода "коллектор—база" позволяет избежать насыщения, что в свою очередь приводит к увеличению падения напряжения на переходе "база—эмиттер". Это уменьшает ток потребления в статическом режиме и, соответственно, потребляемую схемой мощность.

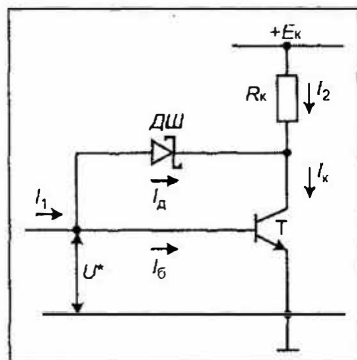


Рис. 7.10. Шунтирование диодом Шоттки перехода "база—коллектор" транзистора

7.3.3. Логические схемы на переключателях тока

Эмиттерно-связанная логика

В логических интегральных схемах, относящихся к эмиттерно-связанной логике (ЭСЛ), для реализации логических операций и других преобразований дискретной информации используются транзисторные переключатели тока с объединенными эмиттерами.

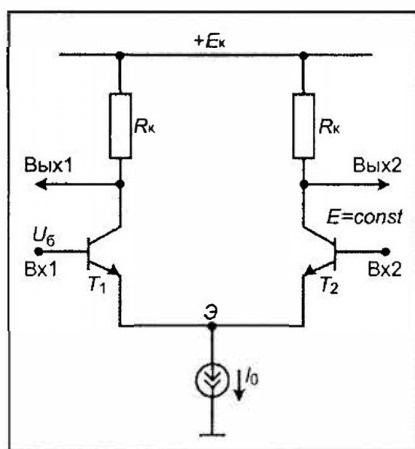


Рис. 7.11. Схема переключателя тока

Переключателем тока называют симметрическую схему, в которой заданный ток I_0 протекает через определенную часть переключателя в зависимости от потенциала $U_б$ на одном из входов. Потенциал E на другом входе имеет постоянную величину.

В отличие от уже рассмотренного ключа в переключателе тока управление осуществляет не током, а напряжением (рис. 7.11).

Если $U_б = E$, то открыты оба транзистора и ток I_0 делится пополам для каждой ветви. Если уменьшить потенциалы $U_б$, то при неизменном потенциале ток транзистора T_1 уменьшится. Транзистор T_1 закроется, а через транзистор T_2 будет протекать полный ток.

При увеличении потенциала базы $U_б$ возрастет потенциал эмиттеров, что приведет к уменьшению тока через транзистор T_2 . Транзистор окажется запертым и весь ток будет протекать по транзистору T_1 . Другими словами, перепад потенциала на базе около средней величины E обеспечивает переключение тока из одного транзистора в другой.

● особенность переключателя тока состоит в том, что транзисторы всегда работают в неаасыщенном режиме — активном режиме. Это обеспечивает повышенную скорость переключения потому, что не тратится время на рассасывание заряда.

На рис. 7.12, а приведена схема двухвходовой логической ячейки, реализованная на переключателях тока.

Роль генератора тока выполняет токозадающий резистор $R_г$. Эмиттеры всех транзисторов соединены в одной точке. В схеме предусмотрены два эмиттерных повторителя, реализованных на транзисторах T_1 и T_2 и резисторах $R_{э,н}$. Ячейка имеет два выхода.

□ Выход 1 инвертирует сигнал и реализует функцию ИЛИ — НЕ ($F_1 = x_1 \vee x_2$).

□ Выход 2 — прямой, ему соответствует логическая функция ИЛИ ($F_2 = x_1 \vee x_2$).

Передаточная характеристика для элемента ЭСЛ представлена на рис. 7.12, б.

Логический перепад достаточно высок, что позволяет сделать схемы ЭСЛ помехоустойчивыми. Эмиттерные повторители на обоих выходах ускоряют процесс зарядки емкости нагрузки. Они же ослабляют зависимость уровня напряжения от числа нагрузок.

Дальнейшее совершенствование логических элементов на переключателях тока привело к разработке схем *эмиттерно-связанной логики с эмиттерными повторителями на вход* (ЭЭСЛ).

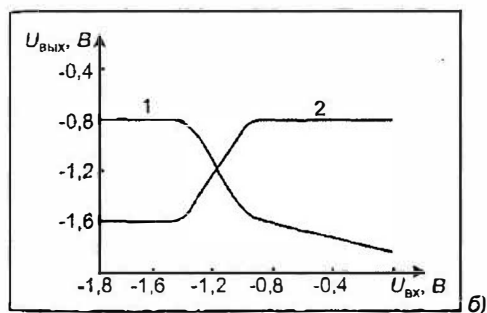
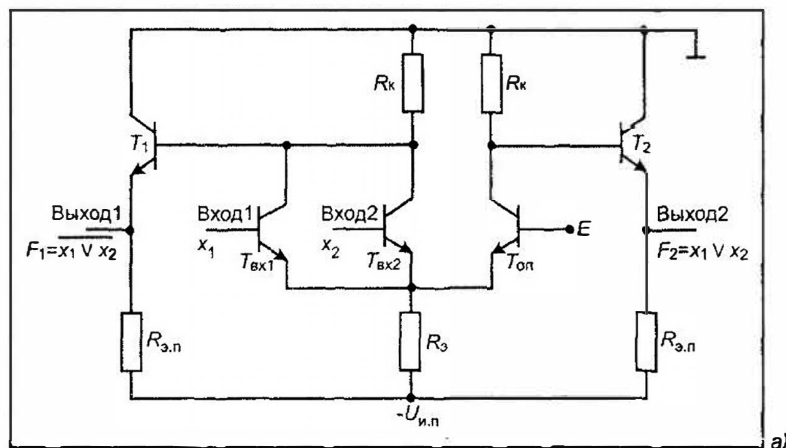


Рис. 7.12. Схема реализована на переключателя тока, к входному транзистору $T_{вх1}$ параллельно подключен входной транзистор $T_{вх2}$: а — схема; б — передаточная характеристика

7.4. Логические элементы, реализованные на МДП-транзисторах

7.4.1. Схемы на транзисторах с каналами одного типа проводимости

К первой группе относятся также логические элементы с использованием МДП-транзисторов. В настоящее время применяются МДП-транзисторы с окисным диэлектриком SiO_2 . В основе МОП-транзисторной логики (НСТЛМ) лежат МОП-транзисторные ключи — инверторы.

Рассмотрим логические элементы одного типа проводимости, например, с индуцированным каналом п-типа. В схемах последовательно с источником питания включают нагрузочный транзистор T_n , используемый как квазилинейный резистор. Для выполнения логических операций применяется транзисторная матрица $T_1, T_2, T_3 \dots T_n$, при последовательном соединении реализующая логическую функцию И — НЕ (рис. 7.13, а).

Если потенциал на входе хотя бы одного из транзисторов T_1, T_2, T_3 меньше порогового напряжения $U_{ин}$, то транзистор остается закрытым. Ток не будут проводить и остальные транзисторы. И только при одновременном отпирании транзисторов происходит переход из закрытого состояния в открытое.

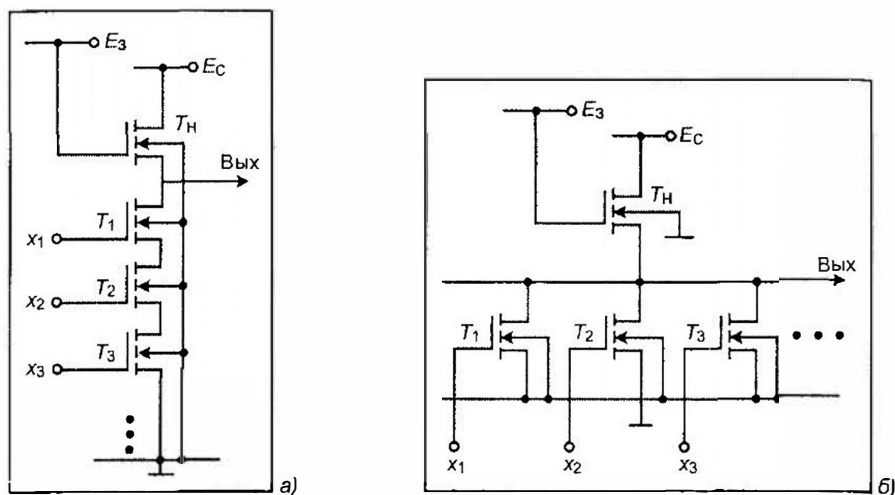


Рис. 7.13. Трехвходовой логический элемент, реализованный на МДП-транзисторах с индуцированным каналом: а — реализация функции И — НЕ; б — реализация функции ИЛИ — НЕ

При параллельном включении МОП-транзисторов, если транзисторы T_1, T_2, T_3 не проводят, на выходе устанавливается высокий потенциал. Когда входной потенциал на затворе хотя бы одного из указанных транзисторов превышает пороговое напряжение $U_{зп\text{пор}}$, то этот транзистор отпирается и выходной потенциал понижается. В этом случае реализуется логическая операция ИЛИ — НЕ (рис. 7.13, б). Быстродействие МОП-логики ограничивается скоростью перезарядки выходной емкости, зависящей от числа нагрузочных элементов логики.

По аналогичной схеме строятся логические элементы на транзисторах со встроенным каналом. Заметим, что каждый из транзисторов T_1, T_2, T_3 вместе с нагрузочным транзистором T_H образуют инвертор. Его передаточная характеристика идентична ранее рассмотренным конструкциям инверторов.

Реализация логических функций с помощью МДП-транзисторов сводится к топологическому управлению межэлектродными проводимостями транзисторов. Ток стока пропорционален межэлектродной проводимости, которая в свою очередь определяется геометрией прибора. Когда МДП-транзистор проводит ток, его межэлектродное сопротивление вместе с сопротивлением нагрузочного резистора образуют делитель напряжения, который определяет величину выходного напряжения. Когда же транзистор заперт, выходное напряжение незначительно отличается от напряжения питания.

Комбинируя последовательное и параллельное соединение МДП-транзисторов можно задать выполнение любых функций.

На рис. 7.14 представлена схема, реализующая функцию

$$F = \overline{x_1 \wedge x_2 \vee x_3}.$$

Часть, относящаяся к функции ИЛИ ($x_2 \vee x_3$), представлена двумя параллельными цепями, а часть И ($x_1 \wedge x_2$) — двумя последовательно включенными транзисторами. Отрицание НЕ обусловлено инверсией входного напряжения на выходе. Обе параллельные цепи должны иметь одинаковые по величине полные сопротивления, что необходимо для поддержания

требуемого соотношения проводимостей логических транзисторов и нагрузочного резистора.

Для сохранения площади маски минимальной, необходимо в большей степени использовать параллельные цепи и меньше последовательные соединения. Схемы на МОП-транзисторах имеют в интегральном исполнении простую конфигурацию, прежде всего, из-за высокого импеданса транзисторов и малого тока через их затвор.

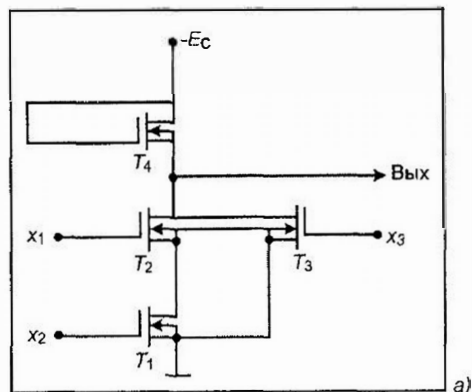
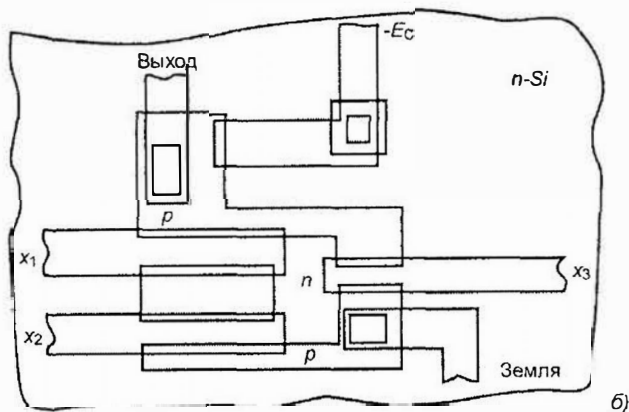


Рис. 7.14. Трехвходовой логический элемент, выполняющий функцию $F = x_1 \wedge x_2 \vee x_3$:

а — схема; б — топология



7.4.2. Схемы на комплементарных транзисторах

Простейшей схемой, реализованной на комплементарных транзисторах, является *комплементарный транзисторный ключ* (рис. 7.15). Если $U_{ax} = 0$, то $U_{зп1} = 0$, а $U_{зп2} = -E_c$ и n-канальный транзистор T_1 закрыт, а p-канальный T_2 открыт.

Ток через транзисторы будет незначительный, т. к. сопротивление закрытого транзистора велико.

Если входное напряжение $U_{ax} > 0$, то $U_{зп1} = E_c$, $U_{зп2} = 0$. В этом случае n-канальный транзистор T_1 открыт, а p-канальный транзистор T_2 закрыт. При этом ток в общей цепи будет по-прежнему мал, потому что уже закрыт p-канальный транзистор.

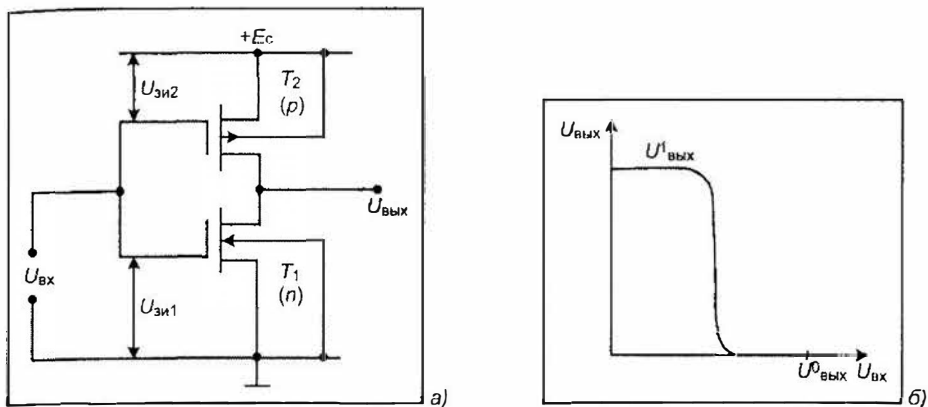


Рис. 7.15. Комплементарный транзисторный ключ (а) и его переключательная характеристика (б)

Важнейшей особенностью комплементарных ключей является тот факт, что они практически не потребляют мощность в обоих состояниях.

Наиболее перспективными микросхемами логики являются интегральные элементы на КМОП-транзисторах. Они отличаются малой мощностью потребления в статическом режиме ($\sim 10^{-9}$ Вт), высоким быстродействием (~ 10 МГц), большой помехоустойчивостью, высокой эффективностью использования источников питания.

В ИС на КМОП-транзисторах логические операции реализуются также путем последовательного включения входных транзисторов для выполнения функции И — НЕ или параллельным их включением для реализации функции ИЛИ — НЕ. На каждый вход при этом требуются два транзистора, образующих ключевой элемент (рис. 7.16).

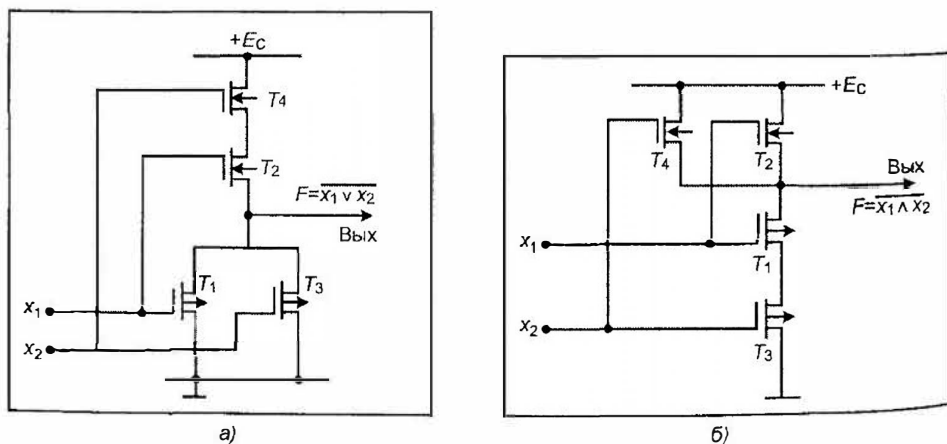


Рис. 7.16. Логический элемент, реализованный на комплементарных ключах с параллельным (а) и последовательным (б) включением логических транзисторов

Следует подчеркнуть закономерность структуры КМОП логических схем, заключающуюся в том, что параллельное соединение одного типа транзисторов сопровождается после-

довательным соединением другого типа. Помимо высокой экономичности КМОП, схемы имеют малые рабочие напряжения ($\sim 2U_0$) и высокое быстродействие.

7.4.3. Логические элементы на арсенид-галлиевых транзисторах

Различают несколько типов логических схем на GaAs.

Логические схемы на полевых транзисторах с непосредственными связями (НСПТ) обычно используют полевые арсенид-галлиевые транзисторы.

На рис. 7.17, а показана базовая схема инвертора с нагрузкой в виде обычного полевого транзистора с затвором. В виде нагрузки могут быть использованы полевые транзисторы с затвором, а также резистивная нагрузка.

Если приложить входное напряжение $U_{вх}$ к затвору переключающего транзистора меньше порогового значения, то транзистор будет в закрытом состоянии. Выходное напряжение составит $U_{вых} \cong E_c$.

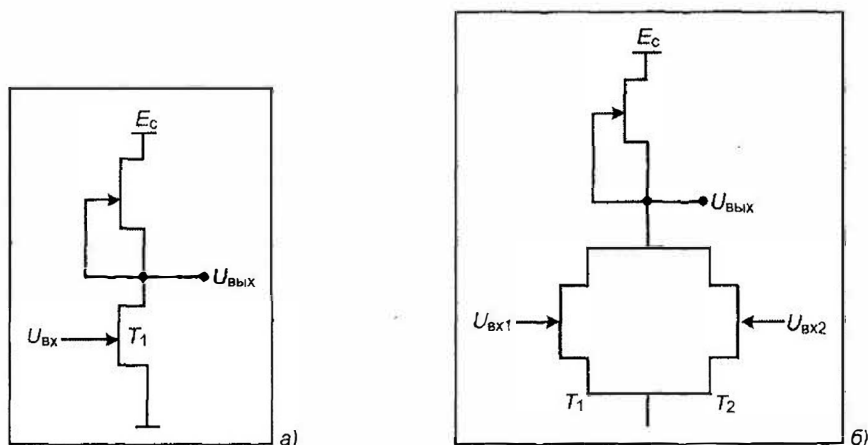


Рис. 7.17. Логические элементы на полевых транзисторах с непосредственными связями: а — инвертор; б — схема, реализующая функцию ИЛИ — НЕ

Когда входное напряжение $U_{вх}$ велико, переключающий транзистор открывается, и выходное напряжение $U_{вых}$ будет мало. Величина выходного напряжения высокого логического уровня ограничивается высотой потенциального барьера в системе "диод—затвор" ПТ. Поэтому НСПТ-схемы характеризуются небольшой величиной напряжений перепада логических сигналов на входе схемы и небольшим запасом помехоустойчивости.

На рис. 7.17, б приведена схема логического элемента ИЛИ — НЕ с коэффициентом объединения по входу равным 2.

Логические схемы на полевых транзисторах с диодами Шоттки (ДШПТ) состоят из транзистора-формирователя верхнего уровня (ФВУ) и транзистора-формирователя нижнего уровня (ФНУ) (рис. 7.18).

ФВУ играет роль нагрузки, а ФНУ соединяет затвор ключевого транзистора с источником питания отрицательной полярности.

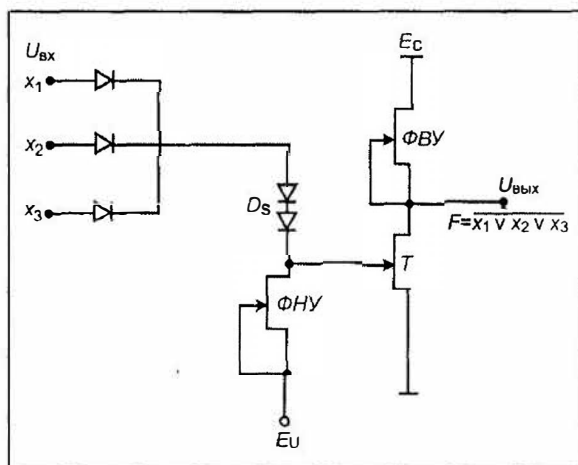


Рис. 7.18. Логическая схема ИЛИ на полевых транзисторах с диодами Шоттки

Диоды сдвига уровня D_s позволяют снизить напряжение на затворе ключевого транзистора T до величины, обеспечивающей отключение транзистора T при малом входном напряжении $U_{вх}$. Для увеличения нагрузочной способности логической схемы выход системы может быть дополнен истоковым повторителем. В ДПШПТ-схемах используются нормально открытые ключевые транзисторы. Это позволяет увеличить логический перепад и, соответственно, помехоустойчивость.

Логический элемент на полевых транзисторах с буферным каскадом (БПТ) представляет собой инвертор статического типа. Другими словами, инверторы БПТ, ДПШПТ, и НСППТ-типов имеют амплитудно-частотную характеристику, аналогичную характеристике фильтра нижних частот.

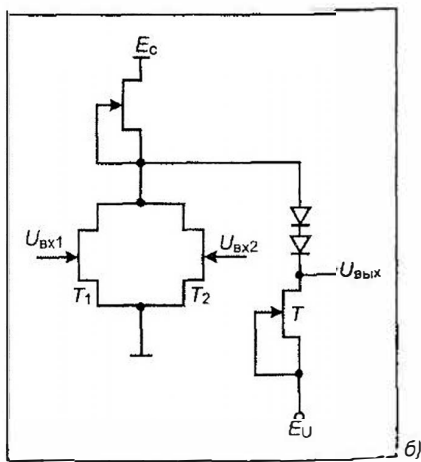
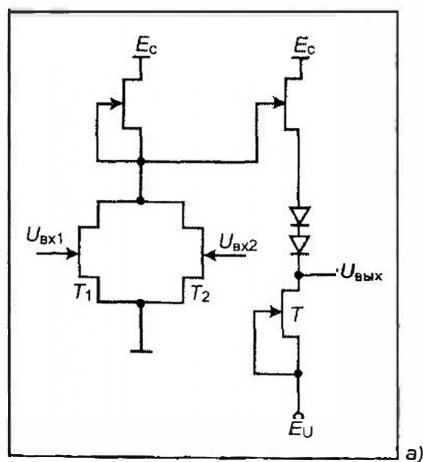


Рис. 7.19. Логические элементы с буферными каскадами на полевых транзисторах:
а — базовая схема инвертора с истоковым повторителем;
б — базовая схема без истокового повторителя

Для работы БПТ-схем требуются два источника питания (рис. 7.19). Базовая схема инвертора включает логический каскад и каскад формирователя сдвига уровня.

В схемах этого типа используются ПТШ, работающие в режиме обеднения. Для согласования входных и выходных уровней напряжения необходим сдвиг уровня выходного напряжения, осуществляемый диодами Шоттки в буферном каскаде (рис. 7.19, а). Число диодов Шоттки определяется напряжением отсечки ключевого транзистора T .

Схемы с меньшими напряжениями отсечки содержат меньшее число диодов сдвига уровня, характеризуются меньшей потребляемой мощностью и меньшей величиной напряжения перепада логических уровней и худшей помехозащищенностью.

Если устранить истоковый повторитель, то можно снизить потребление мощности (рис. 7.19, б). Одновременно ухудшаются времена переключения.

Логические схемы И — НЕ, ИЛИ — НЕ на основе ПТ с длиной затвора 1 мкм имеют время задержки сигнала порядка 100 пс при потребляемой мощности 40 мВт.

7.5. Логические элементы на БикМОП-транзисторах

Разработанная технология совмещения биполярных и КМОП транзисторных структур позволила создать варианты БикМОП логических устройств.

На рис. 7.20, а приведена схема инвертора, реализованная на КМОП и биполярных транзисторах.

Транзисторы T_1 и T_2 формируют КМОП-инвертор с той разницей, что параллельно их каналам включены сопротивления R_1 и R_2 . Эти сопротивления по величине сопоставимы с сопротивлением каналов транзисторов в открытом состоянии.

Выходной каскад сформирован на базе биполярных транзисторов T_3 и T_4 , эмиттерные переходы которых подсоединены к резисторам.

В статическом состоянии токи через транзисторы T_1 и T_2 и, соответственно, через резисторы R_1 и R_2 малы, и поэтому транзисторы T_3 и T_4 закрыты.

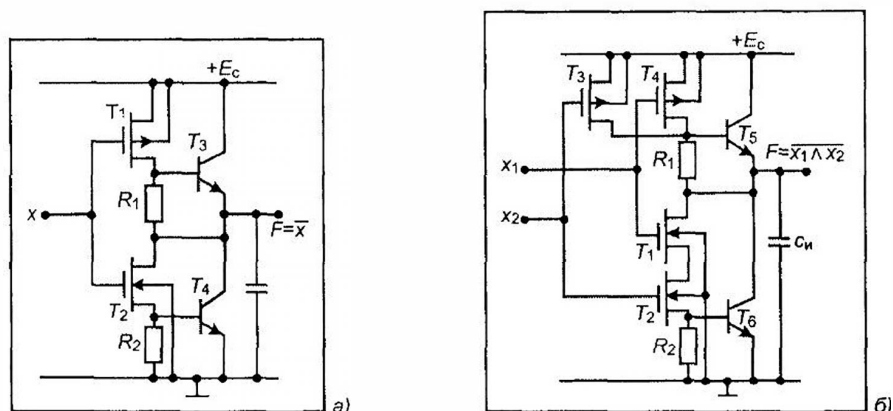


Рис. 7.20. Логические элементы на БикМОП транзисторных структурах:
а — инвертор; б — схема, реализующая функцию И — НЕ

Если на входе X напряжение низкого логического уровня, то р-канальный транзистор T_1 открыт. На выходе инвертора будет высокий логический уровень, и конденсатор C_n будет заряжен.

Если на вход подать высокий логический уровень, откроется n-канальный транзистор T_2 и начнется разряд емкости через резистор R_2 и эмиттерный переход биполярного транзистора T_4 . Часть разрядного тока откроет транзистор T_4 и будет происходить перезарядка емкости нагрузки.

Аналогично происходит процесс переключения при изменении входного напряжения от высокого логического уровня к низкому.

На рис. 7.20, б приведена схема логического элемента И — НЕ. На входе традиционно расположены последовательно соединенные транзисторы, характерные для схем типа И. Транзисторы T_3 и T_4 комплементарны к T_1 и T_2 и выполняют роль нагрузочных резисторов. Биполярные транзисторы T_5 и T_6 на выходе позволяют усилить сигналы, а также нейтрализовать влияние емкостной нагрузки. Именно емкостная нагрузка является фактором ограничения быстродействия КМОП-структур.

7.6. Сравнительный анализ логических элементов

Логические элементы характеризуются конкретным набором параметров, определяющих целесообразность их применения в определенных схемах.

Приведем далее основные параметры, которые помогают выбрать тот или иной тип логического элемента для реализации определенной задачи.

Средняя потребляемая мощность $P_{\text{пот}}$ определяется как

$$P_{\text{пот}} = (P_{\text{пот}}^0 + P_{\text{пот}}^1) / 2,$$

где $P_{\text{пот}}^0$ и $P_{\text{пот}}^1$ — мощности потребляемые логическим элементом в состоянии логического нуля и логической единицы, соответственно.

Среднее время задержки сигнала $t_{\text{зд}}$ определяется как

$$t_{\text{зд}} = (t_{\text{зд}}^{1 \rightarrow 0} + t_{\text{зд}}^{0 \rightarrow 1}) / 2,$$

где $t_{\text{зд}}^{1 \rightarrow 0}$ и $t_{\text{зд}}^{0 \rightarrow 1}$ — время задержки между фронтами выходного и входного сигналов при переходе из состояния логической единицы в логический ноль, и наоборот.

В табл. 7.1 приведен сравнительный анализ параметров рассмотренных логических элементов.

Средняя работа переключения $A_{\text{ср}}$ — определяется как произведение средней мощности переключения на среднюю продолжительность одного переключения. Этот параметр характеризует экономичность и быстродействие логического элемента. Уменьшение этого ключевого параметра позволяет судить о прогрессе в технологии и схемотехнике интегральных схем. Этот параметр называют фактором качества.

Таким образом, в зависимости от выходных параметров интегральной схемы можно выбирать нужные элементы по основным характеристикам: $P_{\text{пот}}$, $t_{\text{зд}}$, $A_{\text{ср}}$.

На диаграмме, показанной на рис. 7.21, приведена зависимость времени задержки от мощности рассеяния для микросхем различных типов.

Видно, что наиболее перспективными по фактору качества являются арсенид-галлиевые схемы.

Таблица 7.1

Тип логики	$P_{\text{пот}}, \text{ мВт}$	$t_{\text{д}}, \text{ нс}$	$A_{\text{ср}}, \text{ пДж}$
РТЛ	20—50	25	250
ДТЛ	8—10	25	220
ТТЛ	1—20	5—20	50—100
ТТЛШ		2—10	10—50
ЭСЛ	20—50	0,5—2	20—50
И ² Л	0,01—0,1	10—100	0,2—2
МОП	1—10	20—200	50—200
КМОП	0,01—0,1	10—50	0,5—5,0
БиКМОП	0,01—0,1	2—10	2—20
МЕР	0,1—0,5	0,15—0,5	0,1—0,5

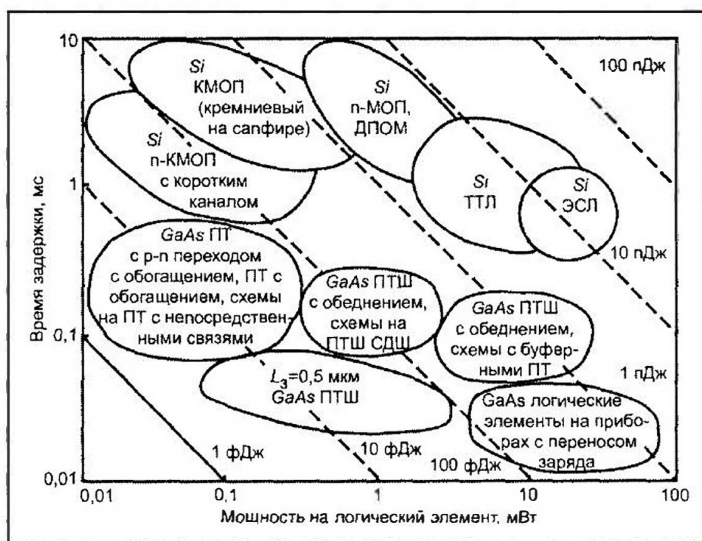
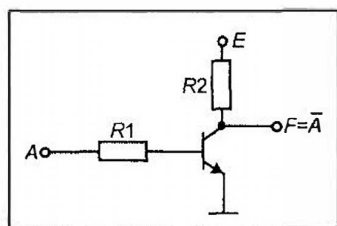


Рис. 7.21. Зависимость времени задержки от мощности рассеяния для различных типов микросхем

Задачи и упражнения

В упражнениях 7.1—7.6 соберите схемы логических элементов в программе Electronics Workbench и проверьте соответствующие таблицы истинности:

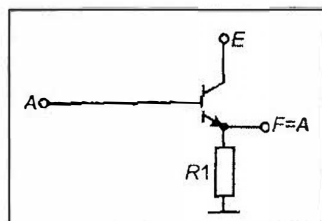
7.1.



A	F
0	1
1	0

$$F = \bar{A}$$

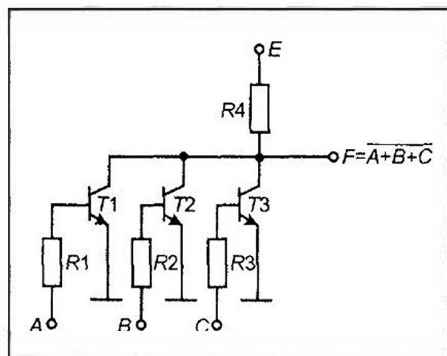
7.2.



A	F
0	0
1	1

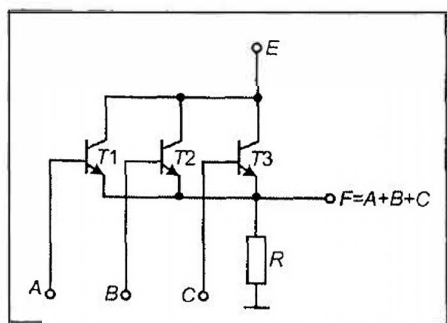
$$F = A$$

7.3.



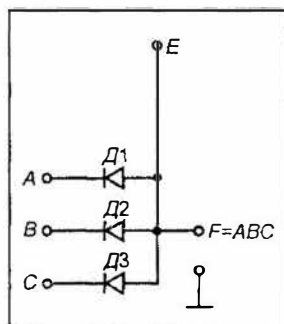
A	B	C	F
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	0

7.4.



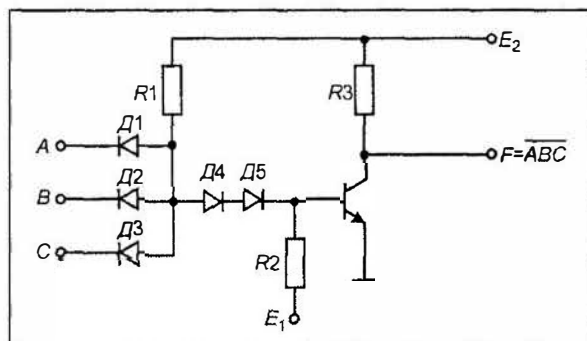
A	B	C	F
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

7.5.



A	B	C	F
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

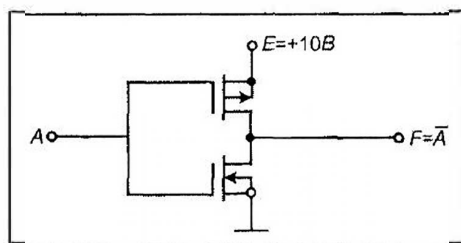
7.6.



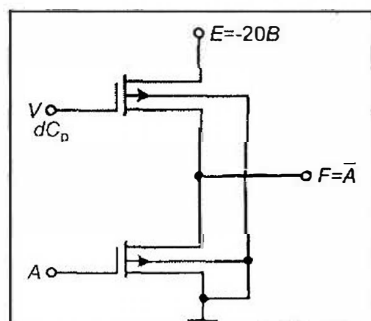
A	B	C	F
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

В упражнениях 7.7—7.11 соберите схемы логических элементов в программе Electronics Workbench и составьте соответствующие таблицы истинности:

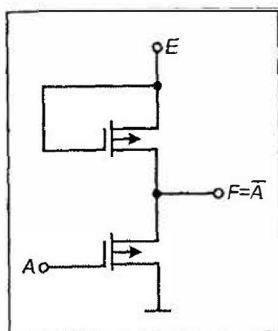
7.7.



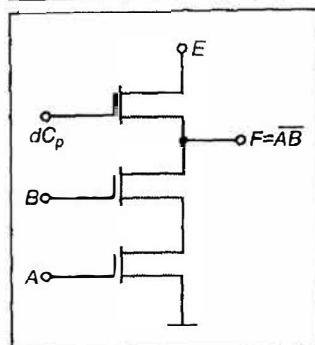
7.8.



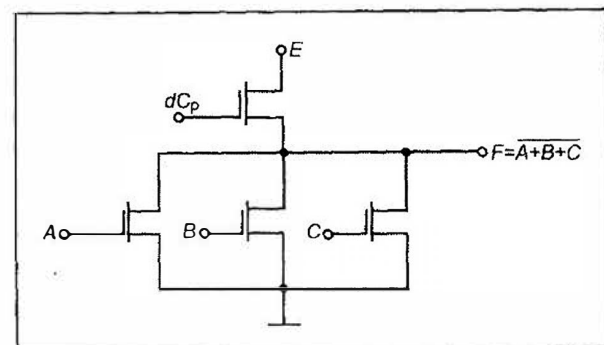
7.9.



7.10.



7.11.



Контрольные вопросы

Что такое интегральная схема?

По каким признакам можно классифицировать интегральные схемы?

Что такое степень интеграции?

Что такое схемы частного применения?

Приведите примеры условного обозначения ИС.

Что такое логический элемент ИС?

Что такое логическая ИС комбинационного типа?

Что такое логическая ИС последовательностного типа?

9. Что такое передаточная характеристика логической схемы?
10. Что такое логическая ИС комбинационного типа?
11. Как работает логическая ячейка типа И²Л? Какую логическую функцию она выполняет?
12. Как работает логическая ячейка типа ДТЛ? Какую логическую функцию она выполняет?
13. Как работает логическая ячейка типа ТТЛ? Какую логическую функцию она выполняет?
14. Как работает логическая ячейка типа ТТЛЦ? Какую логическую функцию она выполняет?
15. Как работает логическая ячейка типа ЭСЛ? Какую логическую функцию она выполняет?
16. Расскажите о работе КМОП-инвертора.
17. Как формируется логическая ячейка типа И — НЕ на МОП-транзисторах?
18. Как формируется логическая ячейка типа ИЛИ — НЕ на МОП-транзисторах?
19. Как формируется логическая ячейка на КМОП-транзисторах?
20. Расскажите о логических элементах на основе GaAs структур.
21. Логические элементы на БиКМОП-структурах.

Рекомендуемая литература

1. Аваев Н. А., Наумов Ю. Е., Фролкин В. Т. Основы микроэлектроники: Учебное пособие для вузов. — М.: Радио и связь, 1991.
2. Алексенко А. Г. Основы микросхемотехники. — 3-е изд. — М.: ЮНИМЕДИАСТАЙЛ, 2002.
3. Карлашук В. И. Электронная лаборатория на IBM PS. Программа Electronics Workbench и ее применение. — 3-е изд. — М.: СОЛОН-Пресс, 2003.
4. Степаненко И. П. Основы микроэлектроники: Учебное пособие для вузов. — 2-е изд. — М.: Лаборатория базовых знаний, 2000.

8. Запоминающие устройства

8.1. Классификация запоминающих устройств

Запоминающее устройство (ЗУ) в вычислительных системах предназначено для записи, хранения и считывания информации. ЗУ в интегральных схемах являются важнейшими элементами цифровой обработки и хранения сигналов. ЗУ, прежде всего, характеризуются емкостью памяти, выраженной в битах. По функциональному назначению основными видами памяти являются оперативные, постоянные и перепрограммируемые ЗУ.

Оперативные запоминающие устройства (ОЗУ) предназначены для быстрого ввода и вывода (записи и считывания) информации в любом, заранее заданном месте памяти. Синонимом ОЗУ является термин — память с произвольной выборкой. Международное обозначение — RAM (Random Access Memory).

Постоянные запоминающие устройства (ПЗУ) предназначены в основном для считывания ранее записанной в них информации. Запись в этом виде ЗУ производится редко с целью постоянного хранения часто используемых данных. Этот класс называется также памятью только для считывания и классифицируется как ROM (Read Only Memory).

Перепрограммируемые запоминающие устройства (РПЗУ) предназначены для частичного и многократного изменения информации в ПЗУ. Они предназначены для внесения необходимых коррективов в постоянную память. К этому классу относятся и перепрограммируемые потребителем запоминающие устройства (ППЗУ). Международное обозначение — PROM (Programmable Read-Only Memory).

На рис. 8.1 приведена примерная иерархия ЗУ различной емкости и их быстродействие.

Центральный блок микропроцессора и входящие в него сдвиговые регистры имеют малую емкость, но большое быстродействие.

Оперативная память вычислительного устройства реализуется на буферном ЗУ большой емкости (ленты, диски и т. п.)

Структура запоминающего устройства, реализованного по полупроводниковой технологии, состоит из накопителя и схем управления. Центральная часть полупроводникового ЗУ представляет собой матрицу, состоящую из запоминающих ячеек, вертикальных и горизонтальных, адресных и разрядных линий.

Накопитель является основой запоминающего устройства, в котором хранится информация в виде двоичных кодов.

Периферия или схемы управления предназначены для ввода и вывода данных. Схемы управления, как правило, состоят из дешифраторов, усилителей, регистров, коммутаторов и других схем, реализованных по совмещенной полупроводниковой технологии. Их работа подробно разбирается в курсах по схемотехнике.

Рассмотрим устройство накопителя, способ его реализации на различных транзисторных структурах (рис. 8.2).

Накопитель имеет форму матрицы, в узлах которой находятся ячейки памяти. Ячейки памяти соединяются с адресной и разрядной шинами с помощью ключа. Каждая ячейка памяти может хранить один бит информации — логический ноль или единицу.

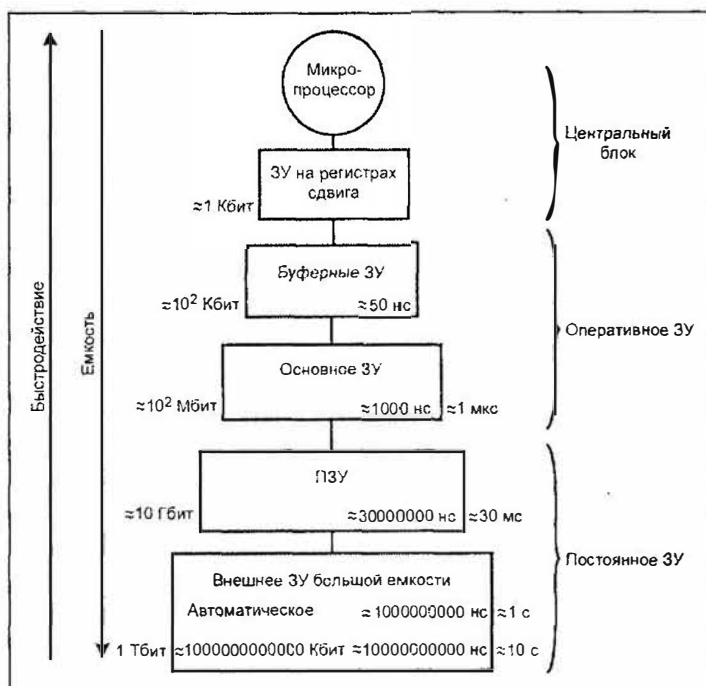


Рис. 8.1. Иерархия ЗУ в вычислительном комплексе

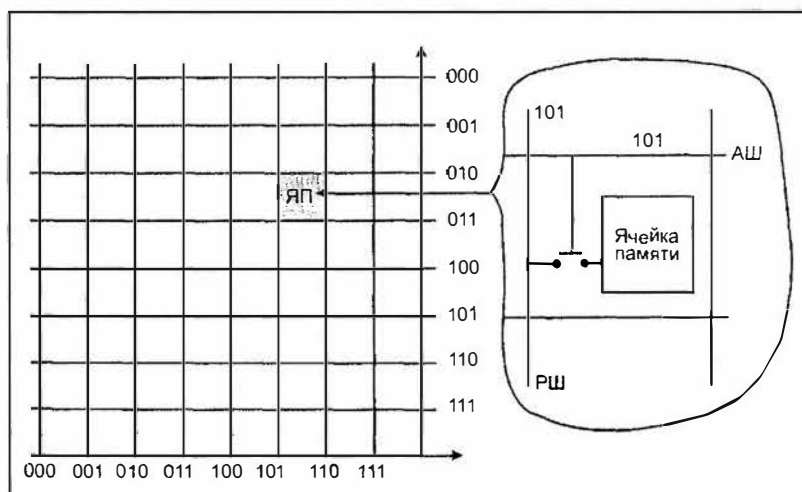


Рис. 8.2. Матричная организация накопителя

Емкость ЗУ определяется количеством ячеек памяти.

Каждая ячейка связана с одной или двумя *адресными* (АШ) и *разрядными* (РШ) шинами. Если используется одна адресная и одна разрядная шины, то речь идет о *динамическом* типе памяти. Если каждая ячейка связана с двумя адресными и разрядными шинами, то речь идет о памяти *статического* типа. В первом способе информацию нужно периодически (во время действия импульсов питания) регистрировать. Во втором способе она хранится постоянно на период подключения источника питания.

Если на определенную адресную шину подать сигнал двух координат адреса необходимой ячейки памяти, то произойдет коммутация по разрядной шине, можно будет записать бит информации или его считать, если он уже записан. Разрядная шина может записать "0" ($PШ_0$) или "1" ($PШ_1$). Адресные и разрядные шины первоначально подключены ко всем ячейкам памяти.

В зависимости от конструкции ячейки памяти, способов ее коммутации с адресными и разрядными шинами, используемых транзисторных структур, различают различные типы ЗУ.

Приведем конструкции наиболее распространенных типов.

8.2. Запоминающие устройства на биполярных транзисторах

Простейшие ячейки памяти на биполярных транзисторных структурах можно реализовать различными способами, в которых используется емкость $p-n$ -перехода для хранения заряда, соответствующего логическому нулю или единице. На рис. 8.3 приведены примеры использования диодов на базе транзисторной структуры для хранения информации.

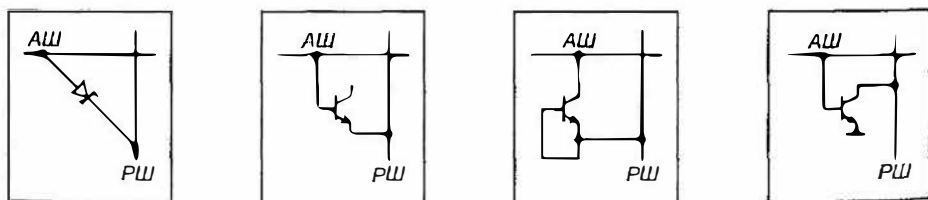


Рис. 8.3. Варианты использования емкости $p-n$ -перехода транзисторной структуры

Такие ОЗУ ненадежны вследствие рассасывания заряда токами утечки. Поэтому наибольшее распространение получили ЗУ, ячейки которых способны хранить один заряд двоичной информации. Такие бистабильные ячейки могут быть реализованы на триггерных схемах и различных типах биполярных транзисторных структур: ТТЛ, ТТЛШ, ЭСЛ, И²Л.

Многоэмиттерные транзисторы удобно использовать в конструкции ячейки памяти ЗУ статического типа. Типовая структура накопителя интегральной схемы ЗУ, реализованная на двух эмиттерных транзисторах T_1 и T_2 , приведена на рис. 8.4, а.

Ячейка памяти состоит из двух транзисторов с перекрестными связями. Вторые эмиттеры соединены с шиной питания. Первые эмиттеры соединены с разрядными шинами $PШ_0$ и $PШ_1$, соответственно, и используются для записи и считывания информации. Коллекторы

транзисторов соединены через резистор с плюсовой массой питания, выполняющей также функции адресной шины $АШ_1$.

В режиме хранения на разрядной шине $РШ_1$ устанавливается напряжение хранения, положительное относительно общей шины микросхемы.

В одном устойчивом состоянии транзистор T_1 открывает, а T_2 закрывает. В другом устойчивом состоянии наоборот.

В режиме считывания повышается напряжение как на шине $АШ_2$, так и на шине $АШ_1$. Если транзистор T_1 открыт, а транзистор T_2 закрыт, то ток в управляющем эмиттере $Э_{22}$ равен нулю, напряжение на шине $РШ_1$ не меняется. В управляющем эмиттере транзистора T_1 появится ток считывания. Напряжение на шине $РШ_0$ повысится. На шинах $РШ_0$ и $РШ_1$ возникнет разность напряжений, которая считывается усилителем.

Ячейка памяти, реализованная на структурах интегральной логики (рис. 8.4, б), работает примерно так же, как и вышеописанная.

Постоянное запоминающее устройство имеет аналогичную с ОЗУ матрицу памяти запоминающих ячеек. ЗУ называется постоянным потому, что в каждой ячейке раз и навсегда записаны логические нуль или единица. Единицы нужно регенерировать в зависимости от типа памяти. А нуль он всегда нуль. Чтобы получить нули в нужных ячейках памяти, необходимо закрыть доступ в нужную ячейку. Это делается, например, с помощью так называемых перемычек.

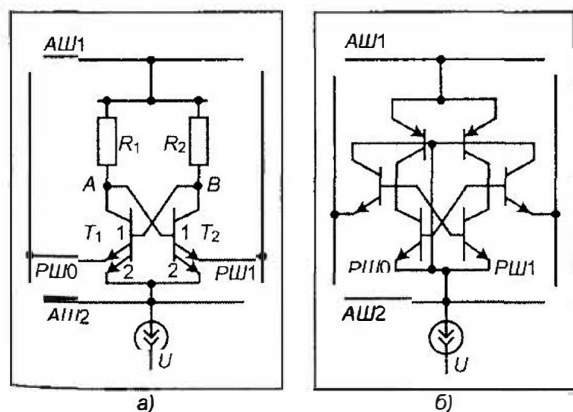


Рис. 8.4. Схема ячейки памяти ОЗУ, реализованная на многозмиттерном транзисторе (а) и на элементах И²П (б)

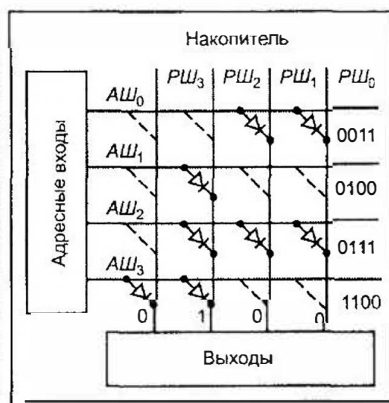


Рис. 8.5. Схема постоянного диодного ЗУ

На рис. 8.5 представлена типичная схема ПЗУ, реализованная на диодах. Она состоит из накопителя, адресных входов и выходов.

Каждая адресная шина предназначена для ввода определенного кода: 0011, 0100, 0111, 1100 и т. д. Запись информационного кода осуществляется с помощью диодов, которые присоединены между адресными шинами и теми разрядными шинами, на которых должна храниться логическая единица.

Например, подадим на адресную шину $АШ_1$ логическую единицу в виде соответствующего напряжения в выбранной строке.

Напряжение с этой адресной шины поступит только на разрядную шину $РШ_2$. На других шинах $РШ_0$, $РШ_1$ и $РШ_3$ будет отсутствовать напряжение. На выход пойдет только код 0100.

Пережигание диодов (перемычек) производится в специальных устройствах в местах расположения логических нулей. На рис. 8.5 штриховыми линиями показаны перегоревшие перемычки.

8.3. Запоминающие устройства на МДП-транзисторах

Наиболее распространенным типом ЗУ являются схемы памяти на МДП-структурах. Их основными преимуществами перед ЗУ на биполярных структурах являются:

- ☐ малая потребляемая мощность;
- ☐ высокая степень интеграции;
- ☐ сравнительная простота технологии (число технологических операций на 30% меньше);
- ☐ низкая стоимость при больших объемах производства;
- ☐ высокий запас помехоустойчивости;
- ☐ энергонезависимость ЗУ.

По режимам работы накопительной матрицы ЗУ различают следующие ИС памяти:

- ☐ динамические ЗУ, в которых информация сохраняется в накопителе в виде зарядов на емкостях, входящих в состав элементов памяти, а регенерация зарядов происходит периодически от источника питания;
- ☐ статические ЗУ, в которых сохранение информации в накопителе обеспечивается с помощью источников питания;
- ☐ квазистатические, в которых информация в накопителе сохраняется в виде зарядов, а их регенерация происходит в определенные периоды, в течение которых происходит считывание и повторная запись считанной информации в каждом элементе накопителя.

Наиболее распространены одностранзисторные ячейки памяти (рис. 8.6, а). Схема состоит из транзистора T и запоминающего конденсатора C .

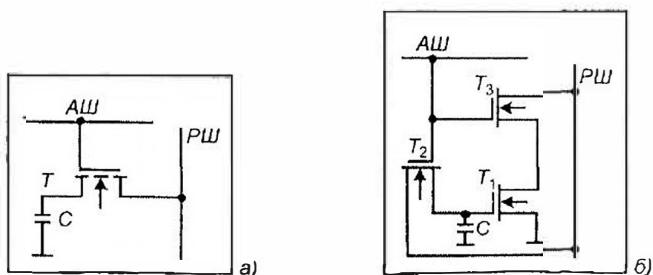


Рис. 8.6. Схема динамических запоминающих элементов, реализованных на одном транзисторе (а) и трех транзисторах (б)

В режиме хранения напряжение на $AШ$ близко к нулю. Транзистор T_1 закрыт. Конденсатор C отключен от разрядной шины. В конденсаторе хранится информация в виде разряда $Q = CU$. В случае хранения логической единицы заряд $Q = CU$ будет постепенно уменьшаться за счет токов утечки в подложку.

При хранении логического нуля заряд $Q = CU^0$, и конденсатор будет слегка заряжаться предпороговым током транзистора от положительно заряженной $PШ$.

Для ЗУ требуется регулярный процесс регенерации для всех элементов строки.

В режиме записи на разрядной шине устанавливается напряжение U^1 или U^0 , а затем подается положительный импульс на $AШ$. Транзистор открывается и на обкладках конденсатора устанавливается нужное напряжение U^1 или U^0 .

В режиме считывания $PШ$ подключена к входу усилителя считывания с высоким входным сопротивлением. При поступлении импульса выборки на $AШ$ транзистор элемента памяти открывается и происходит считывание с разрушением информации.

Для того чтобы избежать разрушения информации во время считывания, разработана трехтранзисторная схема динамического ЗУ. В этой схеме накопительный конденсатор C изолирован от разрядной шины, и считывание информации происходит без ее разрушения. Однако утечки зарядов избежать не удастся и требуется периодическая регенерация информации. Это достигается путем подачи высокого потенциала на $AШ$ и специального подключения $PШ$.

БИС ОЗУ на МДП-транзисторах статического типа приведена на рис. 8.7.

Конструкция ячейки на одногигиных МОП-транзисторах с р-каналом имеет классическую структуру RS-триггера с управляющими ключами T_5 и T_6 , которые заперты в нормальном состоянии и ячейки отключены от разрядных шин $PШ$. При поступлении отрицательного импульса $-E_c$ на адресную шину ключи T_5 или T_6 открываются и подключают ячейку к разрядным шинам. На разрядные шины поступают уровни Q и Q^+ , записанные в ячейке (рис. 8.7, а). В режиме записи на адресную шину также подается импульс, в то время как на разрядные шины подаются взаимно противоположные уровни. Другими словами, роль тактового импульса играет импульс на адресной шине в обоих режимах. Это самые простые по схемотехнической и технологической реализации реализации схемы ОЗУ. Однако они обладают низким быстродействием и малой степенью интеграции.

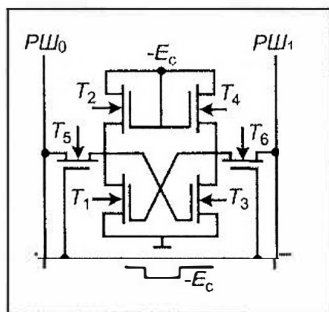


Рис. 8.7. Ячейка ОЗУ статического типа, реализованная на МДП-транзисторах

Использование КМОП-транзисторов в схемах ячеек памяти позволяет существенно повысить быстродействие, получить практически нулевую рассеивающую мощность в статическом режиме и повышенную помехоустойчивость.

Разновидность БИС ПЗУ определяется типом ячейки памяти, способом записи и стирания информации. Состав этих устройств памяти аналогичен БИС ПЗУ с двухкоординатной выборкой. Обрамление матричного накопителя, как правило, состоит из дешифраторов строк и столбцов, адресных формирователей, усилителей считывания и других схем управления. Запись информации в ПЗУ осуществляется на этапе изготовления кристаллов с помощью заказного последнего фотошаблона путем формирования нужной конфигурации металлизированной разводки. Можно воспользоваться также процессом селективного вскрытия контактных окон под металлизацию. Такой тип ПЗУ называют масочным, потому что программирование осуществляется маской — последним фотошаблоном.

В программируемые ИС ПЗУ (или ППЗУ) нужная информация записывается электрическим способом, путем пережигания перемычек, либо пробоем р—п-перехода. Перемычки изготавливаются из нихрома, поликристаллического кремния или алюминия.

ППЗУ на основе МДП-транзисторов обладает достаточно большой информационной емкостью и низкой потребляемой мощностью.

Для широкого использования БИС ЗУ необходимо создать память, способную многократно перепрограммироваться и сохранять информацию при отключенном питании. На основе МДП-структур разработаны репрограммируемые постоянные запоминающие устройства (РПЗУ), допускающие многократную перезапись и хранение информации при отключенном питании. В основе лежит идея создания бистабильных МДП-транзисторов, которые могут находиться в одном из двух состояний, соответствующих хранению логической единицы или нуля.

РПЗУ в зависимости от физического принципа работы подразделяются на два класса. К первому классу относятся устройства на транзисторных структурах, в которых используется захват заряда на естественной границе раздела двух диэлектриков. Конструкция транзистора с диэлектриком, имеющим структуру сэндвича из слоев нитрида кремния и диоксида кремния (МНОП-транзистор), представлена на рис. 8.8, а.

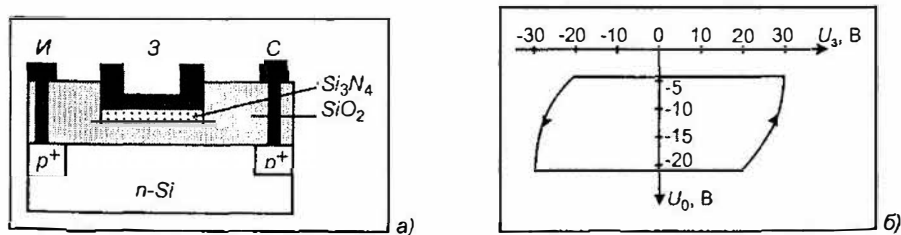


Рис. 8.8. Структура МНОП-транзистора (а) и зависимость порогового напряжения от напряжения на затворе (б)

В основе работы МНОП-транзистора лежит явление накопления заряда на границе нитридного и оксидного слоев. Накопление происходит из-за несимметричных токов проводимости в нитридном и диоксидном слоях. При большом отрицательном напряжении на затворе, на границе накапливается положительный заряд, способствующий увеличению отрицательного порогового напряжения (рис. 8.8, б). При большом положительном напряжении на затворе, на границе накапливается отрицательный заряд, что приводит к уменьшению отрицательного порогового напряжения.

Благодаря гистерезисной зависимости порогового напряжения от напряжения на затворе, МНОП-транзистор с помощью управляющих импульсов можно переводить из рабочего

состояния в запертое и обратно. Таким образом, можно осуществлять доступ в любую ячейку памяти накопителя ЗУ.

Ко второму классу РПЗУ относятся устройства на транзисторных структурах с плавающим затвором. В этих конструкциях управляющий заряд хранится в тонком проводящем слое — *плавающем затворе*, расположенном в диэлектрике в затворной части МДП-структуры. Заряд, аккумулированный на плавающем затворе, позволяет изменять пороговое напряжение, необходимое для открывания затвора.

На рис. 8.9 приведена схема МОП-транзистора с плавающим затвором, выполненным из поликремния. Если на затвор МОП-транзистора подать высокое напряжение ($U_z \sim 25$ В), то формируется канал. Под воздействием поля, формируемого напряжением стока, в канале образуются *горячие* электроны. Они проникают сквозь туннельно-тонкий диэлектрик на плавающий затвор. На затворе формируется отрицательный заряд, способный перекрыть канал "исток—сток", и транзистор закрывается (рис. 8.9, а).

Если транзистор осветить ультрафиолетовым светом, энергия квантов которого больше ширины запрещенной зоны полупроводника $h\nu > E_g$, то плавающий затвор будет очищен от осевших там электронов. Электроны, получив энергию излучения, способны преодолеть потенциальный барьер между плавающим затвором и каналом. Канал транзистора будет восстановлен и транзистор будет открыт (рис. 8.9, б).

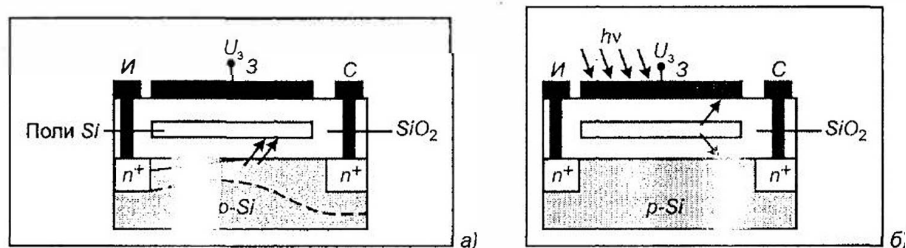


Рис. 8.9. Схема работы электрически программируемого ЗУ (а) со стиранием информации УФ-излучением (б)

Такая технология получения РПЗУ получила название FAMOS (Floating gate avalanche injection MOS) или МОП-транзистор с плавающим затвором и лавинной инжекцией заряда.

Другой конструкцией электрически репрограммируемых ЗУ является ЭСППЗУ — электрически стираемые программируемые ПЗУ. Технология их производства называется FLOTOX (Floating gate tunnel-oxide) или плавающий затвор с туннелированием в окисле.

В транзисторной структуре этого типа плавающий поликремниевый затвор отделен от диффузионной области n -типа тонким слоем диэлектрика (~ 200 Å). Если на затвор подать высокое напряжение (~ 20 В), то на плавающий затвор начнут дрейфовать туннельные электроны из стока. На плавающем затворе сформируется статический заряд (рис. 8.10, а). Если на затвор подать низкое напряжение, а на сток высокое, то произойдет обратное туннелирование электронов с плавающего затвора. Произойдет процесс стирания информации (рис. 8.10, б).

Процессы программирования и стирания информации имеют одинаковую скорость и отличаются только местом приложения напряжения и направлением движения электронов.

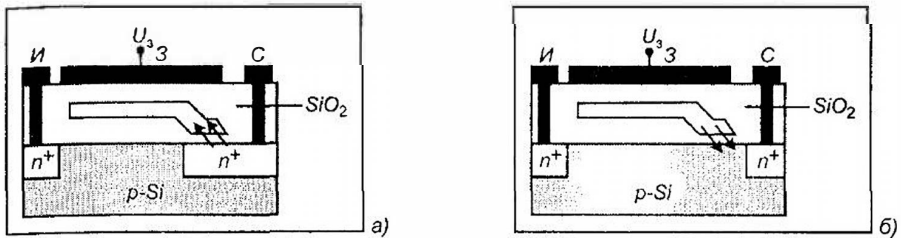


Рис. 8.10. Схема работы электрически стираемого ЗУ в процессе программирования (а) и стирания информации (б)

Анализ разработок БИС ППЗУ и РПЗУ на МДП-транзисторах показывает, что разнообразие схематического построения БИС зависит от элементарной базы накопителя, другими словами, от конструктивного решения ячейки памяти.

Именно выбор ячейки памяти определяет величину напряжения программирования, время хранения информации, число циклов перезаписи информации и возможность использования одного или нескольких источников питания.

Максимальным временем хранения информации после отключения источника питания обладают РПЗУ, в основе конструкции которых лежат транзисторные структуры с плавающим затвором.

8.4. Запоминающие устройства на арсенид-галлиевых структурах

В настоящее время разработаны различные структуры ЗУ на арсенид-галлиевых транзисторных структурах. На рис. 8.11, а представлена статическая ячейка ●ЗУ, реализованная на полевом транзисторе (ПТ) и туннельном диоде (ТД).

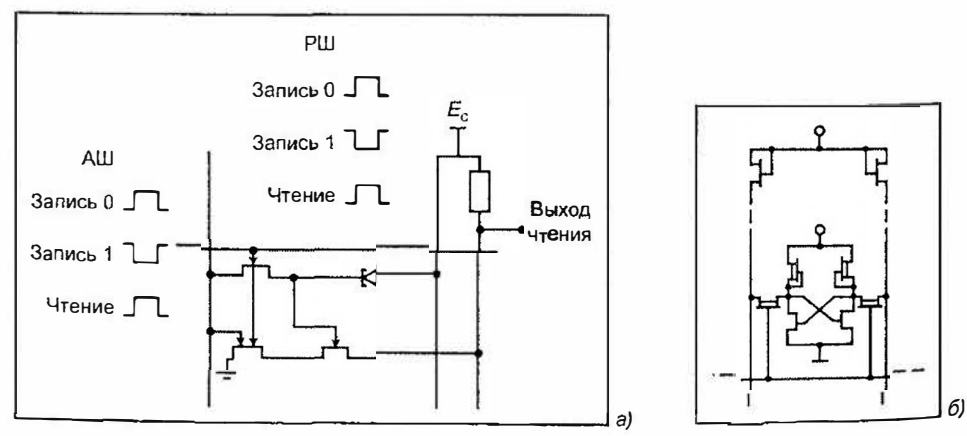


Рис. 8.11. Ячейки ЗУ на основе ПТ и ТД (а) и статического типа на основе триггера (б)

Отрицательное сопротивление туннельного диода позволяет достичь минимального потребления мощности схем в стационарном состоянии и в то же время значительный ток

переключения. Время задержки сигнала составило $\sim 10^2$ пс, что соответствует энергии переключения ~ 1 фДж.

В схеме ячейки используется переключающий транзистор, работающий в режиме обеднения, а также двухзатворный и буферный ПТ, работающий в режиме обогащения. Показана комбинация сигналов, приложенных в *III* и *III*, с помощью которых вырабатываются сигнальные записи, стирания и вывод данных из ячейки.

Другой пример ЗУ на арсенид-галлиевых структурах приведен на рис. 8.11, б. Ячейка памяти представляет собой схему триггера с перекрестными связями на шести транзисторах. Переключающие ПТ имеют длину затвора ~ 2 мкм. Время выборки из ЗУ составляет ~ 4 нс.

Приведенные триггеры свидетельствуют о весьма сложном процессе создания интегральных схем на арсенид-галлиевых транзисторных структурах.

Задачи и упражнения

8.1. Организация ЗУ на интегральных схемах

На рис. 8.12 представлены интегральные схемы ОЗУ, ПЗУ и шинные формирователи для организации ЗУ. В табл. 8.1—8.5 приведены их основные характеристики.

1. Для каких целей используются ОЗУ и ППЗУ?
2. Какие типы ОЗУ вы знаете?
3. Какие выходные каскады используются в микросхемах статических ОЗУ и ППЗУ для параллельного соединения выводов?
4. Какие выходные каскады используются в шинных формирователях для присоединения к двунаправленной шине?
5. Разрушается ли информация в микросхемах статических ОЗУ и ППЗУ при выключении источника питания, при считывании информации?
6. Как присоединяются микросхемы ОЗУ и ППЗУ с открытым коллекторным входом к двунаправленной шине?
7. Что такое время выборки адреса в ОЗУ и ППЗУ?
8. Можно ли к двунаправленной шине присоединить одновременно статическое ОЗУ и ППЗУ?
9. Можно ли использовать ППЗУ для выполнения логических операций?
10. Можно ли использовать ППЗУ для выполнения арифметических операций?

Таблица 8.1. Назначение выводов микросхем ОЗУ КМ132РУ8А, КМ132РУ8Б

Обозначение	Назначение
$A_0 \dots A_9$	Адресные входы
$DI_0 \dots DI_3$	Вход — выход данных
\overline{CS}	Выбор микросхемы
$\overline{WR/RD}$	Сигнал запись — считывание

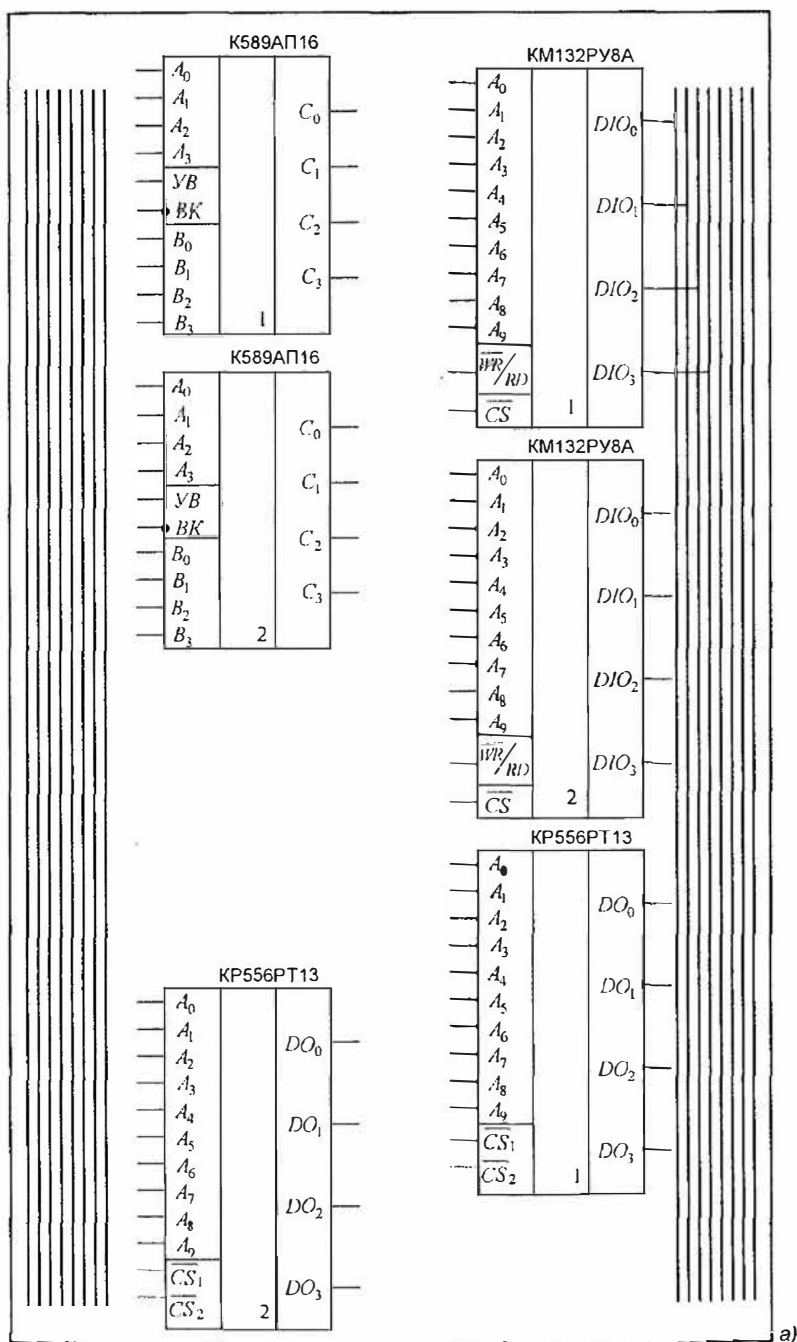


Рис. 8.12. Интегральные схемы ОЗУ, ПЗУ (а) (см. продолжение)

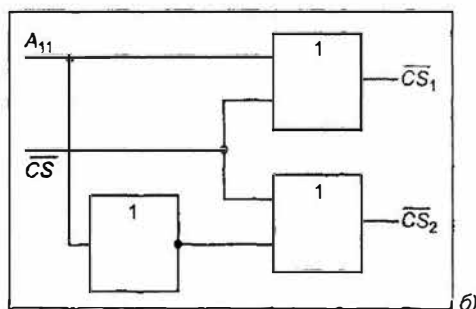


Рис. 8.12. Шинные формирователи для организации ЗУ (б)

Таблица 8.2. Таблица истинности микросхем ОЗУ КМ132РУ8А, КМ132РУ8Б

\overline{CS}	$\overline{WR/RD}$	$A_0...A_9$	$DIO_0...DIO_3$	Режим работы
1	Ф	Ф	R_{off}	Хранение
0	0	A	0	Запись 0
0	0	A	1	Запись 1
0	1	A	Данные в прямом коде	Считывание

Примечание: Ф — безразличное значение сигнала; R_{off} — выходное сопротивление, равное ∞ .

Таблица 8.3. Назначение выводов микросхемы ППЗУ КР556РТ13

Обозначение	Назначение
$A_0...A_9$	Адресные входы
$DO_0...DO_3$	Выход данных
$\overline{CS}_1, \overline{CS}_2$	Выбор микросхемы

Таблица 8.4. Таблица истинности микросхемы ППЗУ КР556РТ13

\overline{CS}_1	\overline{CS}_2	$A_0...A_9$	$DO_0...DO_3$	Режим работы
М	М	Ф	R_{off}	Хранение
0	0	A	Данные в прямом коде	Считывание

Примечание: ММ — любая комбинация, кроме 00; R_{off} — выходное сопротивление, равное ∞ .

Таблица 8.5. Таблица истинности микросхемы шинного формирователя К589АП16

ВК	УВ	Направление передачи	Значение А, В и С	Выходы в выключенном состоянии
0	0	От входа A_i на выход B_i	$B_i = A_i$	C_i
0	1	От входа B_i на выход C_i	$C_i = B_i$	—
1	0	Передача отсутствует	$B_i = R_{off}, C_i = R_{off}$	B_i, C_i
1	1	Передача отсутствует	$B_i = R_{off}, C_i = R_{off}$	B_i, C_i

Примечание: R_{off} — выходное сопротивление, равное ∞ .

8.2. В запоминающем устройстве используются микросхемы ОЗУ КМ132РУ8А (см. рис. 8.12). Как определить информационную емкость каждой микросхемы?

Решение

Количество 4-разрядных слоев, хранимых в каждой микросхеме ОЗУ, составляет

$$N = 2^n = 2^{10} = 1024,$$

где n — число адресных выходов микросхемы. Информационная емкость микросхемы составляет

$$1024 \cdot 4 = 4096 \text{ бит.}$$

8.3. В ЗУ КМ132РУ8А необходимо хранить только 512 8-разрядных слов (см. рис. 8.12).

1. Как должны быть соединены выводы микросхемы ОЗУ?
2. Какое число адресных входов необходимо использовать?

Решение

1. Все адресные входы и входы управления микросхемы 1 должны быть соединены с одноименными входами микросхемы 2. Выводы $DIO_0 \dots DIO_3$ микросхемы 1 ОЗУ должны быть соединены с соответствующими двунаправленными шинами с номерами 0...3. Микросхема 2 ОЗУ присоединяется к шинам 4...7.
2. Девять входов ($N = 2^9 = 512$).

8.4. Как должны быть соединены микросхемы ОЗУ КМ132РУ8А (см. рис. 8.12), каждая из которых может хранить 1024 слова, чтобы можно было хранить 2048 4-разрядных слов? Сколько адресных шин должно иметь ОЗУ?

Решение

Все выводы $DIO_0 \dots DIO_3$ микросхемы 1 должны быть соединены с одноименными выводами микросхемы 2 и присоединены к двунаправленной шине. Адресные входы одной микросхемы $A_0 \dots A_9$ должны быть соединены с одноименными адресными входами другой микросхемы.

Всего адресных шин должно быть 11 ($2^{11} = 2048$). Сигнал старшего разряда адреса A_{10} используется совместно с сигналом управления \overline{CS} для формирования сигналов управления, подаваемых на микросхемы 1 и 2 (\overline{CS}_1 и \overline{CS}_2).

При записи информации в микросхему 1 и чтении информации из нее ($\overline{CS}_1 = A_{10} = 0$) на вход \overline{CS} микросхемы 2 (\overline{CS}_2) должен подаваться сигнал 1, чтобы она была в режиме хранения (табл. 8.2), а при сигнале $A_{10} = 1$ наоборот. При сигнале $\overline{CS} = 1$, поступающем из устройства управления, оба ОЗУ должны работать в режиме хранения, для чего $\overline{CS}_1 = \overline{CS}_2 = 1$ (см. рис. 8.12).

8.5. Как построить ОЗУ на 1024 12-разрядных слов, используя микросхемы КМ132РУ8А (см. рис. 8.12)?

Решение

Необходимо использовать три микросхемы. Все одноименные адресные входы и выходы соединить параллельно. Выводы $DIO_0...DIO_3$ присоединить к двунаправленным шинам с номерами 0...11, так, чтобы каждая микросхема была присоединена к четырем шинам, т. е. чтобы каждая микросхема использовалась для хранения четырех из двенадцати разрядов слова.

8.6. Как необходимо соединить микросхемы ОЗУ КМ132РУ8А (см. рис. 8.12), чтобы получить два ОЗУ, каждое емкостью 1024 4-разрядных слова, между которыми можно было бы производить обмен информацией?

Решение

Выводы $DIO_0...DIO_3$ одной микросхемы соединить с одноименными выводами другой и присоединить их к 4-разрядной двунаправленной шине. Для обмена информацией необходимо подавать сигналы управления так, чтобы при считывании информации с одной микросхемы производилась запись в другую. Адресные шины должны быть разными для возможности записи по адресу, отличному от адреса, по которому осуществляется считывание.

8.7. Как должны быть соединены микросхемы (см. рис. 8.12), чтобы можно было считывать информацию как с ППЗУ 1, так и с ОЗУ 1 на общую двунаправленную шину, а также осуществлять запись информации в ОЗУ?

Решение

Выводы $DIO_0...DIO_3$ микросхемы ОЗУ должны быть соединены с соответствующими выводами ППЗУ и присоединены к общей двунаправленной шине. Адресные входы микросхемы ОЗУ должны быть соединены с соответствующими адресными входами микросхемы ППЗУ. При считывании информации из одного ЗУ входное сопротивление другого должно быть равно бесконечности. Для этого должны подаваться сигналы управления, приведенные в табл. 8.2 и 8.4.

Контрольные вопросы

1. Что такое ЗУ и какие функции оно выполняет?
2. Дайте определение ОЗУ. Какие функции оно выполняет?
3. Дайте определение ПЗУ. Какие функции оно выполняет?
4. Дайте определение РПЗУ. Какие функции оно выполняет?
5. Как вы понимаете матричную организацию ЗУ?
6. Как формируются ячейки памяти на МОП-транзисторах?
7. Как формируются ячейки памяти на биполярных транзисторах?
8. Что такое ЗУ динамического типа?
9. Что такое ЗУ статического типа?
10. Как организовано энергонезависимое ЗУ?
11. Как организовано ЭСППЗУ?
12. Какие типы интегральных схем памяти вы знаете?

Рекомендуемая литература

1. Аваев Н. А., Наумов Ю. Е., Фролкин В. Т. Основы микроэлектроники: Учебное пособие для вузов. — М.: Радио и связь, 1991.
2. Валиев К. А., Орликовский А. А. Полупроводниковые интегральные схемы памяти на биполярных транзисторах. — М.: Советское радио, 1979.
3. Гольденберг Л. М., Малев В. А., Малько Г. Б. Цифровые устройства и микропроцессорные системы. Задачи и упражнения: Учебное пособие для вузов. — М.: Радио и связь, 1991.
4. Степаненко И. П. Основы микроэлектроники: Учебное пособие для вузов. — 2-е изд. — М.: Лаборатория базовых знаний, 2000.

9. Триггеры

9.1. Бистабильные ячейки

Триггер (от английского trigger) — спусковое устройство с набором устойчивых состояний, в которых оно может находиться сколь угодно долго.

Если таких устойчивых состояний два, то такой триггер называется *бистабильным*.

Воздействие внешнего сигнала переводит триггер в противоположное состояние. Каждому состоянию триггера соответствуют определенные сигналы на выходах, отличающиеся высоким или низким уровнями.

Если у триггера более двух устойчивых состояний, то говорят о *мультистабильных* триггерах.

В основе электронного триггера лежит схема бистабильной ячейки (рис. 9.1).

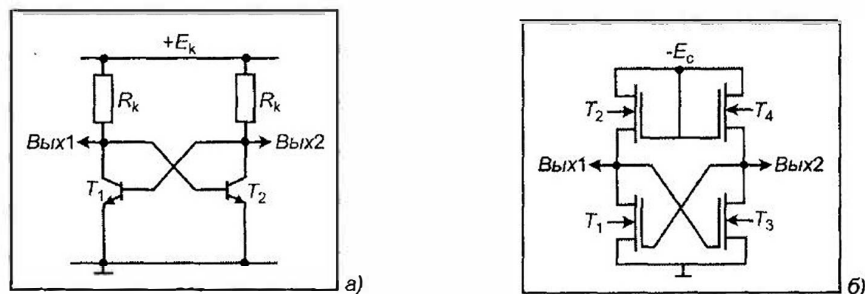


Рис. 9.1. Схема бистабильной ячейки, реализованной на биполярных (а) и МДП-транзисторах (б)

Эти схемы имеют симметричную конфигурацию и перекрестные обратные связи. Устойчивость состояний бистабильной ячейки обеспечивается тем, что один из ключей постоянно заперт, а второй открыт и насыщен. Переход бистабильной ячейки из одного устойчивого состояния в другое происходит в соответствии с соотношением:

$$t_{пер} = t_p + t_{зф} + t_n + t_{нф}$$

где t_p — время рассасывания зарядов в транзисторе, $t_{зф}$ — время задержки фронта, $t_{нф}$ — длительность фронта, t_n — время накопления зарядов.

Анализ дает следующие значения: $t_{зф} \sim 1$ нс, $t_{нф} \sim 1$ нс, $t_n \sim 90$ нс, $t_p \sim 20$ нс.

Другими словами, инерционность триггера определяется в основном временем накопления и рассасывания зарядов в транзисторных структурах.

9.2. Триггер Шмитта

Триггер может быть реализован не только на транзисторных ключах, но и на переключающих токах. На рис. 9.2 представлена схема *триггера Шмитта*, реализованного на переключающих токах на основе транзисторов T_1 и T_2 . Обратная связь реализуется с помощью делителя напряжения на резисторах R_1 и R_2 .

Если в исходном состоянии транзистор T_1 заперт, а транзистор T_2 открыт, то $U_{k1} = E_k$, $E = \gamma E_k$, $U_{k2} = E_k - I_0 R_k$, $U_{k2} \geq E$. Исходное состояние сохраняется пока транзистор T_1 заперт и $U_{вх} < E$. Если на базу транзистора T_1 подать напряжение отпирающего транзистора $U_{вх}^+ = E - \delta = \gamma E_k - \delta$, то при $\delta = 0,1$ В, $E_k = 5$ В и $\gamma = 1/2$, $U_{вх}^+ = 2,4$ В. Если $U_{вх} > U_{вх}^+$, то наблюдается приращение тока в коллекторной цепи и начинает развиваться регенеративный процесс, в результате которого ток I_0 открывает транзистор T_1 . Благодаря обратной связи через делитель напряжения на основе резисторов R_1 и R_2 напряжение на базе транзистора T_2 будет мало и он закроется.

Напряжение $U_{вх}^+$, при котором происходит скачкообразное переключение тока I_0 из транзистора T_2 в транзистор T_1 , называют *напряжением срабатывания триггера*. После срабатывания распределение потенциалов примет вид:

$$U_{k1} = U_{k0}, \quad E = \gamma U_{k0}, \quad U_{k2} = E_k,$$

где U_{k0} — коллекторный потенциал открывшегося транзистора T_1 . Чтобы вернуть триггер в исходное состояние, необходимо уменьшить входной сигнал до значения $U_{вх} \sim E$. В этом случае начинает отпираться транзистор T_2 . Напряжение $U_{вх} \approx 1,3$ В называется *напряжением отпускания триггера*.

Для триггера Шмитта напряжение отпускания меньше напряжения срабатывания. В общем случае речь идет о порогах срабатывания триггера. На рис. 9.2, б показана передаточная характеристика триггера Шмитта, являющегося пороговым устройством. Характерной особенностью триггера Шмитта является гистерезисный характер его передаточной характеристики.

Если амплитуда входных сигналов больше порога срабатывания, то триггер Шмитта играет роль формирователя импульсов прямоугольной формы с короткими фронтами.

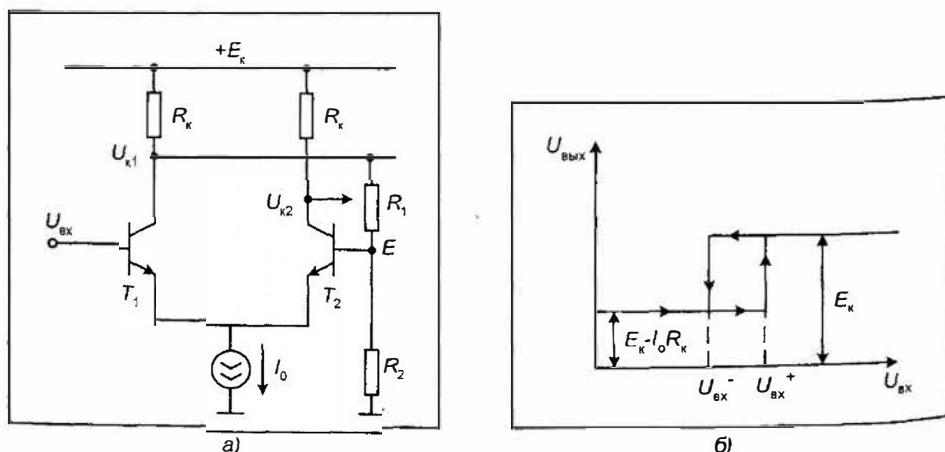


Рис. 9.2. Схема триггера Шмитта (а) и его передаточная характеристика (б)

9.3. RS-триггер

RS-триггер имеет два входа S (set — установка) и R (reset — сброс) и относится к асинхронному типу.

Из анализа схем триггеров следует, что любой триггер является совокупностью нескольких определенным образом соединенных интегральных логических элементов (ИЛЭ). Для выполнения заданных функций триггер собирается из разного количества ИЛЭ, соединенных по законам булевой алгебры.

Для удобства изложения будем пользоваться установленными условными обозначениями ИЛЭ. Рассматривать принципиальные электрические схемы триггеров различного назначения достаточно сложно.

По способу записи информации различают *асинхронные* и *синхронные* (тактируемые) триггеры.

Состояние асинхронных триггеров меняется при поступлении сигналов на входы.

Состояние синхронных триггеров изменяется, прежде всего, вследствие появления тактирующих импульсов на специальных входах триггера.

На рис. 9.3, а приведена структурная схема RS-триггера.

RS-триггер формируется на основе ИЛЭ типа И — НЕ, охваченных обратной связью, и инверторах.

При поступлении импульса на вход S ($S = 1$, $R = 0$) на выходах ИЛЭ получим $\bar{S} = 0$, $\bar{R} = 1$. В целом триггер будет соответствовать состоянию $Q = 1$ и $\bar{Q} = 1$, которое соответствует окончанию импульса на входе S . Выход Q называют главным, и он определяет состояние триггера в целом.

Если на вход R подать импульс ($R = 1$, $S = 0$), то триггер установится в состояние $Q = 0$ и $\bar{Q} = 1$.

Функционирование RS-триггера описывается характеристическим уравнением вида:

$$Q^{n+1} = S + \bar{R}Q^n,$$

где Q^n и Q^{n+1} — собственно сигнал на входе до и после поступления входного импульса. Если одновременно $R = S = 1$, то во время действия этих импульсов

$$\bar{S} = \bar{R} = 0 \text{ и } Q = \bar{Q} = 1.$$

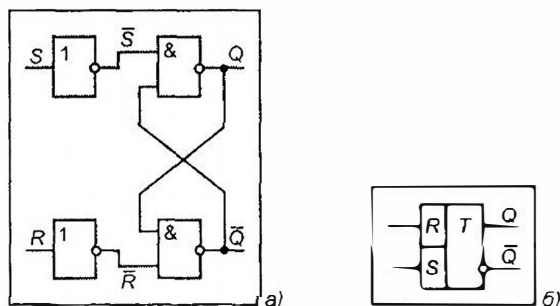


Рис. 9.3. Структура RS-триггера (а) и его условное обозначение (б)

По окончании действия импульсов RS-триггер случайным образом установится либо в состояние $Q = 0$ или $Q = 1$. Такая комбинация $R = S = 1$ приводит к неопределенному состоянию RS-триггера и поэтому одновременная подача импульсов на оба входа не допускается.

Максимальная частота переключения триггера определяется минимально доступным временным интервалом между двумя сигналами, которые поступают на входы триггера.

9.4. RST-триггер

Триггеры, у которых состояние меняется только путем поступления тактовых импульсов, называют синхронными триггерами.

К ним относятся *RST-триггер*, имеющий времязадающий вход C (clock). В промежутках между тактовыми импульсами изменения сигналов на входах не вызывают переключения триггера. На рис. 9.4 приведена схема RST-триггера, состоящего из двухвходовых ИЛЭ типа И — НЕ.

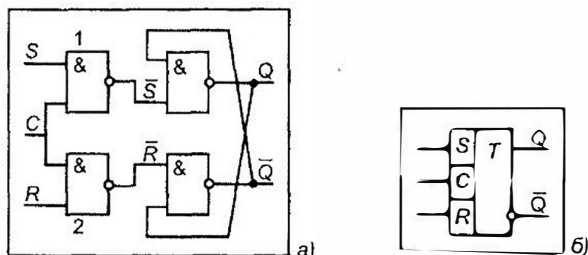


Рис. 9.4. Структура RST-триггера (а) и его условное обозначение (б)

При отсутствии тактового импульса $C = 0$ и на входах \bar{S} и \bar{R} не меняется. Поэтому состояния на выходах Q и \bar{Q} остаются постоянными.

Если на вход C подать импульс ($C = 1$), то элементы 1 и 2 работают как инверторы. Поведение RST-триггера будет идентично поведению RS-триггера. Характеристическое уравнение RST-триггера имеет вид:

$$Q^{n+1} = C(S + \bar{R}Q^n).$$

Для RST-триггера недопустима комбинация сигналов $S = R = 1$.

9.5. D-триггер

D-триггер (от delay — задержка) имеет тактовый вход C и информационный вход D .

На рис. 9.5 представлена структура D-триггера. Тактовый вход соединен одним из входов ИЛЭ 1, а элементы 1 и 2 охвачены обратной связью.

Если тактовый импульс отсутствует ($C = 0$), то сигналы \bar{R} и \bar{S} не зависят от входного сигнала D и триггер сохраняет то состояние, в котором он переключится при поступлении предыдущего тактового импульса. Пусть к моменту поступления тактового импульса

($C = 1$) на информационном входе $D = 1$, тогда $S = 0$, а $\bar{R} = 1$. Состояние триггера будет характеризоваться $\bar{Q} = 1$. Если при поступлении тактового импульса ($C = 1$) на входе $D = 0$, то $\bar{S} = 1$, а $\bar{R} = 0$. Триггер установится в состояние $Q = 0$. Другими словами, при действии тактового импульса в D-триггер записывается информация, которая ранее была на входе D .

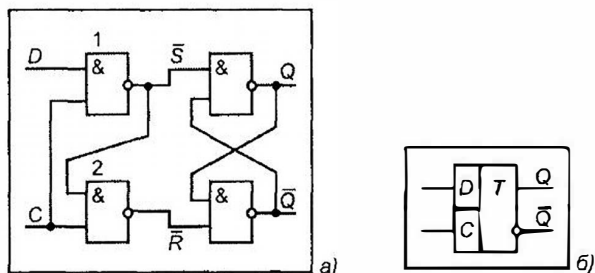


Рис. 9.5. Структура D-триггера (а) и его условное обозначение (б)

Характеристическое уравнение D-триггера имеет вид:

$$Q^{n+1} = D.$$

Устойчивая работа D-триггера обеспечивается постоянством сигнала D на время действия тактового импульса.

9.6. Т-триггер

Счетный *T-триггер* предназначен для изменения своего состояния на противоположное каждый раз, когда на единственный вход поступит управляющий импульс ($C = 1$). Другими словами, он способен считать импульсы на входе.

В основе конструкции Т-триггера лежат два RST-триггера, соединенные с инвертором (рис. 9.6). Триггер T_1 называют ведущим, а T_2 — ведомым.

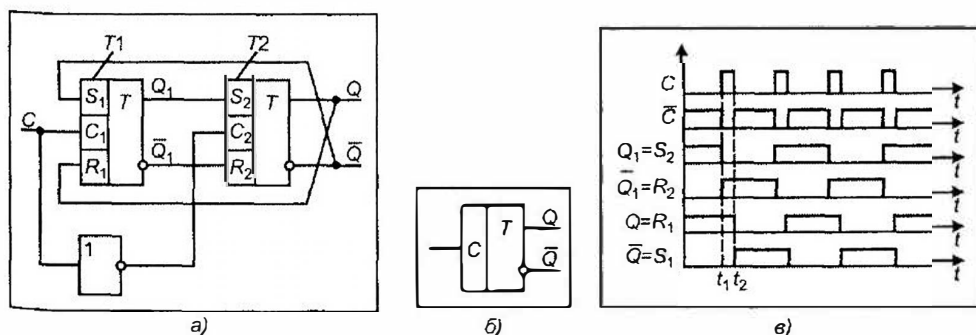


Рис. 9.6. Структура Т-триггера (а), его условное обозначение (б) и временные диаграммы работы (в); t_1 — время поступления тактового импульса, t_2 — время восстановления сигнала

В интервале между импульсами $C = 1$ на входах триггера T_2 и T_1 сигналы одинаковы и тогда $Q = Q_1$, $\bar{Q} = \bar{Q}_1$. В момент поступления импульса $C = 1$ информация с выходов триггера T_2 записывается в T_1 , и тогда $Q_1 = \bar{Q}$ и $\bar{Q}_1 = Q$. По окончании действия импульса вновь имеем $C = 1$ и информация из T_1 переписывается в T_2 так, что состояние изменяется на противоположное. В результате действия каждого управляющего импульса Т-триггер переключается в противоположное состояние с задержкой. Задержка равна длительности импульса.

Характеристическое уравнение Т-триггера имеет вид:

$$Q^{n+1} = CQ^n + \bar{C}\bar{Q}^n,$$

где CQ^n — значение во время тактового импульса, а $\bar{C}\bar{Q}^n$ — по его окончании.

9.7. JK-триггер

JK-триггеры (jump — переброс, keep — держать) являются универсальными и относятся к категории синхронных. Другими словами, выходные уровни устанавливаются только при поступлении тактовых импульсов.

В основе JK-триггера лежит структура Т-триггера, входы которого связаны с выходами перекрестными обратными связями. JK-триггер имеет два информационных входа: J и K , а также тактовый вход C (рис. 9.7).

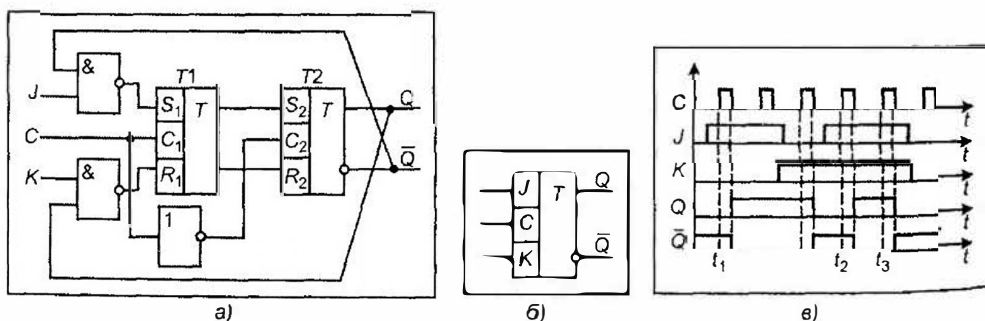


Рис. 9.7. Структура JK-триггера (а), его условное обозначение (б) и временные диаграммы (в); $t_1 - t_2$ — интервал работы в качестве RST-триггера; $t_2 - t_3$ — интервал работы Т-триггера

При наборе $J = 1$, $K = 1$ схемы И повторяют сигналы на входах T_2 , что характерно для Т-триггера (интервал $t_2 - t_3$). При остальных наборах величины J и K схема работает как RST-триггер (интервал $t_1 - t_2$), где уровень J устанавливает значение $Q = 1$, а уровень K — значение $Q = 0$.

Характеристическое уравнение JK-триггера имеет вид:

$$Q^{n+1} = J\bar{Q}^n + \bar{K}Q^n.$$

В частности, при $J = 1$ и $K = 1$ получим $Q^{n+1} = \bar{Q}^n$, что соответствует режиму работы Т-триггера.

Задачи и упражнения

9.1. На рис. 9.3 приведена схема RS-триггера, реализованного на элементах типа И—НЕ.

1. Объясните принцип работы триггера.
2. Составьте таблицу переключений триггера.
3. На вход триггера поданы сигналы $S = R = 0$ — запрещенная комбинация. Затем подаются сигналы $S = R = 1$. В каком состоянии окажется триггер?
4. На вход триггера поданы сигналы $S = R = 1$ — запрещенная комбинация. Затем подаются сигналы $S = R = 0$. В каком состоянии окажется триггер?

Решение

Работа триггера описывается в табл. 9.1.

Таблица 9.1

S^n	R^n	Q^n	Q^{n+1}	\bar{Q}^{n+1}
0	0	0	1	1 запрет
0	0	1	1	1 запрет
0	1	0	1	0
0	1	1	1	0
1	0	0	0	1
1	0	1	0	1
1	1	0	0	1
1	1	1	1	0

После подачи сигналов $S = R = 1$ триггер будет в состоянии $Q = 0$, $\bar{Q} = 1$, либо в состоянии $Q = 1$, $\bar{Q} = 0$.

После подачи сигналов $S = R = 0$ триггер будет в состоянии $Q = 0$, $\bar{Q} = 1$, либо в состоянии $Q = 1$, $\bar{Q} = 0$.

9.2. Собрать схему (рис. 9.8) и исследовать характеристики триггера Шмитта в программе Electronics Workbench.

Контрольные вопросы

1. Что такое бистабильная ячейка?
2. Что такое RS-триггер?
3. Чем отличается RST-триггер от RS-триггера?

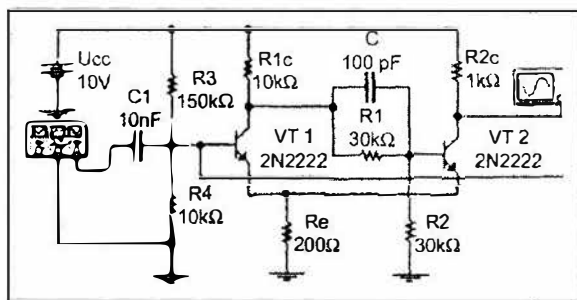


Рис. 9.8. Триггер Шмитта

4. Что такое JK-триггер?
5. Какие функции выполняет Т-триггер?
6. Какие функции выполняет D-триггер?

Рекомендуемая литература

1. Алексенко А. Г. Основы микросхемотехники. — М.: ЮНИМЕДИАСТАЙЛ, 2002.
2. Гольденберг Л. М., Малев В. А., Малько Г. Б. Цифровые устройства и микропроцессорные системы. Задачи и упражнения: Учебное пособие для вузов. — М.: Радио и связь, 1991.
3. Карлшук В. И. Электронная лаборатория на IBM PC. Программа Electronics Workbench и ее применение. — 3-е изд. — М.: СОЛОН-пресс, 2003.

10. Цифровые устройства на основе триггеров

Класс интегральных схем, предназначенных для обработки цифровой информации в вычислительных устройствах, отнесем к *цифровым устройствам*.

Цифровые устройства могут сопрягаться с архитектурой вычислительного устройства, а также технологически быть совместимыми с другими БИС микропроцессорного комплекта. Эти устройства относятся к интегральным системам последовательностного типа и реализуются на основе триггеров.

10.1. Счетчики

Каждая вычислительная система содержит несколько счетчиков, предназначенных для подсчета числа событий, временных интервалов, а также упорядочивания событий в хронологической последовательности. Счетчики состоят из триггеров.

Функциональная схема счетчика, имеющая 16 различных выходных состояний, представлена на рис. 10.1. Схема состоит из четырех JK-триггеров, каждый из которых работает в режиме переключения ($J = K = 1$).

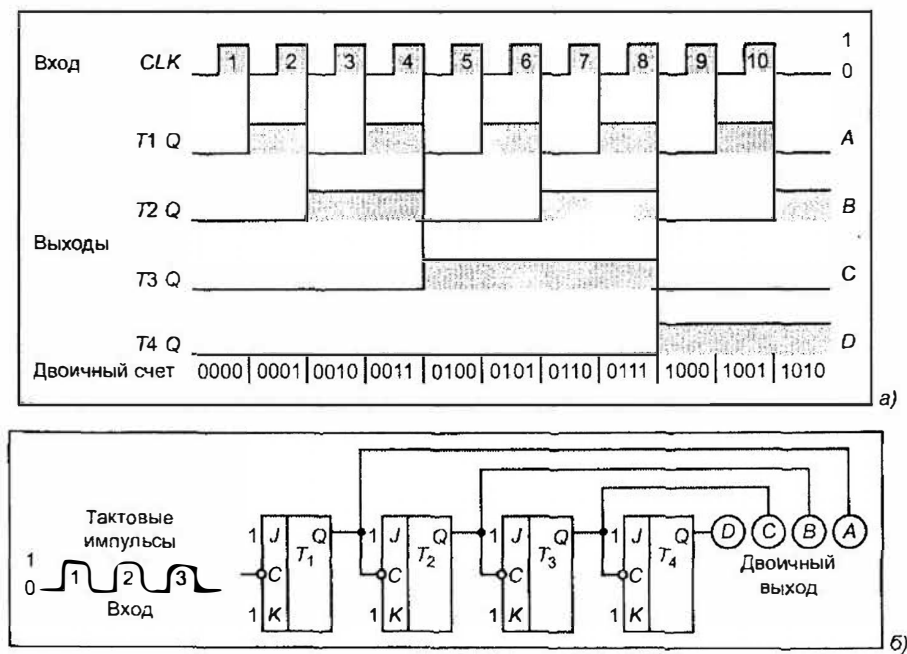


Рис. 10.1. Функциональная схема счетчика (а) и временные диаграммы (б)

Пусть в начальный момент состояния выходов счетчика соответствует двоичному числу 0000. При поступлении тактового импульса 1 на синхронизирующий вход C триггера T_1 этот триггер переключается и на индикаторе появится число 0001.

Тактовый импульс 2 возвращает триггер T_1 в исходное состояние 0 ($Q = 0$), что стимулирует переключение триггера T_2 в состояние 1 ($Q = 1$). На индикаторе появится число 0010.

Срез сигнала на выходе предыдущего триггера запускает следующий триггер. Триггер T_1 переключается с приходом каждого такого импульса, триггер T_2 переключается в два раза реже триггера T_1 и т. д. Под диаграммами, показанными на рис. 10.1, указаны двоичные числа, соответствующие различным состояниям счетчика.

Из диаграммы также следует, что тактовые импульсы запускают только триггер T_1 , триггер T_1 запускает триггер T_2 , триггер T_2 , соответственно, T_3 и т. д. Изменение состояния последовательно распространяется по цепочке триггеров, и такие счетчики называются *счетчиками со сквозным переносом* или *асинхронными счетчиками*.

Это четырехразрядный счетчик, на что указывает число двоичных разрядов на выходе. Для быстродействующих цифровых устройств важно, чтобы все ступени счетчика срабатывали одновременно. Такие счетчики называют *синхронными*. С этой целью тактовые импульсы одновременно подаются на выходы C всех триггеров. В данном режиме JK-триггеры используются как при $J = K = 1$, так и при $J = K = 0$.

10.2. Регистры сдвига

Регистры сдвига или *сдвиговые регистры* предназначены для сдвига числа на одну позицию каждый раз, когда вводится новое число.

Регистры сдвига обладают временной памятью. Регистры сдвига реализуются на триггерах. На рис. 10.2 приведена схема режима сдвига, реализованного на JK-триггерах.

За четыре полных такта логическая единица с входа первого разряда передается на выход Q последнего разряда. На рис. 10.2, б приведены временные диаграммы работы четырехразрядного регистра сдвига.

Предполагается, что первоначальное состояние регистра нулевое, а логическое состояние на входе является логической единицей и не изменяется в течение четырех тактов синхронизации.

Если данные вводятся поразрядно, то регистр называется *последовательным регистром сдвига*.

Существует и параллельная загрузка, при которой информационные разряды вводятся в регистр одновременно по команде одного тактового импульса.

10.3. Сумматоры

Сумматор представляет собой логический операционный узел, выполняющий арифметическое сложение кодов двух чисел в цифровых устройствах.

При арифметическом сложении выполняются различные дополнительные операции: учет знаков чисел, выравнивание порядков слагаемых и т. д.

Сумматоры квалифицируют по различным признакам, например по числу входов и выходов.

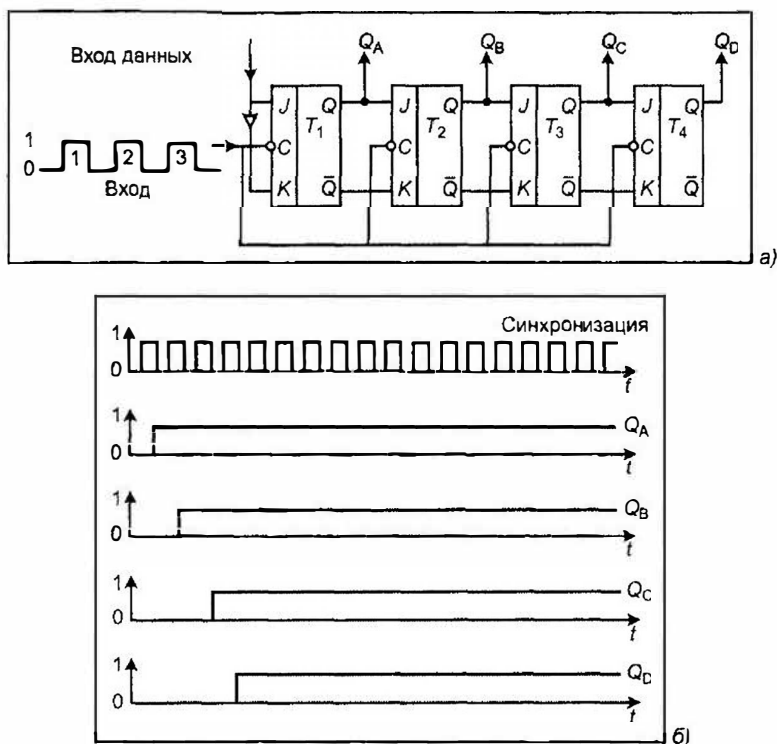


Рис. 10.2. Функциональная схема регистра сдвига и временные диаграммы его работы

На рис. 10.3 показан четвертьсумматор с двумя входами для двух одноразрядных чисел и одним выходом, на котором реализуется их сумма.

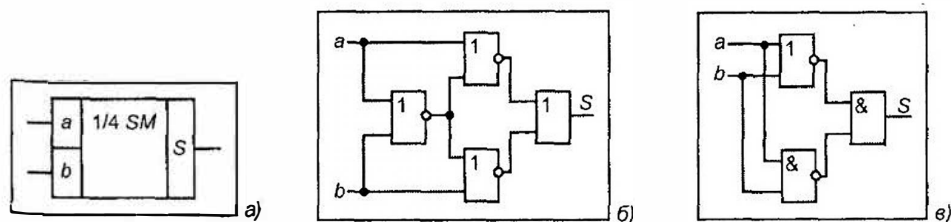


Рис. 10.3. Четвертьсумматор (а) и схема его реализации на элементах И — НЕ и ИЛИ — НЕ (б, в)

На рис. 10.4, а показан полусумматор с двумя входами, на которые подаются одноименные разряды двух чисел, и двумя выходами, на одном из которых реализуется арифметическая сумма, а на другом — перенос в следующий более старший разряд. Обозначением полусумматора служат буквы HS (halfsum — полусумма).

Полусумматор может быть реализован на элементе исключающих ИЛИ и одном двухвходовом вентиле И (рис. 10.4, б).

Полные одноразрядные сумматоры характеризуются наличием трех входов, на которые подаются одномерные разряды двух складываемых чисел и перенос из предыдущего бо-

лее младшего разряда, и двумя выходами, на одном из которых реализуется арифметическая сумма в данном разряде, а на другом — перенос в более старший разряд.

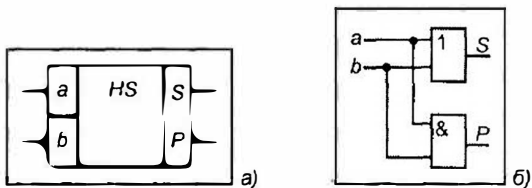


Рис. 10.4. Полусумматор (а) и его структурная схема (б)

Схема полного сумматора, реализованного на двух полусумматорах, показана на рис. 10.5, а.

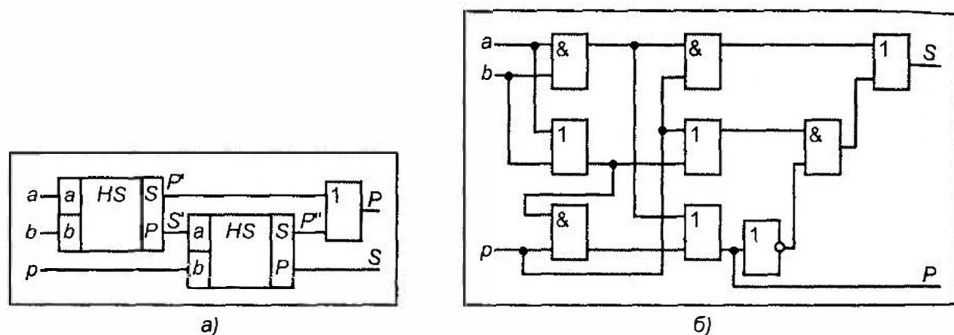


Рис. 10.5. Сумматор, реализованный на двух полусумматорах (а), и один из вариантов его реализации на ИЛЭ типа ИЛИ — НЕ (б)

Сумматор на схемах обозначается SM. На входы a и b сумматора подаются два слагаемых, на вход p — перенос из предыдущего несущего разряда. Выходы обеспечивают суммирование по каналу S и перенос в старший разряд по каналу P .

Особенностью выходных сигналов полного двоичного сумматора является их самодвойственность как функций алгебры логики. Другими словами, выходные сигналы S и P способны инвертировать свое значение при инвертировании всех переменных, от которых они зависят.

Полные сумматоры совместно с инверторами используются для проведения операций вычитания. Сумматоры и вычитатели реализованы в виде интегральных схем.

10.4. Шифраторы и дешифраторы

Шифратором в вычислительной технике называют комбинационный узел, предназначенный для преобразования одного из совокупности входных функционально однотипных однопозиционных кодов в двоичный код.

Одним из аппаратных средств для ввода данных в цифровую систему является клавиатура, которая переводит десятичные цифры в код. На рис. 10.6 приведена структура шифратора клавиатуры, выполненная на логических элементах.

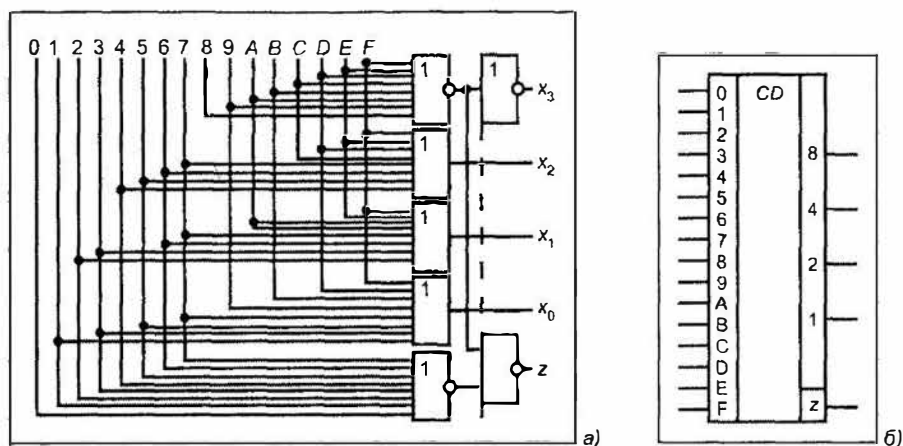


Рис. 10.6. Структура шифратора клавиатуры для ввода шестнадцатеричных чисел (а) и ее условное обозначение (б)

Клавиатура предназначена для ввода шестнадцатеричных чисел. Переменные 4-разрядного двоичного кода с весами соответственно 8, 4, 2, 1 обозначены переменными x_3 , x_2 , x_1 , x_0 . Осведомительный сигнал Z принимает значение 1 при нажатии клавиши и значение 0 при всех нажатых клавишах. Этот сигнал называется *сигнал-универсум*.

Шестнадцатеричные цифры состоят из десятичного числа от 0 до 9 и шести цифр, обозначенных большими латинскими буквами A, B, C, D, E, F. Схема шифратора состоит из шести логических ячеек типа ИЛИ — НЕ и одной типа И — НЕ. Интегральная схема может быть обозначена буквами CD (coder). Приведенный тип шифратора предназначен для обработки зависимых однопозиционных сигналов. Это означает, что из всех клавиш может быть нажата только одна. Такие шифраторы называются приоритетными. При большом числе кнопок на клавиатуре код каждой клавиши поддерживается или программным сканированием столбцов и строк, или специализированным контроллером.

Дешифраторам в цифровой технике называют комбинационный узел, преобразующий n -разрядный двоичный код в однопозиционный или в совокупность однопозиционных кодов. Примерами однопозиционных кодов могут быть адрес ячейки памяти, однозначный сигнал установки счетчика в нулевое состояние и т. д.

Двоичный n -разрядный код имеет 2^n наборов и столько же однопозиционных кодов. Дешифраторы, как и шифраторы, выполняются в виде интегральных схем.

Полным дешифратором называется дешифратор, имеющий $N = 2^n$ выходов, если $N < 2^n$, то — *неполным*.

На рис. 10.7 приведены структурные схемы полного дешифратора, преобразующего двухразрядный двоичный код в однопозиционный.

Переменные дешифруемого двухразрядного двоичного кода обозначены x_0 и x_1 , однопозиционный выход дешифратора обозначен как J_0 . Однопозиционный выход дешифратора вырабатывает активный сигнал 0 или 1. Выходные сигналы J_1 , J_2 , J_3 определяются функцией алгебры логики и зависят от переменных x_0 и x_1 . Условное обозначение дешифратора приведено на рис. 10.7, в. Цифры 2 и 1 слева обозначают двоичные веса разрядов дешифруемого двоичного кода, а кружки справа свидетельствуют о том, что активными сигналами выхода являются уровни 0 положительной логики. Их отсутствие свидетельст-

ует о том, что активными сигналами выхода являются 1. На интегральных схемах дешифратора маркируются буквы DC — decoder.

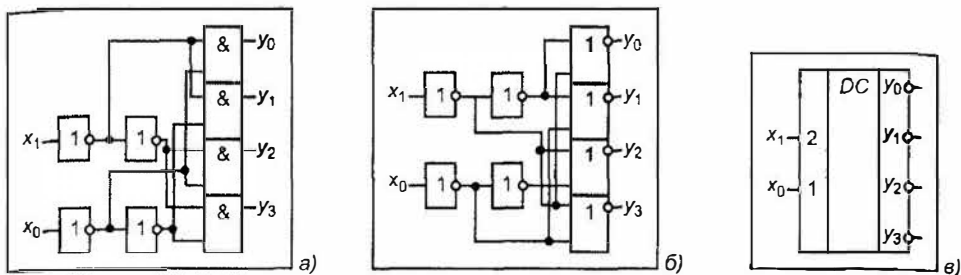


Рис. 10.7. Структурные схемы дешифраторов (а, б) и их условное обозначение (в)

10.5. Мультиплексор

Мультиплексором в вычислительной технике называется цифровой комбинационный узел, осуществляющий адресную передачу (коммутацию) данных от одного из многих входов в один единственный выход.

Передача n -разрядного слова может быть осуществлена с помощью n -однобитовых мультиплексоров. Мультиплексор может быть реализован как логический узел, составленный из интегральных логических элементов, в которых адрес задается двоичным кодом. На рис. 10.8, а приведена структурная схема мультиплексора.

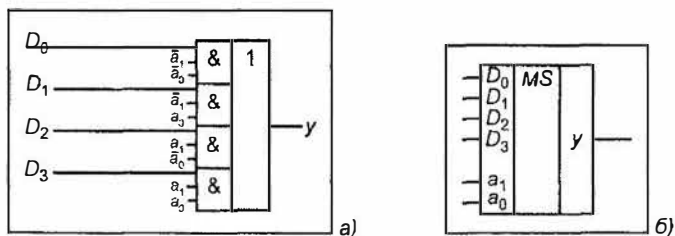


Рис. 10.8. Структурная схема мультиплексора на четыре позиции (а) и ее условное обозначение (б)

Мультиплексор на четыре входных однобитовых данных D_0, D_1, D_2, D_3 реализован на четырех 3-входовых элементах И и одном 4-входовом элементе ИЛИ.

Адресные переменные a_1 и a_0 характерны для разрядности адреса входного направления равного двум. Мультиплексор на четыре позиции обозначают MS4 → 1 (рис. 10.8, б). Число информационных входов мультиплексора может быть увеличено как за счет комбинационного включения мультиплексоров MS4 → 1, так и за счет увеличения числа состояний выхода до трех. На основе мультиплексоров могут быть созданы различные функциональные цифровые устройства. Например, на основе мультиплексора можно создать многоканальный селектор, позволяющий осуществить коммутацию n входов с m выходами. При этом возможна передача данных с любого входа на любой выход, при условии использования между входами и выходами минимального числа линий связи.

Мультиплексоры используются в составе микроконтроллера для тестирования состояния сложного объекта, в качестве сдвигового регистра, в режиме кольцевого счетчика и т. д.

Задачи и упражнения

10.1. На рис. 10.9 приведена схема 4-разрядного сдвигающего регистра на D-триггерах К155ТМ2 и его условное обозначение:

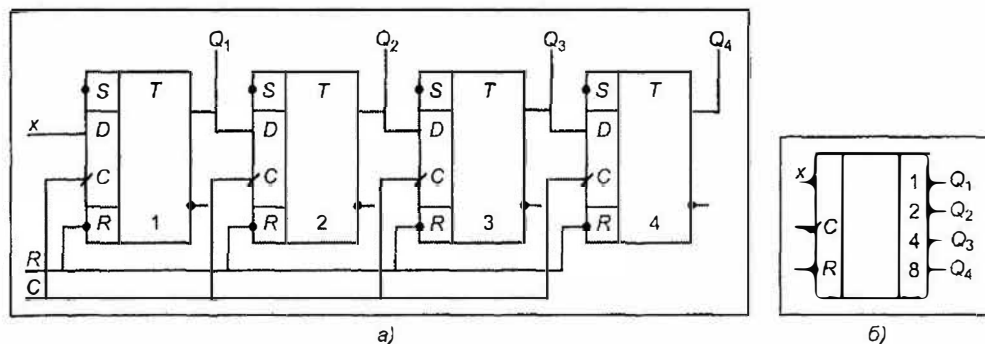


Рис. 10.9. Четырехразрядный сдвигающий регистр (а) и его условное обозначение (б)

1. Пояснить принцип работы регистра на примере продвижения одной единицы.
2. Построить временные диаграммы сигналов на выходах регистра Q при подаче шести продвигающих импульсов C и $X = 1$. Исходное состояние регистра $Q_1 = Q_2 = Q_3 = Q_4 = 0$.
3. Определить максимальную задержку выходных сигналов относительно продвигающего импульса C ?
4. Можно ли осуществить запись новой информации при считывании старой в последовательном коде?
5. Можно ли в регистре заменить триггеры К155ТМ2 триггерами К155ТМ7 или К155ТМ5?
6. Можно ли в регистре использовать JK-триггеры?

Решение

1. Временные диаграммы приведены на рис. 10.10.
2. Задержка выходных сигналов для всех выходов одинакова и равна задержке, создаваемой одним триггером.
3. Можно. Новая информация записывается в освобождающиеся триггеры.
4. Нельзя. Триггеры ТМ5 и ТМ7 — одноступенчатые со статическими входами C . При сигналах $x = 1$ и $C = 1$ единица запишется во все триггеры.
5. Можно. Передача информации от триггера к триггеру должна осуществляться по двум проводам (с выходов Q и \bar{Q} поступать на входы J и K).

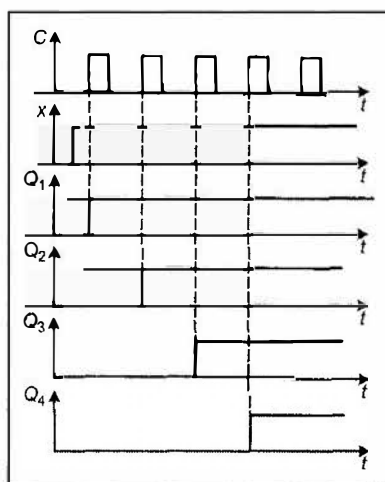


Рис. 10.10. Временные диаграммы на выходе регистра

Контрольные вопросы

1. Что такое сдвиговый регистр?
2. Что такое мультиплексор?
3. Что такое дешифратор?

Рекомендуемая литература

1. Алексенко А. Г. Основы микросхемотехники. — М.: ЮНИМЕДИАСТАЙЛ, 2002.
2. Гольденберг Л. М., Малев В. А., Малько Г. Б. Цифровые устройства и микропроцессорные системы. Задачи и упражнения: Учебное пособие для вузов. — М.: Радио и связь, 1991.
3. Токхейм Р. Основы цифровой электроники. — М.: Мир, 1988.

11. Микропроцессоры и компьютеры

Цифровой электронике принадлежит решающая роль в системах обработки информации различных уровней — от микроконтроллера до супер-ЭВМ.

Микропроцессор является самостоятельной или составной частью вычислительного устройства, осуществляющей обработку информации и управление этим процессом, которая выполнена в виде интегральной схемы.

Микропроцессор наряду с запоминающими устройствами, интерфейсом является основой ряда вычислительных устройств, микро-ЭВМ, персональных компьютеров.

Структурная схема микропроцессора включает в себя арифметико-логическое устройство (АЛУ), устройство управления (УУ), регистры (Р) и устройство ввода-вывода информации (УВВ) или интерфейс (рис. 11.1).

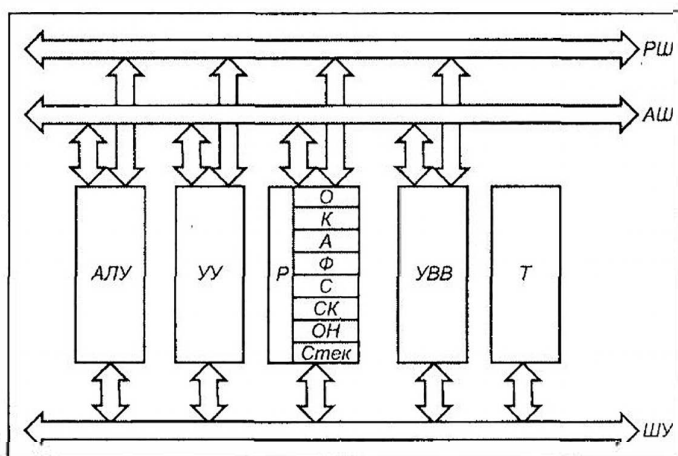


Рис. 11.1. Обобщенная структура (архитектура) микропроцессора

Микропроцессор управляется сигналами трех видов:

- ☐ адресными, передающимися по адресным шинам (АШ);
- ☐ информационными, передающимися по разрядным шинам (РШ);
- ☐ сигналами, управляющими микропроцессором соответственно через шины управления (ШУ).

Шины представляют собой группу линий связи, число которых определяется, прежде всего, разрядностью обрабатываемого информационного сигнала. Шины — двунаправ-

тенные, позволяющие передавать информацию в обоих направлениях. Иногда шины объединяют в одну или две.

Арифметическое логическое устройство (АЛУ) предназначено для выполнения всех операций вычислительного процесса. В АЛУ осуществляется обработка информации по кодам программы, совершаются арифметические и логические операции над числами и адресами, представленными в двоичном коде. В набор команд АЛУ входят арифметические и логические сложения и умножения, сдвиги уровней, сравнения и т. п. Арифметические операции выполняются по правилам двоичной арифметики, логические — по правилам алгебры логики.

Устройство управления (УУ) управляет работой АЛУ и всех других элементов микропроцессора.

Из программы, хранящейся в ПЗУ в УУ, поступают команды в виде двоичных сигналов, стимулирующие выполнение команды. УУ распределяет также очередность выполнения команд во времени.

Команда представляет собой двоичное слово из 8, 16, 24 или 64 разрядов.

Команда состоит из кода операции и операндов, представляющих собой исходный элемент данных, над которыми выполняется операция.

Регистры (Р) микропроцессора представляют собой ячейки памяти, служащие для сверхмгновенного хранения только текущей информации (СОЗУ). Набор регистров выполняет различные функции. Регистр — аккумулятор или *регистр операндов* О в процессе выполнения операции в АЛУ хранит два логических числа. Одно из чисел после завершения процесса в АЛУ заменяется результатом и накапливается в регистре или аккумулируется. Отсюда и название — аккумулятор или накопитель. Содержимое второго регистра операндов заменяется в процессе следующей операции другим операндом.

Регистр команд К предназначен для хранения нескольких разрядов командного слова. Адресная часть командного слова содержится в регистре адреса А.

Регистр адреса А предназначен для хранения командных слов операций.

Регистр Ф или *регистр флажков* служит для регистрации ситуации, когда в процессе выполнения операции разрядность результата оказывается больше разрядности каждого из операндов. Эта ситуация разрешается программистом, который устраняет возникшее переполнение регистра. Регистр флажков представляет собой набор триггер-защелок, состояние которых зависит от результата операции в АЛУ. Выход каждого триггера действует, как флажок: флажок нуля ZF, переполнения OF, флажок отрицательного результата NF, флажок переноса CF.

Регистр состояния С предназначен для фиксации состояния микропроцесса. Дело в том, что в системе команд микропроцессора есть команды переходов к выполнению определенного участка программы по специальным признакам и условиям. Это так называемые команды условных переходов. Роль регистра состояний С заключается в информировании устройства управления о переходе процесса вычисления к команде, адрес которой содержится в регистре счета команд.

Регистр-счетчик команд СК служит для формирования адреса следующего командного байта. При выборе каждого командного байта производится автоматический инкремент программного счетчика.

Регистр общего назначения предназначен для хранения промежуточных результатов и команд, которые могут возникать в процессе выполнения программы. Эти регистры коммутируются как с другими регистрами, так и с устройствами ввода-вывода информации.

Указатель стека предназначен для взаимодействия с внешним запоминающим устройством с произвольной выборкой, в котором хранятся промежуточные данные. Стек представляет собой динамическую структуру, размер которой меняется в процессе обработки информации.

Стек работает по принципу "последний пришел — первый ушел". Данные "проталкиваются" в стек, а затем "вытаскиваются" из него. Стек позволяет получать правильную последовательность выполнения различных арифметических действий. Последовательность выполнения определяется старшинством операций: сложение старше умножения, умножение старше сложения и т. д.

Если стек разместить не в микропроцессоре, а в памяти, то увеличивается такой параметр, как глубина стека. Чем больше глубина стека, тем выше возможность микропроцессора.

В состав микропроцессора в качестве вспомогательного устройства включается *таймер* Т.

Его работа определяет во времени порядок работы всех разрядных, адресных и управляющих сигналов, а также синхронизирует работу устройства управления.

Как вычислительное устройство, микропроцессор характеризуется производительностью, разрядностью массива обрабатываемых данных и выполняемых команд, числом команд, количеством внутренних регистров, возможностью обеспечения режима прерывания и числом уровней прерывания, объемом адресуемой памяти, наличием канала прямого доступа к памяти, способом управления и наличием и видом программного обеспечения.

Микропроцессоры классифицируются и по применимости. *Универсальными* называются микропроцессоры, которые могут использоваться в системах обработки разнотипных информационных массивов. *Специальные* микропроцессоры используются для построения определенного типа вычислительного устройства, предназначенного для обработки специализированного информационного массива.

По способу управления различают микропроцессоры со схемным и микропрограммным управлением. *Микропроцессоры со схемным управлением* работают в режиме постоянного набора команд и соответствующей электрической схемой. Эти микропроцессоры отличаются высоким быстродействием. *Микропроцессоры с микропрограммами управления* работают под управлением определенной последовательности микрокоманд. Это уникальный тип микропроцессоров, позволяющих легко перестраиваться с одной программы на другую.

Микропроцессоры могут выполняться в виде одной большой интегральной схемы и несут название однокристалльные микропроцессоры. Если микропроцессоры выполнены по принципу секционирования, позволяющего расширение разрядности и увеличение ЗУ, то говорят о секционировании микропроцессора. В этом случае формируется микропроцессорный комплект интегральных схем. То есть *микропроцессорный комплект интегральных схем* представляет собой конструктивно и электрически совместимые интегральные схемы. Он позволяет сформировать как отдельные микропроцессоры, так и служат основой создания микро-ЭВМ и других вычислительных устройств с заданными техническими характеристиками. Микропроцессорный комплект может быть функционально расширен за счет других совместимых типов ИС ЗУ, интерфейсными ИС, контроллерами и т. д.

11.2. Мировой рынок микропроцессоров

Первый микропроцессор родился на фирме Intel, получивший название 8008. Это было начало семидесятых годов прошлого столетия. Затем появились Intel-8080, Intel-8085 и

Intel-8086. Именно Intel-8086 стал основой персональных компьютеров на платформе IBM PC, распространенных во всем мире. С выходом IBM PC произошла техническая революция не только в средствах вычислительной техники, но и во всех отраслях, перечислять которые займет много времени. Фирма Intel и ее микропроцессоры в течение многих лет были основой, на которых создавались новые модели IBM PC, и не только персональные компьютеры этой фирмы. До конца прошлого века фирмы, специализирующиеся на выпуске микропроцессоров, могли создавать только дешевые клоны микропроцессоров Intel, такие как Intel 286, Intel 386.

Фирма Intel в основу своих микропроцессоров положила архитектуру IA-32 (Intel Architecture-32) с системой команд X86, которая на долгие годы была промышленным аппаратом для всех 32-разрядных микропроцессоров (от Intel 386).

В табл. 11.1 приведены характеристики процессоров фирмы Intel с архитектурой IA-32.

Таблица 11.1

Модель, начало выпуска	Число транзисторов	Тактовая частота, МГц	Объем внутренней кэш-памяти
1386, октябрь 1985 г.	275 тыс.	до 40	нет
1486, апрель 1989 г.	1,2 млн.	до 100	8 Кбайт — команды 8 Кбайт — данные
Pentium, март 1993 г.	3,1 млн.	до 200	8 Кбайт — команды 8 Кбайт — данные
Pentium Pro, ноябрь 1995 г.	5,5 млн.	до 200	8 Кбайт — команды 8 Кбайт — данные
Pentium MMX, январь 1997 г.	4,5 млн.	до 233	8 Кбайт — команды 8 Кбайт — данные
Pentium II, май 1997 г. (Xeon, июнь 1998 г.)	7,5 млн.	до 450	16 Кбайт — команды 16 Кбайт — данные
Celeron, апрель 1998 г.		до 750	128 Кбайт — общий
Pentium III, февраль 1999 г. (Xeon, март 1999 г.)	8,5 млн.	до 1000 (до 700)	16 Кбайт — команды 16 Кбайт — данные
Pentium 4, ноябрь 2000 г. (Foster, 2001 г.)	42 млн.	до 1500 (до 2000)	256 Кбайт — общий 12 К — микрокоманды 8 Кбайт — данные

Первые модели процессоров Pentium выполняли только поочередную обработку данных о принципе "одна команда — одни данные" (SISD, Single Instruction — Single Data). С созданием микропроцессора Pentium MMX был реализован принцип групповой обработки "одна команда — много данных" (SIMD, Single Instruction — Multiple Data).

В 1999 году фирмы AMD (Advanced Micro Devices) и Cyrix приступили к выпуску собственного оригинального процессора K7 Athlon, который поддерживал систему команд x86. Однако эти процессоры не были обременены грузом генетических проблем архитектуры IA-32 и успешно конкурировали с продукцией фирмы Intel по цене, тактовой частоте и производительности.

В 2000 году фирма Intel выпустила процессор Pentium III (ядро Coppermine, кэш-память объемом 2 Мбайт), а фирма AMD воспроизвела процессор Athlon на основе Thunderbird с кэш-памятью L-2256 Кбайт. Эти процессоры работали на частоте 1,3 ГГц. Однако продукция фирмы AMD была дешевле.

В этом же 2000 году фирма Intel представила процессор Pentium 4 с принципиально новым ядром Netburst (пакетно-сетевая), ориентированным на эффективную работу с интернет-приложениями. В Pentium IV используется гиперконвейерная технология выполнения, позволяющая одновременно выполнять 20 команд, находящихся на различных стадиях их реализации.

На кристалле процессора Pentium 4 располагается 42 млн транзисторов, выполненных по КМОП-технологии с топологическими нормами 0,18 мкм. Кристаллы работают на частоте 1,5 ГГц и размещаются в 423 выводных корпусах типа PPGA (Plastic Pin Grid Array). В перспективе планируется переход на 0,13 мкм технологию с использованием шестислойной системы медных соединений. Это позволит обеспечить работу микропроцессора на тактовой частоте 3 ГГц.

Микропроцессор Pentium 4 предназначен для работы с ОЗУ типа RAMBUS (RDRAM) с частотой системной шины 400 МГц. Это дорогая память.

В это же время процессоры фирмы AMD Athlon работают с ОЗУ типа DDR SDRAM на скоростях до 266 МГц и поэтому гораздо дешевле.

Для улучшения своих позиций на рынке фирма Intel снизила цены вдвое и тем самым атаковала фирму AMD, на что AMD ответила разработкой нового процессора с ядром Palomino, работающим на частоте 1,533 ГГц. Одновременно AMD выпустила процессор Duron, который по цене начал теснить Celeron.

В последние годы появилась тайваньская компания VIA, которая уже поглотила фирму Cyrix и выпустила микропроцессор VIA Cyrix III (Samuel) по 0,15 мкм технологии. Тактовая частота этого процессора составляет 733 МГц, и он имеет 128 Кбайт кэш-памяти первого уровня и 64 Кбайт кэш-памяти второго уровня. Отличительными особенностями этого процессора являются рекордно низкое энергопотребление (5 Вт) и низкая цена.

11.3. Микропроцессорные системы

Микропроцессоры и устройства на его основе в современной микроэлектронике являются наиболее значимыми и распространенными изделиями.

Однако применение микропроцессора без системной поддержки ряда других устройств употребляется весьма ограничено.

Микропроцессор, прежде всего, нуждается в расширении оперативной постоянной памяти. Оперативное ЗУ и запоминающее устройство с произвольной выборкой (ЗУПВ) используется с микропроцессорами для хранения системных переменных и в качестве рабочей области поля оперативной памяти. ОЗУ в операционной системе требуется для размещения управляющей программы в целях организации стека для временного хранения

данных. Часть ОЗУ предназначена для оперативного хранения небольших программ и данных. При использовании растрового дисплея часть ОЗУ применяется для экранной памяти.

Постоянное запоминающее устройство ПЗУ в микропроцессорных системах используется для хранения управляющих программ, операционных систем и интерпретаторов языков программирования высокого уровня. Поэтому применяются различные типы ПЗУ: программируемые ПЗУ с плавкими перемычками и масочные ПЗУ, стираемые ПЗУ, репрограммируемые ПЗУ.

Микроконтроллер представляет собой микропроцессорную систему, предназначенную для управления, которая выполнена на основе одной или нескольких микропроцессорных БИС и способна к программированию.

Другими словами, если к микропроцессору добавить ОЗУ и ПЗУ, можно сформировать микроконтроллер.

Для микроконтроллеров характерны малая потребляемая мощность, расширенные возможности работы с памятью, низкая стоимость.

Большинство современных микроконтроллеров имеют RISC-ядро (Reduced Instruction Set Computer, ядро с сокращенным набором команд) и соответствуют промышленному стандарту производительностью 5 млн операций в секунду.

В основе RISC-архитектуры лежат четыре основополагающих принципа:

- ☐ любая операция выполняется за один такт;
- ☐ система команд содержит минимальное число инструкций одинаковой длины;
- ☐ операция обработки данных реализуются в формате "регистр — регистр";
- ☐ результаты формируются со скоростью одно слово за такт.

Если в системные шины встроить интерфейс, генератор сигналов, дисплей и блок питания, то можно говорить о микро-ЭВМ.

Микро-ЭВМ представляет собой конструктивно завершенное вычислительное устройство, реализованное на базе микропроцессорного комплекта СБИС и оформленное в виде автономного устройства с источником питания, интерфейсом, устройством отображения информации и комплектом программного обеспечения.

Микро-ЭВМ получили название персональных компьютеров (ПК) и широко распространены во всех сферах человеческой деятельности.

На рис. 11.2 представлена структурная схема (архитектура) микро-ЭВМ, сформированная на базе микропроцессоров и микропроцессорного комплекта БИС.

На базе микропроцессоров создаются мультипроцессорные системы, в которых используются более одного микропроцессора и которые предназначены для обработки информации, в том числе и для параллельной обработки.

Указанное архитектурное решение вычислительной машины повторяет классическую фон-неймановскую систему, в которой исповедуется принцип последовательного обмена информацией между процессором и памятью. Однопроцессорная архитектура фон Неймана имеет предел скорости обработки информационного массива. Этот предел, в частности, определяется скоростью распространения электрического сигнала по физическим линиям связи между отдельными узлами вычислительного устройства.

Важной является задача разработки такой архитектуры, которая позволяла бы предельно распараллелить процесс обработки информационного массива.

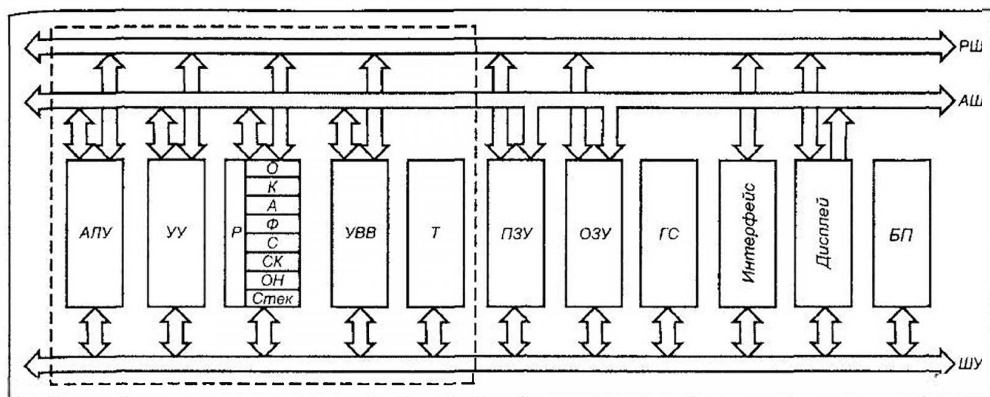


Рис. 11.2. Архитектура микро-ЭВМ (штриховкой выделен микропроцессор)

Разработан новый тип микропроцессора — *транспьютер*, позволяющий осуществлять обработку больших информационных массивов. Транспьютер содержит процессор с архитектурой RISC, имеющий быстродействие более 10^6 опс/с, а также аппаратные средства, обеспечивающие параллельные вычисления. На кристалле транспьютера размещается память, а также подключение внешнего ЗУ.

Транспьютер содержит четыре высокоскоростных канала связи, которые служат для обмена сообщениями с другими транспьютерами. Благодаря четырем линиям связи транспьютеры можно соединять между собой в различные транспьютерные сети. Для транспьютерных сетей разработана специальная, так называемая *ползающая* программа. Она загружается от управляющей ЭВМ в первый транспьютер сети (рис. 11.3).

По мере продвижения по сети программа может обрабатывать данные, возвращаемые пользующей программой и интерпретировать их на экране дисплея. Программа обработки данных позволяет процессору А анализировать линии связи и загружать другие процессоры, например процессор В. Когда процессор В завершит анализ своих связей, он найдет процессор С и зарегистрирует его.

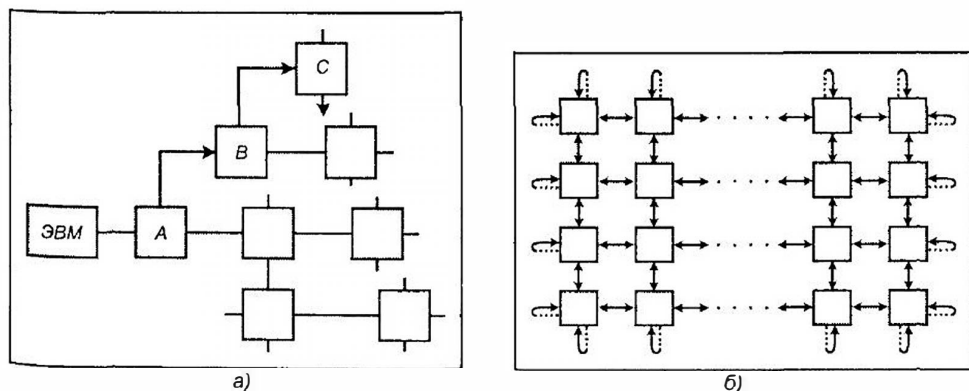


Рис. 11.3. Последовательность загрузки ползающей программы (а) и одна из моделей реализации транспьютерной сети (б)

Все процессоры анализируют свои связи и распараллеливают поток данных для обработки другими процессорами.

Набор аппаратных средств транспьютеров позволяет создавать микропроцессорные (транспьютерные сети со сложными межсоединениями).

На рис. 11.3, б приведена прямоугольная матрица транспортной сети. Однако можно включить каналы связи и по другой схеме.

Чтобы каждый транспьютер мог эффективно использоваться, необходимо создать такие конфигурации сети, которые соответствовали бы поставленной задаче обработки данных.

По вычислительной мощности компьютеры условно делятся на персональные и суперкомпьютеры.

Персональные компьютеры ориентированы как на широкое применение, так и на решение специальных задач, например, в бортовых системах.

Суперкомпьютер — вычислительное устройство общего назначения, выполняющее большие вычислительные задачи с числом операций порядка триллиона.

Производительность или вычислительная мощность суперкомпьютеров измеряется в следующих единицах:

□ мегафлоп (MFlops) — миллион операций в секунду;

□ гигафлоп (GFlops) — миллиард операций в секунду;

□ терафлоп (TFlops) — триллион операций в секунду.

Суперкомпьютеры являются стратегическим товаром и редко пересекают границы государств-производителей.

Принято считать, что эру суперкомпьютеров открыла матричная система ILLIAC IV, созданная NASA в Иллинойском университете (США). Производительность этой машины составила 20 MFlops.

Последовавшие затем многочисленные попытки совершенствовать матричную структуру показали ограниченность области применения таких матричных супер-ЭВМ.

В 1972 году после раскола компании Control Data ее ведущий сотрудник Саймур Край организовал собственную фирму Cray Research. Она захватила две трети рынка средств вычислительной техники сверхвысокой производительности. В 1974 году увидел свет первый суперкомпьютер CRAY-1. В основе его лежало три принципа: векторно-конвейерная архитектура, блестящая инженерная разработка архитектуры и оптимизированные программные средства. Затем появились CRAY-2, CRAY-3, CRAY-4 и CRAY-Y-MPC90. Еще позднее появились супер-ЭВМ японского производства, которые уступали по производительности, но были дешевле.

Произошла "микропроцессорная революция", сущность которой заключалась в объединении сотен и более стандартных микропроцессоров в вычислительную систему сверхвысокой производительности.

Так в системе ТЗЕ фирмы Gray Research используется 2048 микропроцессоров, а пиковая производительность такой супер-ЭВМ достигает 2,5 TFlops. Создание и использование суперкомпьютеров является одной из стратегических задач государства, позволяющая ему обеспечить независимость как при проведении фундаментальных исследований, так и при создании высокотехнологической продукции.

Очевидно, что для создания отечественных суперкомпьютеров и сокращения отставания нашей страны в этой области необходима государственная программа. В 2001 году в России был введен в строй суперкомпьютер MBC-100M с пиковой производительностью в

1 TFlops (10^{12} операций с плавающей запятой). Машина была реализована на сотнях микропроцессоров Alpha 21164 фирмы Samsung.

Эти примеры и фрагменты истории создания суперкомпьютеров свидетельствуют об огромном интересе к микропроцессорам — одному из главных продуктов микроэлектроники.

Задачи и упражнения

11.1. Найдите в справочниках структурную схему микропроцессора МП КР580ВМ80 (рис. 11.4). Объясните назначение основных узлов: АЛУ, аккумулятора, РОИ, РК, РП, РА, СК, указателя стека.

Решение

Центральный процессорный элемент КР580ВМ80 (сокращенно КР580) представляет собой однокристалльный восьмиразрядный МП с фиксированной системой команд, в котором совмещены операционное, управляющее устройство и сверхоперативная память (1, 2, 5, 8). Шина данных — 8-разрядная двунаправленная, адресная шина — 16-разрядная однонаправленная, обеспечивает возможность обращения к памяти, содержащей 64 К ячеек.

Сигналы, необходимые для управления МПС, снимаются частично с управляющих выходов МП, частично — с шины данных. МП допускает использование 256 портов ввода и 256 портов вывода.

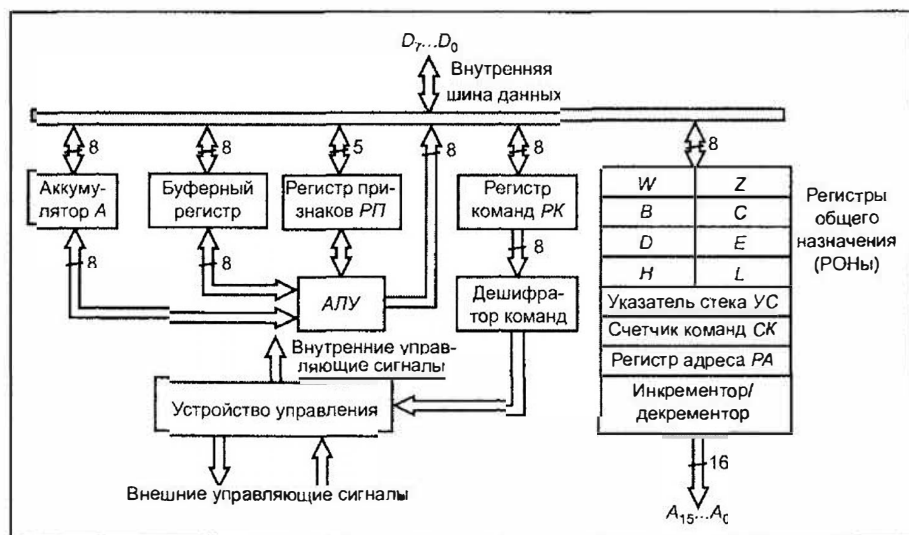


Рис. 11.4. Структурная схема МП КР580ВМ80

Основные характеристики МП:

- число команд — 78;
- максимальная тактовая частота — 2 МГц;

- ☐ время выполнения команд — 2—9 мкс;
- ☐ напряжения питания — + 5, -5, + 12 В;
- ☐ число вводов — 40.

Система команд МП1 приведена в приложении [см. литературную ссылку 4].

Контрольные вопросы

1. Что такое микропроцессор и как он структурно сформирован?
2. Нарисуйте структуру МП.
3. Нарисуйте структуру микро-ЭВМ.
4. Что такое универсальный микропроцессор?
5. Что такое специальный микропроцессор?
6. Что такое микропроцессорный комплект? Приведите пример.
7. Что такое микроконтроллер?
8. Что такое микро-ЭВМ?
9. Что такое транспьютер?
10. Что такое супер-ЭВМ?

Рекомендуемая литература

1. Алексенко А. Г. Основы микросхемотехники. — 3-е изд. — М.: ЮНИМЕДИАСТАЙЛ, 2002.
2. Гольденберг Л. М., Малев В. А., Малько Г. Б. Цифровые устройства и микропроцессорные системы. Задачи и упражнения: Учебное пособие для вузов. — М.: Радио и связь, 1992.
3. Вычислительная и микропроцессорная техника: Учебник для вузов. Под ред. Э. В. Евсеева. — М.: Радио и связь, 1991.
4. Калабеков Б. А. Микропроцессоры и их применение в системах передачи и обработки сигналов: Учебное пособие для вузов. — М.: Радио и связь, 1988.

12. Аналоговые интегральные схемы

12.1. Классификация аналоговых схем

Аналоговые интегральные схемы предназначены для преобразования сигналов, заданных в виде непрерывной функции.

Номенклатура аналоговых ИС включает различные генераторы сигналов, усилители, детекторы, задержки сигналов, модуляторы и т. д. Более подробная классификация приведена на рис. 6.3. В аппаратуре широко используются стандартные ИС указанных классов. В основе их конструкций лежат различные схематические решения, заимствованные из радиотехники дискретных элементов.

Технология аналоговых микросхем как самостоятельное направление микроэлектроники развивалось с определенным отставанием от технологии цифровых ИС. Среди причин такого отставания явился более ограниченный набор элементов полупроводниковых микросхем. В частности, в микроэлектронике не используются индуктивные элементы.

Однако особенности технологии микроэлектроники, позволяющей получать групповым способом на одной подложке совокупность элементов с взаимосогласованными характеристиками, позволяли создать широкую номенклатуру аналоговых ИС. Были разработаны типовые структуры, подобные интегральным логическим элементам в цифровых ИС, позволявшие унифицировать аналоговую микросхемотехнику.

Рассмотрим некоторые классы аналоговых схем.

Генератор — микроэлектронное устройство, предназначенное для создания электрических колебаний заданной формы и частоты.

В зависимости от формы формируемых колебаний различают генераторы гармоничных и релаксационных (импульсных) колебаний.

В зависимости от функционального назначения генератора радиотехническими методами и микроэлектронной техникой формируют соответствующие интегральные схемы (см. рис. 6.1).

Генераторы могут быть сформированы не только схематическими методами, но и с использованием приборов функциональной электроники. Так на диодах Ганна может быть сформирован генератор ВЧ- и СВЧ-колебаний.

Усилитель — микроэлектронное устройство, предназначенное для усиления сигналов в заданном диапазоне частот.

Различают усилители постоянного тока, высокочастотные усилители, широкополосные, низкочастотные, усилители промежуточной частоты, видеоусилители, усилители считывания и воспроизведения.

Эти типы усилителей, как правило, в микроэлектронику пришли из классической радиотехники и благодаря использованию микроэлектроникой технологии стали бытовыми с низким энергопотреблением, оформленными в виде интегральных схем.

Особое внимание с точки зрения идеологии микроэлектроники заслуживают дифференциальные и операционные усилители. Они будут рассмотрены подробнее.

Детектор — микроэлектронное устройство, служащее для преобразования электрических колебаний, в результате чего из него выделяются нужные составляющие амплитудно-частотных характеристик спектра сигналов.

В частотном и фазовом детекторах частотно-модулированные и фазово-модулированные колебания преобразуются в амплитудно-модулированные колебания, которые затем детектируются.

В этих интегральных схемах широко используются разработанные в радиотехнике схемы, которые с помощью микроэлектронной технологии облечены в форму интегральных схем.

Селектор импульсов представляет собой микроэлектронное устройство, предназначенное для выделения из множества импульсов только тех, которые обладают заданными свойствами.

Например, амплитудные селекторы выделяют только те импульсы, амплитуда которых превышает заданный уровень или порог селекции. Может стоять и задача выделения импульсов, которые не достигают порога селекции. Схемы селекторов весьма разнообразны и выполняются по микроэлектронной технологии в виде интегральных схем.

Фильтр — микроэлектронное устройство, предназначенное для разделения электрических колебаний различных частот.

Из спектра поданных на вход электрических колебаний фильтр выделяет или пропускает на выход только колебания в заданной области частот, называемой *полосой пропускания*. Все остальные колебания фильтром подавляются.

По виду частотной характеристики фильтры делятся на фильтры верхних частот, пропускающие колебания с частотами выше заданной. Фильтры нижних частот пропускают колебания не выше заданной граничной частоты.

Полосовые фильтры пропускают колебания в заданном интервале частот. Режекторные или заграждающие фильтры задерживают колебания в заданной полосе частот. В радиотехнике достаточно хорошо разработаны методы расчета фильтров и их синтеза. Микроэлектронная технология позволяет создать фильтры в виде интегральных схем.

Модулятор — микроэлектронное устройство, осуществляющее управление заданным параметром колебательного процесса в соответствии с сигналами передаваемого сообщения.

Воздействие модулирующих сигналов на параметры модулируемых колебаний осуществляется посредством нелинейного управляющего элемента. Различают амплитудные, частотные, фазовые, импульсные и другие типы модуляторов.

Преобразователь — микроэлектронное устройство, предназначенное для преобразования параметров сигналов: частоты, фазы, длительности мощности и т. д.

Эти преобразования осуществляются методами радиотехнической схемотехники. Особое место в этом классе приборов занимает аналого-цифровые преобразователи (АЦП) и цифроаналоговые преобразователи (ЦАП). Эти устройства позволяют стыковать приборы, обрабатывающие информацию в аналоговой и цифровой формах. ЦАП и АЦП являются сугубо микроэлектронными устройствами, реализованными в виде интегральных схем.

12.2. Операционный усилитель

12.2.1. Дифференциальный усилитель

Операционный усилитель представляет собой микронэлектронное устройство, предназначенное для усиления как постоянного тока, так и электрических колебаний.

Операционный усилитель обязательно имеет внешние цепи, предназначенные для выполнения некоторых линейных и нелинейных операций. В некотором смысле операционный усилитель является элементной базой для аналоговых преобразователей.

В основе схемы операционного усилителя лежит дифференциальный усилительный каскад (рис. 12.1).

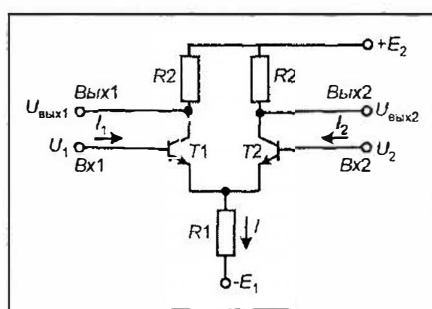


Рис. 12.1. Схема дифференциального усилителя

Два идентичных транзистора T_1 и T_2 и диффузионные резисторы R_1 и R_2 формируют два транзисторных усилителя, включенных симметрично. Отношение резисторов коллекторной цепи должно быть постоянным. У дифференциального усилителя есть два входа и два выхода.

При передаче на входы дифференциального усилителя одинаковых (синфазных) сигналов напряжение на входах практически не меняется.

Разность напряжений, формируемых на входах (базах транзисторов T_1 и T_2) называют *дифференциальным входным сигналом*, а полусумму этих напряжений — *синфазным входным сигналом*:

$$U_{\text{ин}} = U_1 - U_2;$$

$$U_{\text{сф}} = \frac{U_1 + U_2}{2}.$$

Выходным сигналом дифференциального усилителя является напряжение между коллекторами транзисторов.

В идеальном дифференциальном усилителе дрейф выходного напряжения отсутствует, хотя в каждом из плеч он может быть сравнимо большим. Симметрия не меняется при синфазном изменении U_1 и U_2 — дифференциальный усилитель нечувствителен к синфазному сигналу.

Появление на входе дифференциального сигнала приводит к нарушению симметричного режима работы усилителя.

Сигнал на входе дифференциального усилителя пропорционален разности входных токов

$$\Delta I = I_2 - I_1.$$

Синфазная составляющая входного сигнала определяется как часть входного тока, которая одинакова для каждого входа дифференциального каскада или

$$I_c = \frac{I_1 + I_2}{2}.$$

Тогда

$$I_1 = I_0 - \frac{\Delta I}{2},$$

$$I_2 = I_0 + \frac{\Delta I}{2}.$$

Наиболее важными параметрами дифференциального усилителя являются:

коэффициент усиления $K_u = \frac{dU_{вых}}{dU_{вх}} = \frac{\alpha I R_2}{2\varphi_T}$

и входное сопротивление $R_{вх} = \frac{dU_{вх}}{d(\Delta I)} = \frac{2\varphi_T(1+\beta)}{I},$

где α — коэффициент передачи тока транзистора в схеме с общей базой, β — коэффициент передачи тока базы биполярного транзистора в схеме с общим эмиттером, I — ток в эмиттерной цепи; R_2 — коллекторный резистор.

Тепловой потенциал φ_T определяется соотношением

$$\varphi_T = \frac{kT}{q}.$$

Дифференциальный усилитель мощности можно реализовать и на полевых транзисторах. Полевые транзисторы по сравнению с МДП-транзисторами обладают большей стабильностью характеристик и малым уровнем собственных шумов.

Схема на полевых транзисторах аналогична схеме на дифференциальном усилителе на биполярных транзисторах: подобная симметрия плеч, сохранение чувствительности усилителя к синфазному входному сигналу.

Дифференциальный усилитель на полевых транзисторах имеет высокое входное сопротивление. Это позволяет во входных цепях всегда обеспечить режим холостого хода, в котором разность потенциалов затворов полевых транзисторов совпадает с разностью напряжений источников входных сигналов.

12.2.2. Структурная схема операционного усилителя

Операционный усилитель имеет в основе дифференциальный усилитель, и поэтому способен реагировать только на дифференциальный сигнал.

Структурная схема операционного усилителя является типовой и отвечает основным принципам микросхемотехники (рис. 12.2).

Дифференциальный усилитель является базовой структурной единицей операционного усилителя. Дифференциальный усилитель может быть реализован на биполярных, поле-

вых транзисторах, а также на их сочетании при условии полной технологической и схематической совместимости.

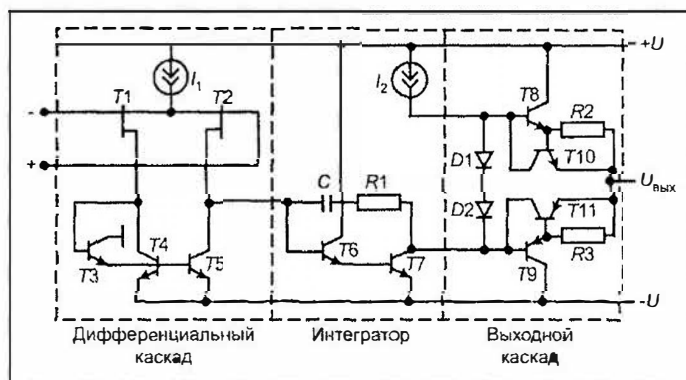


Рис. 12.2. Типовая схема операционного усилителя

Дифференциальный каскад реализован на полевых транзисторах T_1 и T_2 , входы которых условно обозначаются как инвертирующий, обозначаемый знаком "−", и неинвертирующий или прямой, обозначаемый знаком "+".

В данном случае дифференциальный усилитель выполнен по схеме с однофазным выходом, чтобы выходной ток обеспечивал перезарядку интегрирующей емкости C . Емкость интегрирующего конденсатора C должна быть такой, чтобы петлевой коэффициент усиления был меньше единицы.

В качестве коллекторной нагрузки дифференциального усилителя используются источники тока I_1 .

Согласование дифференциального каскада и интегратора достигается при условии сохранения симметрии плеч дифференциального усилителя. С этой целью погрешность работы отражателей тока, обусловленная базовыми токами транзисторов T_4 и T_5 , компенсируется входным током интегратора. Это возможно при полной идентичности параметров транзисторов T_4 , T_5 и T_7 и равенности токов I_1 и I_2 . В этом случае суммарный ток без транзисторов T_4 и T_5 будет совпадать с током базы транзистора T_7 .

Для увеличения статического коэффициента усиления операционного усилителя в базовую цепь интегрирующего транзистора T_6 для сохранения симметрии плеч в отражатель тока также включается эмиттерный повторитель, реализованный на транзисторе T_3 .

Выходной каскад выполнен по схеме двухтактного эмиттерного повторителя, реализованного на комплементарных вставках транзисторов T_{10} и T_{11} .

Это позволяет снизить мощность, рассеиваемую каскадом в статическом режиме.

Эмиттерный повторитель подключается к выходу интегратора через диоды смещения D_1 и D_2 .

Для уменьшения нелинейных искажений транзисторы в режиме покоя приоткрывают, применяя диодные смещения переходов "база—эмиттер" транзисторов T_8 и T_9 . Заметим, что в качестве диодов D_1 и D_2 , как правило, используются транзисторы в диодном включении.

Выходное напряжение операционного усилителя снимается с делителя, реализованного на резисторах-ограничителях R_2 и R_3 . Эти резисторы одновременно защищают выходные каскады от короткого замыкания.

Операционный усилитель питается от двухполярного источника питания.

Коэффициент усиления операционного усилителя K лежит в пределах 10^4 — 10^9 , входное сопротивление достигает 100 МОм, а выходное сопротивление составляет 10^2 Ом. Операционный усилитель имеет малый уровень собственных шумов, сильное подавление синфазной составляющей (≈ 60 дБ), широкую полосу пропускания от 0 до 10 МГц.

Вышеперечисленные свойства операционного усилителя позволяют получать высокую точность выполнения операций. Операционный усилитель является микроэлектронным устройством универсального применения. Операционные усилители конструктивно выполняются в виде интегральных схем средней степени интеграции.

12.2.3. Аппаратурные включения операционных усилителей

Использование операционных усилителей в виде интегральных схем позволяет повысить качество и надежность электронной аппаратуры, унифицировать комплектующие изделия.

В определенном смысле операционный усилитель можно представить как элементную базу, на основе которой можно построить большое число аналоговых устройств различного функционального назначения.

Различают линейные и нелинейные включения операционного усилителя (ОУ).

Линейные включения ОУ

Инвертирующий усилитель представляет собой операционный усилитель, охваченный обратной связью (рис. 12.3).

Из схемы включения следует, что $I_{вх} = I_0$. Поскольку $I_{вх} = U_{вх} / R_1$, а $I_0 = U_{вых} / R_2$, то

$$U_{вых} = -U_{вх} R_2 / R_1.$$

Тогда коэффициент усиления по напряжению K_n

$$K_n = -R_2 / R_1. \quad (12.1)$$

Ясно, что коэффициент усиления может быть выбран сколь угодно большим.

Если $R_1 = R_2$, то схема будет работать в режиме инвертирующего повторителя, выполняющего функцию преобразователя импедансов.

Неинвертирующий усилитель позволит усилить сигнал без изменения полярности (рис. 12.4).

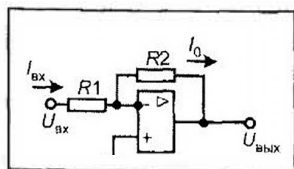


Рис. 12.3. Схема инвертирующего усилителя

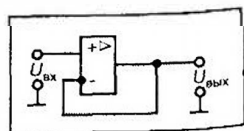
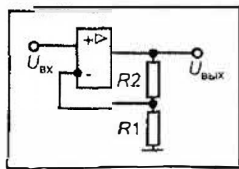


Рис. 12.4. Схема неинвертирующего усилителя (а) и повторителя напряжения на его основе (б)

Из рис. 12.4, а следует

$$U_{\text{ав}} = \frac{R_1}{R_1 + R_2} U_{\text{вх}}.$$

И тогда коэффициент усиления

$$K_u = 1 + \frac{R_2}{R_1}. \quad (12.2)$$

Другими словами, коэффициент усиления можно задать любым: от единицы ($R_1 \gg R_2$) до бесконечности ($R_2 \gg R_1$). Если $R_1 = \infty$, а $R_2 = 0$, то получаем схему повторителя напряжения (рис. 12.4, б).

Аналоговый интегратор предназначен для интегрирования сигналов (рис. 12.5).

В основе действия интегратора лежит накопление заряда на конденсаторе C под действием приложения тока I_2 . Для идеального ОУ ток, протекающий через сопротивление R , равен току заряда конденсатора C , а напряжение в точке соединения равно нулю.

$$U_{\text{вых}} = -U_c = -\frac{1}{RC} \int_0^t U_{\text{вх}}(t) dt; \quad (12.3)$$

где $\tau = RC$ — временная характеристика скорости заряда конденсатора.

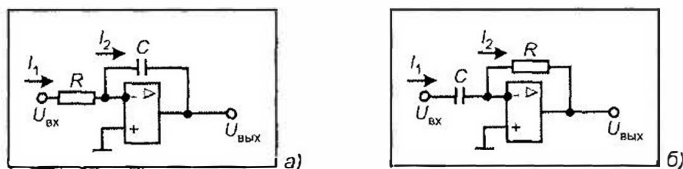


Рис. 12.5. Схема аналогового интегратора (а) и дифференциатора (б)

Дифференциатор предназначен для дифференцирования по времени электрических сигналов (рис. 12.5, б). Ток через емкость пропорционален производной приложенного к ней напряжения, т. е.

$$I_1 = \frac{dU_c}{dt},$$

где

$$U_c = U_{\text{вх}}, \quad I_1 = C \frac{dU_{\text{вх}}}{dt}, \quad U_{\text{вых}} \approx -RC \frac{dU_{\text{вх}}}{dt}. \quad (12.4)$$

Сумматор предназначен для проведения аналогового сложения. В зависимости от способа включения входных резисторов можно производить сложение или вычитание сигналов.

В основу структуры сумматора положены усилительные схемы.

В инвертируемом сумматоре (рис. 12.6, а)

$$I_0 = I_1 + I_2 + \dots + I_m; \quad U_{\text{вых}} = -I_0 R_0; \quad U_1 = I_1 R_1; \quad U_2 = I_2 R_2; \quad \dots \quad U_m = I_m R_m$$

или

$$U_{\text{вых}} = U_1 \frac{R_0}{R_1} + U_2 \frac{R_0}{R_2} + \dots + U_m \frac{R_0}{R_m}.$$

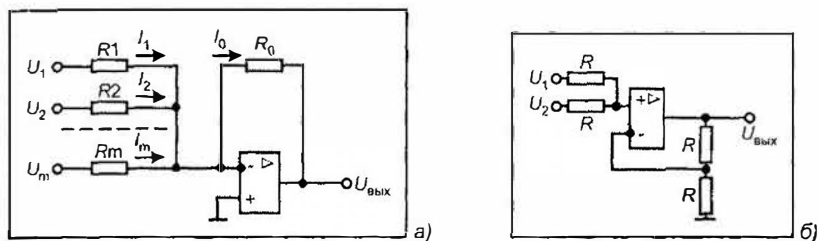


Рис. 12.6. Схема сумматора с инвертированием (а) и без инвертирования (б) сигнала

В неинвертирующем сумматоре в стационарном режиме

$$U_n = U_0 = U_{вых} \frac{R_1}{R_1 + R_2}.$$

Неинвертирующий вход ОУ не должен потреблять ток и тогда

$$\frac{U_1 - U_n}{R} + \frac{U_2 - U_n}{R} + \dots + \frac{U_m - U_n}{R} = 0.$$

Отсюда

$$U_1 + U_2 + \dots + U_m = m U_n = m U_{вых} \frac{R_1}{R_1 + R_2}$$

или

$$U_{вых} = \left(1 + \frac{R_2}{R_1}\right) \frac{1}{m} (U_1 + U_2 + \dots + U_m). \quad (12.5)$$

Другими словами, напряжение на выходе ОУ равно сумме напряжений на входе, умноженной на коэффициент $\frac{1}{m} \left(1 + \frac{R_2}{R_1}\right)$.

Для двух сигналов, если $R_1 = R_2$

$$U_{вых} = U_1 + U_2.$$

Вычитатель предназначен для вычитания сигналов, и его схема приведена на рис. 12.7.

В стационарном режиме $U_0 = U_n$.

$$U_n = U_2 \frac{R_2}{R_2 + R_2} = \frac{U_2}{2};$$

$$I_1 = \frac{U_1 - U_0}{R_1} = \frac{U_0 - U_{вых}}{R_1}.$$

Тогда

$$U_0 = U_n = \frac{U_2}{2},$$

$$U_2 = U_1 + U_{вых}.$$

Или

$$U_{вых} = U_2 - U_1. \quad (12.6)$$

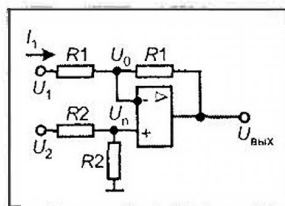


Рис. 12.7. Схема аналогового вычитателя

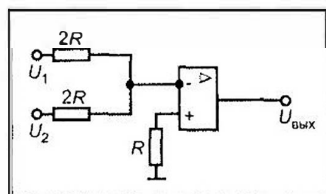


Рис. 12.8. Схема компаратора

Меняя значения резисторов в цепях, можно производить аналоговое вычитание с весовыми коэффициентами.

Компаратор предназначен для сравнения входных напряжений с опорным.

В зависимости от входных напряжений, на выходе можно получить напряжение, соответствующее логическому нулю или логической единице.

На рис. 12.8 представлена схема компаратора.

К неинвертирующему входу подключается резистор для уменьшения значения ошибки за счет входных токов ОУ.

Компаратор работает так, что

- при $U_1 + U_2 < 0$, $U_{вых} > 0$, что соответствует "1";
- при $U_1 + U_2 > 0$, $U_{вых} < 0$, что соответствует "0".

Нелинейные включения операционного усилителя

Логарифмический усилитель предназначен для логарифмирования входных и выходных сигналов.

В процессе логарифмирования используются нелинейные свойства вольтамперной характеристики р—п-перехода.

Ток через р—п-переход определяется зависимостью

$$I = I_0 \left[\exp \left(\frac{U}{m\phi_T} \right) - 1 \right],$$

где I_0 — ток насыщения, U — напряжение, $\phi_T = kT/q$ — термический потенциал, m — коэффициент, характеризующий поверхностную рекомбинацию, 1,0—1,3.

Схема логарифмического усилителя приведена на рис. 12.9, а.

Если $\exp \left(\frac{U}{m\phi_T} \right) \gg 1$, то для схемы логарифмического усилителя будем иметь

$$I_1 = \frac{U_{вых}}{R} = I_0 \exp \left(-\frac{U_{вых}}{m\phi_T} \right).$$

Тогда

$$\frac{U_{вых}}{m\phi_T} = -\ln \left(\frac{U_{вых}}{I_0 R} \right),$$

или окончательно

$$U_{\text{вых}} = a_0 \ln \left(\frac{U_{\text{вх}}}{I_0 R} \right),$$

где a_0 — константа. При переводе к десятичному логарифму получаем

$$U_{\text{вых}} = a \lg \left(\frac{U_{\text{вх}}}{I_0 R} \right),$$

где $a = -2,3 \cdot m\varphi_T$.

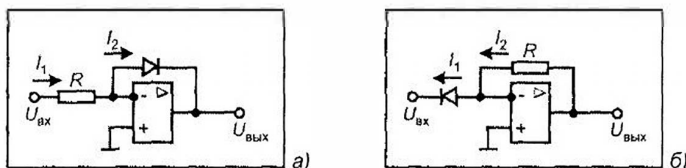


Рис. 12.9. Схема логарифмического усилителя (а) и вычислителя антилогарифмов (б)

Для вычисления антилогарифмов используется схема, представленная на рис. 12.9, б. Запишем очевидное уравнение:

$$I_1 = I_0 \exp(-U_{\text{вх}} / m\varphi_T) = I_2 = U_{\text{вых}} / R.$$

Тогда

$$m\varphi_T \ln U_{\text{вых}} = U_{\text{вх}} - m\varphi_T \ln(I_0 R)$$

или

$$\ln U_{\text{вых}} = \ln(I_0 R).$$

Окончательный вид выражения для выходного напряжения следующий:

$$\lg U_{\text{вых}} = a U_{\text{вх}} + a,$$

где $a = 1 / (2,3 \cdot m\varphi_T)$, $a = \lg(I_0 R)$.

Логарифмический умножитель представляет собой устройство умножения двух или более аналоговых членов путем использования сложения логарифмов этих сигналов с последующим их использованием. Логарифмический умножитель является аналогом логарифмической ячейки (рис. 12.10).

Два аналоговых сигнала вводятся в параллельные логарифмические усилители, на входе которых имеются значения $\lg x$ и $\lg y$. Прологарифмированные сигналы подаются на сумматор, на выходе которого получаем логарифм произведения этих сигналов $\lg(x \cdot y)$.

Сумматор представляет собой ранее рассмотренную схему неинвертирующего сумматора, позволяющего получить на выход значение $\lg x + \lg y = \lg(x \cdot y)$.

Далее этот сигнал поступает на усилитель, вычисляющий антилогарифмы. Выражение $\lg(x \cdot y)$ потенцируется, и на выходе получаем $Z = X \cdot Y$ — произведение двух аналоговых сигналов, поступивших на вход логарифмического умножителя.

Существует большое число схем включения операционного усилителя, позволяющих производить различные операции с аналоговыми сигналами.

Приведенные примеры достаточно полно проиллюстрировали возможности операционных усилителей, которые являются основой аналоговых вычислительных устройств.

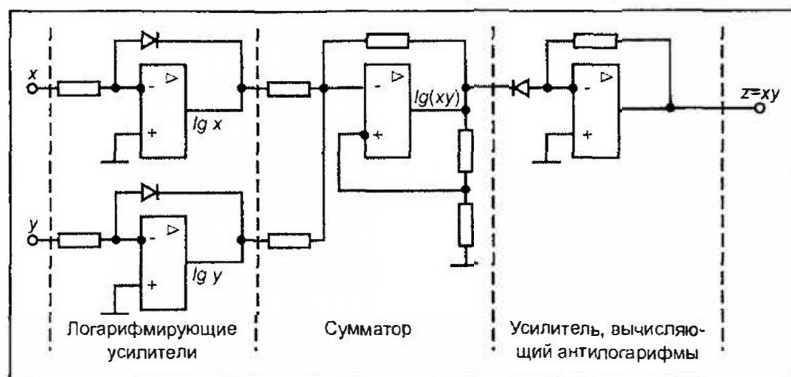


Рис. 12.10. Схема логарифмического умножителя

12.3. Преобразователи типа ЦАП — АЦП

12.3.1. Цифроаналоговый преобразователь

Цифроаналоговый преобразователь (ЦАП) представляет собой микросэлектронное устройство для автоматического преобразования числовых кодов в эквивалентные им значения определенной физической величины.

Коды обычно представляются в двоичной, десятичной и какой-либо другой системе исчисления.

Выходные физические величины представляются в виде временных интервалов, угловых перемещений, напряжений или токов и т. п.

Существуют разнообразные конструкции ЦАП.

На рис. 12.11 представлена схема ЦАП на основе делителя типа $R-2R$ и операционного усилителя. Ключи A , B , C и D подключают в нужном порядке резисторы $2R$ к источнику эталонного напряжения, когда соответствующий разряд двоичного числа равен единице. Если разряд числа равен нулю, то ключ замыкается на землю. На рис. 12.11 положение ключей соответствует числу 1101. На инвертирующем входе $\ominus U$ создаются напряжения, соответствующие делению эталонного напряжения.

Операционный усилитель работает по схеме неинвертирующего сумматора (см. рис. 12.6, б).

12.3.2. Аналого-цифровой преобразователь

Аналого-цифровой преобразователь (АЦП) является микросэлектронным устройством, осуществляющим автоматическое преобразование непрерывно меняющегося аналогового сигнала в цифровой код. Процесс аналого-цифрового преобразования включает этапы дискретизации (квантования) непрерывного сигнала по времени, по уровню или по обоим параметрам одновременно. Квантовый сигнал в дальнейшем кодируется.

Часто в качестве исходной величины используются также основные сигналы, такие как напряжение или ток, частота или фаза электрических колебаний.

Процессы квантования и кодирования в микросэлектронных АЦП производится с помощью аналоговых устройств в микросэлектронном исполнении. На рис. 12.12 приведена

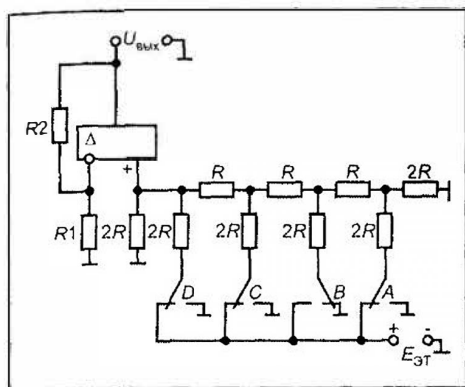


Рис. 12.11. Схема цифроаналогового преобразователя

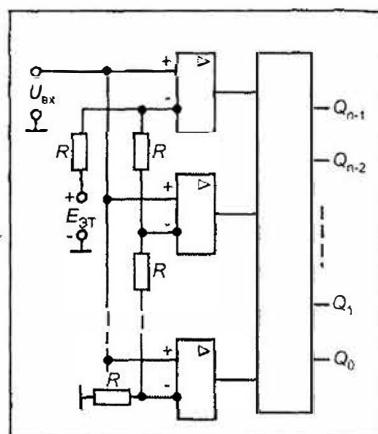


Рис. 12.12. Схема аналого-цифрового преобразователя

схема АЦП параллельного кодирования, позволяющая преобразовывать аналоговое напряжение в n -разрядное двоичное число.

Эталонное напряжение $E_{эт}$ с помощью резисторного делителя делится на 2^n градаций и подается на инвертирующие входы операционных усилителей. Для 2^n градаций должно быть $2^n - 1$ операционных усилителей. Не инвертирующие входы ОУ соединены и на них подается входное аналоговое напряжение, равное максимально возможному.

На инвертирующем входе верхнего усилителя напряжения соответственно будет

$$E_{эт} \frac{2^n - 1}{2^n} \text{ и т. д.}$$

Операционные усилители включены по схеме компараторов и позволяют сравнивать аналоговое напряжение с частью эталонного напряжения.

Если аналоговое напряжение превышает напряжение на инвертирующем входе ОУ, то на выходе появляется положительное напряжение. Положительный сигнал интерпретируется как логическая единица.

В противном случае на выходе ОУ появится отрицательное напряжение, интерпретируемое как логический ноль.

Приоритетный шифратор является комбинационной схемой. Эта схема формирует на выходе двоичное число по старшему входу, имеющему логическую единицу. В схеме рис. 12.12 старшим является верхний вход шифратора, затем второй и т. д.

Быстрота преобразований по такой схеме оборачивается сложностью коммутации большого числа операционных усилителей.

Задачи и упражнения

12.1. Проведите моделирование двухвходового сумматора в программе Electronics Workbench. При указанных на рис. 12.13 параметрах проведите расчет и сравните полученные данные с результатом моделирования.

12.2. Проведите моделирование в программе Electronics Workbench дифференциатора и снимите его АЧХ (рис. 12.14).

Исследуйте влияние на его АЧХ коэффициента усиления при разомкнутой отрицательной обратной связи.

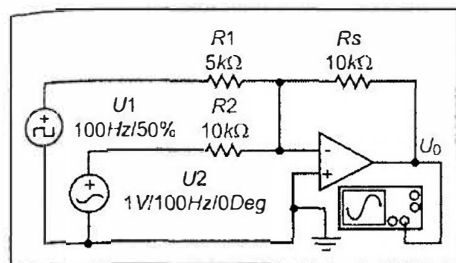


Рис. 12.13. Схема двухвходового сумматора

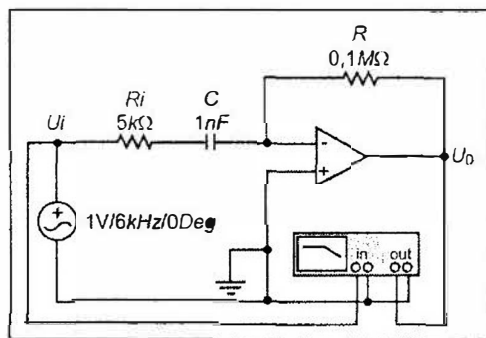


Рис. 12.14. Схема дифференциатора

12.3. Соберите схему инвертирующего усилителя в программе Electronics Workbench и проверьте формулу (12.1).

12.4. Соберите схему неинвертирующего усилителя в программе Electronics Workbench и проверьте формулу (12.2).

12.5. Соберите схему сумматора в программе Electronics Workbench и проверьте формулу (12.5).

12.6. Соберите схему вычитателя в программе Electronics Workbench и проверьте формулу (12.6).

Контрольные вопросы

1. Что такое аналоговые ИС?
2. Какие ИС относятся к аналоговым?
3. Что такое дифференциальный усилитель?
4. Что такое синфазный сигнал?
5. Что такое операционный усилитель?
6. Нарисуйте схему инвертирующего усилителя и напишите значение его коэффициента усиления.
7. Нарисуйте схему неинвертирующего усилителя и напишите значение его коэффициента усиления.
8. Нарисуйте схему сумматора и напишите значение суммы.

Рекомендуемая литература

1. Аваев Н. А., Наумов Ю. Е., Фролкин В. Т. Основы микроэлектроники: Учебное пособие для вузов. — М.: Радио и связь, 1991.
2. Алексеев А. Г. Основы микросхемотехники. — 3-е изд. — М.: ЮНИМЕДИАСТАЙЛ, 2002.
3. Карлачук В. И. Электронная лаборатория на IBM PC. Программа Electronics Workbench и ее применение. — 3-е изд. — М.: СОЛОН-Пресс, 2003.
4. Степаненко И. П. Основы микроэлектроники: Учебное пособие для вузов. — 2-е изд. — М.: Лаборатория базовых знаний, 2000.

13. Базовые матричные кристаллы

13.1. Классификация базовых матричных кристаллов

Развитие технологии и схемотехники СБИС сопряжено как с ростом степени интеграции и объемов производства, так и с повышением стоимости разработок, а также стоимости обработки одного бита информации. До появления САПР противоречия в этих тенденциях решались путем создания строго специализированных БИС, нацеленных на выполнение определенных функций. В настоящее время широко применяется другой подход, основывающийся на достоинствах программного и аппаратного подходов: программное начало закладывается в архитектуру и алгоритмы обработки, аппаратная — в параметры и архитектуру БИС. Такие БИС получили название *базовых матричных кристаллов* (БМК). Они сняли с повестки дня ряд проблем, характерных для развития микроэлектроники. Например, увеличение сложности микросхемы делает ее более специализированной, уменьшает гибкость ее применения. БМК же позволяют гибко использовать архитектуру БИС под конкретную задачу.

При производстве БМК реализуется идея, высказанная еще в 80-х годах Гордоном Муром из фирмы Intel и заключающаяся в создании "кремниевых мастерских". В соответствии с этой идеей проектирование осуществляет потребитель, а изготовление БИС в "кремниевой мастерской". Такой подход позволяет снизить стоимость производства, создать многофункциональные схемы их типовых ячеек, решить проблему оптимизации числа внешних выводов схемы, использовать ранее разработанные и многократно используемые конструктивные схемотехнические решения.

Специализированные БИС придают изделиям уникальные свойства, решают конкретные целевые функции.

На рис. 13.1 приведена классификация специализированных БИС по способу проектирования.



Рис. 13.1. Классификация специализированных БИС

Полузаказная интегральная схема представляет собой класс интегральных схем, имеющих постоянную и переменную части. Постоянная часть представляет собой заранее

спроектированную схему. Переменная часть интегральной схемы определяется требованиями заказчика. К этим типам ИС относятся БИК и ПЛИС.

Заказная интегральная схема представляет собой класс интегральных схем, содержащих стандартные или специально созданные элементы или узлы по заранее заданной функциональной схеме.

Существуют два способа изготовления заказных ИС: на основе стандартных элементов или полностью заказные. Стандартные элементы выбираются из заранее спроектированной библиотеки элементов.

Функциональная схема вычислительных структур и систем требует использования сложных СБИС, проектирование которых не всегда экономически оправдано. Зачастую "приспосабливают" готовую схему в нужную СБИС. Это сопряжено с риском неточной адаптации схемы. Когда речь идет о цифровых схемах, то такая проблема может быть решена с помощью *программируемых пользователем логических интегральных схем* (ПЛИС). Технология изготовления ПЛИС позволяет обеспечить высокую степень интеграции (10^5 — 10^6 элементов/кристалл), высокое быстродействие (10^8 оп/с), возможность реализации в одном корпусе комбинационных И/ИЛИ или последовательных схем. К первому поколению ПЛИС относятся два типа структур: *вентильные матрицы* и *программируемые логические матрицы*.

Базовые матричные кристаллы состоят из нескольких основных элементов, расположенных на кристалле в зависимости от его конструктивно-технологического исполнения (рис. 13.2).

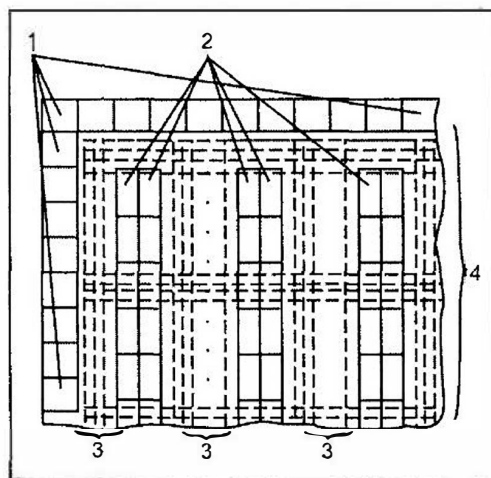


Рис. 13.2. Фрагмент структурной схемы базового матричного кристалла; 1 — периферийные ячейки и контактные площадки; 2 — базовые ячейки; 3 — вертикальные трассы в первом слое металлизации; 4 — горизонтальные трассы во втором слое металлизации

Матрица базовых ячеек занимает основную площадку кристалла и состоит из отдельных ячеек, выполняющих определенные функции. Различают конструктивно однородную матрицу, состоящую из функционально-однородных или неоднородных ячеек. Неоднородная матрица имеет ячейки, либо фиксированные по определенной координате, либо имеющие неодинаковые размеры. Базовые ячейки представляют собой определенные конструктивно-технологические варианты логических ячеек, ячеек памяти и т. п. Базовые ячейки группируются в макроячейки в виде группы из четырех симметрично расположенных ячеек, либо в линейку ячеек. Выбор конфигурации оптимален в том случае, если

все выводы сосредотачиваются на границах. Большинство выводов ячеек имеют эквивалентные пары на противоположных сторонах ячейки, что существенно облегчает последующую трассировку. В матрице могут быть размещены специальные буферные ячейки, микроячейки, реализующие типовые функциональные узлы, например, запоминающие устройства, регистры, аналоговые блоки и т. п.

Многослойная система металлических проводников, соединяющая базовые ячейки в соответствующие функциональные структуры и подключающая их к шинам электрического питания, носит названия *трасс*, которые располагаются на нескольких уровнях. Вертикальные трассы располагаются в первом слое металлизации, горизонтальные — во втором. В ряде разработок трассировка производится и в большем числе слоев.

Трассы межсоединений могут различаться емкостью каналов, а также конструктивным исполнением. Помимо алюминиевых тонкопленочных шин используются поликремниевые шины. В конструкциях БМК на МОП-транзисторах с молибденовыми затворами используются молибденовые шины. Наиболее низкоомные электрические соединения формируют в наружном слое.

Необходимые электрические контакты между шинами различных технологических слоев осуществляются в местах пересечения вертикальных и горизонтальных отрезков трасс. Подключение источника питания к ячейкам матрицы может быть как индивидуальным, так и групповым.

Вспомогательные схемы (схемы обрамления) и контактные площадки, расположенные по периферии кристалла, представляют собой схемы контроля и диагностики, источники опорного напряжения, трансляторы уровней и т. п. Все схемы связаны с контактными площадками и через них осуществляется ввод-вывод информации.

Элементной базой БИК служат различные варианты логических схем, ячеек памяти, расположенных в одном или двух ярусах. Существует большое число конструктивно-технологических решений схем базовых ячеек (рис. 13.3).

Процесс развития технологии изготовления микросхем стимулировал работы по объединению в одном кристалле биполярных и МОП-транзисторных структур. Это направление получило название БиМОП-технологии, а при использовании комплементарных транзисторов — БиКМОП-технологии.

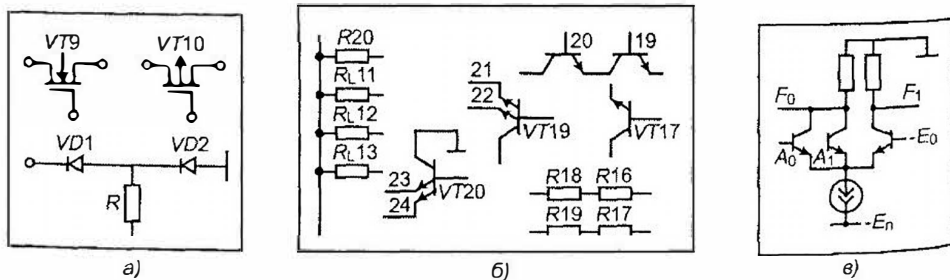


Рис. 13.3. Примеры наборов элементов для базовых ячеек на МОП-транзисторах (а), ТТЛ логике (б) и ЭСЛ логике (в)

Главным достоинством этих приборов микроэлектроники является соединение воедино преимуществ МОП и биполярных транзисторных структур. Действительно, высокая плотность интеграции, низкая рассеиваемая мощность, высокая помехоустойчивость,

характерные для МОП-структур, сочетаются с хорошей выходной нагрузочной способностью и высоким быстродействием биполярных структур.

Концепция построения базовой ячейки вентиля на БикМОП-структурах основана на дополнении к стандартному КМОП-вентилю усилителя формирователя тока на биполярных транзисторах (рис. 13.4). Аналогично можно сформировать БикМОП-инвертор. На основе таких элементов формируют схемы вентиляей, применяемых в матричных кристаллах. В БикМОП снижение тока стока МОП-транзисторов может быть компенсировано улучшением коэффициента усиления биполярных транзисторов.

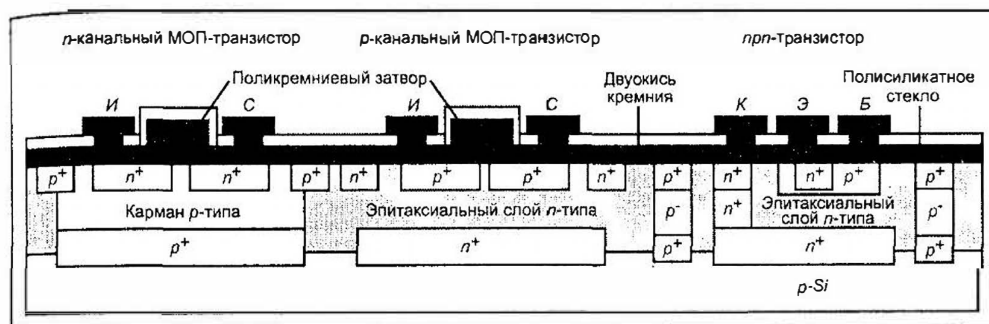


Рис. 13.4. Вариант БикМОП-технологии для БМК

Таким образом, основой для всех базовых ячеек являются вентили различных конструкций. Говорят о море вентиляей в БМК.

Достоинства вентиляльных матриц состоят в следующем: конкретная схема может реализовываться при небольшом числе технологических операций, регулярность структуры позволяет потребителю быстро разработать свою конкретную подсистему с высокой вероятностью функционирования (95%). Таким образом, вентиляльные матрицы представляют собой некоторые полуфабрикаты изделий микроэлектроники, на которые нужно нанести разводку по требованию заказчика.

Программируемые логические матрицы представляют собой готовые изделия, конструктивно содержащие две вентиляльные матрицы, например, матрицу элементов типа И и матрицу элементов типа ИЛИ, все узловые точки которых соединены диодами. В ходе программирования по заданию заказчика каждое такое соединение либо размыкается, либо остается без изменения. Итогом коммутации является нужная логическая структура типа И / ИЛИ. ПЛМ позволяет получить заданную комбинаторную булеву логику.

Программируемые логические матрицы выпускаются как стандартные изделия, не требующие дополнительных затрат времени и средств на специализацию. В этом случае длительность цикла разработки БИС сводится к минимуму. Программируемые логические матрицы удобны для малосерийных системных изделий.

Проектирование схем на стандартных элементах заключается в подборе оптимизированных функциональных блоков или стандартных ячеек, их размещении и коммутации. Процесс изготовления схем на стандартных элементах идентичен процессу изготовления схем, проектируемых вручную. Стандартные ячейки выполняются в виде аналоговых соответствующих устройств малой и средней степени интеграции. Это могут быть арифметическо-логические устройства, регистры и т. п., называемые иногда макроэлементами.

Можно интегрировать стандартный БИС. В этом случае на одном кристалле можно разместить систему на основе микропроцессора. Такой метод получил название *метода процессорного ядра* или *метода суперинтеграции*.

Второе поколение ПЛИС более сложно с точки зрения архитектуры и технологии изготовления.

Микросхемы, относящиеся к заказным, могут разрабатываться также по ячеечному принципу. Ячеечные микросхемы представляют собой своеобразный гибрид из стандартных ячеек, который конструируется одновременно с разработкой микросхемы.

С точки зрения экономики производства этот тип БМК занимает промежуточное положение между полностью заказными схемами и схемами на стандартных ячейках. Иногда отождествляется проектирование микросхем на стандартных ячейках и ячеечное проектирование, а соответствующие микросхемы называют *ячеечными*.

13.2. Программируемые логические интегральные схемы

Программируемые логические интегральные схемы (ПЛИС) родились в жесткой конкурентной борьбе, когда в микроэлектронике существуют две противоречивые тенденции. С одной стороны, нужно сократить жизненный цикл изделия микроэлектроники, чтобы удовлетворить быстро меняющиеся требования заказчика. С другой стороны, надо постоянно повышать требования к сложности, быстродействию, потребляемой мощности, надежности и стоимости изделий микроэлектроники. Поэтому необходимо сокращать цикл проектирования с тем, чтобы на рынке появились новые изделия, соответствующие требованиям времени и прогресса.

Масочные вентиляльные матрицы теряют свои преимущества при переходе на субмикронную технологию. Дело в том, что с масштабированием размеров свойства транзисторов улучшаются, а свойства межсоединений ухудшаются. Это происходит потому, что с уменьшением топологической нормы уменьшается сечение проводника и соответственно увеличивается его сопротивление. В результате задержка в линиях межсоединений доминирует над задержкой сигнала в вентиле. В межсоединениях нужно вводить дополнительный масочный слой для увеличения трассировочных ресурсов. Каждый дополнительный слой при 0,35 топологической норме обходится не менее 15 тысяч долларов. Это обстоятельство осложняет производство БМК.

Выход найден в программируемых логических микросхемах. Программируемая логика обладает возможностью внутрисистемной репрограммируемости, что в сочетании с высоким быстродействием и уровнем интеграции, а также с малой потребляемой мощностью и низкой стоимостью позволяет ПЛИСам найти широкое применение.

Наибольшее распространение получают МОП ПЛИС, имеющие более 100 тысяч вентилялей, а также встроенную память и ядро системы-на-чипе: процессор, контроллер, сигнальный процессор. Эти схемы можно быстро переконфигурировать.

Логическая емкость определяется числом эквивалентных вентилялей типа 2И — НЕ.

По способам коммутации элементов логических матриц различают несколько классов ПЛИС.

Программируемые логические матрицы (FPLA. Field Programmable Array) состоят из базовых ячеек типа И и ИЛИ. Такая архитектура недостаточно полно использует программируемую матрицу ИЛИ.

Программируемая матричная логика (PAL, Programmable Array Logic) содержит программируемую матрицу И и фиксированную матрицу ИЛИ. К этому типу относится большинство ПЛИС небольшой степени интеграции.

Программируемые коммутируемые матрицы (CPLD, Complex Programmable Logic Divieces) содержат матричные логические блоки, объединенные коммутационной матрицей. Это ИС высокой степени интеграции с программируемой матрицей И и фиксированной матрицей ИЛИ.

Программируемые вентиляемые матрицы (FPGA, Field Programmable Gate Array) состоят из логических блоков (ЛБ) и коммутирующих соединений блока ввода/вывода информации.

Конфигурируемый логический блок (CLB, Configured Logic Block) представляет собой комбинацию таких элементов, как ОЗУ, регистры, функциональные генераторы (рис. 13.5).

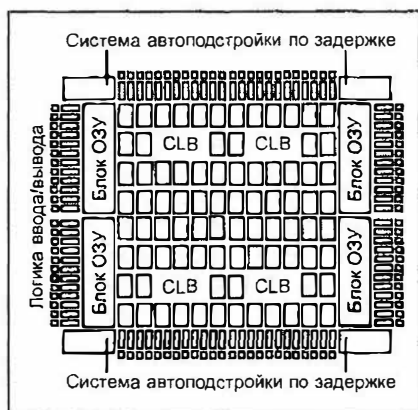


Рис. 13.5. Вариант базового блока ПЛИС типа FPGA: CLB — конфигурируемый логический блок

Для программирования логического устройства используются СОЗУ, ОЗУ базового блока, размещенные вблизи логических ячеек, выполняют задачи конфигурации системы, функции управления.

Такие ПЛИС имеют более 10^6 вентилях, что позволяет их легко реконфигурировать, встраивать нужные функции. ПЛИС выполняются по КМОП-технологии по 0,35 или 0,25 мкм топологическим нормам с пятислойной металлизацией. Напряжение питания логических устройств составляет 2,5 В, для контактных площадок ввода/вывода требуется источник питания 3,3 В.

Фирмы, выпускающие ПЛИС, разрабатывают свою архитектуру, основными элементами которой являются конфигурируемый логический блок, блоки ОЗУ, блоки ввода/вывода информации и т. д.

Стремительное развитие архитектурных решений ПЛИС вызвало к жизни новые их разновидности, связанные с расширением их функциональных возможностей. В ПЛИС встраиваются специализированные блоки программируемых логических ядер. Они представляют собой программируемые устройства, содержащие ячейки специализированных ИС и блоки программируемой логики. Такое ядро называется *soft-ядром*. Это виртуальный компонент, не имеющий строгих геометрических границ. Для выполнения *soft-ядром* логических функций не существует физических атрибутов.

Встроенное программируемое аппаратное ядро (*hard-ядро*) позволяет реализовать различные функции по требованию заказчика. Hard-ядро является виртуальным компонентом с заданными параметрами, которые описаны на физическом уровне.

По уровню характеристик soft-ядро уступает аппаратным hard-ядрам.

Развитие ПЛИС позволяет сократить цикл проектирования современных цифровых систем, сохранить гибкость конструкции и использовать новейшие технологические решения. Эти вопросы решаются благодаря способности ПЛИС вносить изменения в конструкцию системы на любом этапе процесса проектирования. При этом ПЛИС отвечает сочетанию таких характеристик, как быстродействие, минимальная потребляемая мощность, уровень интеграции и стоимость.

Контрольные вопросы

1. Что такое БМК?
2. Что такое заказная ИС?
3. Что такое полужаказная ИС?
4. Какова структура БМК?
5. Что такое программируемая логическая ИС?

Рекомендуемая литература

1. Домрачев В. Г., Мальцев П. П., Новаченко И. В., Пономарев С. Н. Базовые матричные кристаллы и матричные БИС. — М.: Энергоатомиздат, 1992.

14. Интегральные схемы СВЧ-диапазона

14.1. Общие положения

Техника СВЧ широко используется в быту, научных исследованиях, системах связи и обработки информации, особенно специального применения. Широкое использование СВЧ-устройств связано, прежде всего, с возможностью концентрации высокочастотного излучения в узкий луч. В перспективе это позволит создавать экономичные системы связи, радиолокационные станции обнаружения и сопровождения цели. Большая информативная емкость СВЧ-диапазона позволяет уплотнить число каналов связи, организовать многоканальную передачу телевизионных каналов.

Многие из этих перечисленных проблем могут быть эффективно решены методами и технологией микроэлектроники.

В последнее время существенно повысился интерес к твердотельной электронике СВЧ вообще и к монолитным ИМС СВЧ-диапазона в частности. Этот интерес вызван потребностью в развитии электронного оборудования спутникового вещания и связи, бортовой электроники самолетов и ракет (как гражданского, так и специального применения), радиолинейных линий, связанного оборудования, подвижных объектов и т. д. Одним из серьезных стимулов для развития монолитной микроволновой микроэлектроники является повышенный интерес к развитию техники фазированных антенных решеток (ФАР), для создания которых необходимо большое количество (тысячи и десятки тысяч) однотипных дешевых приемопередающих модулей.

Основным материалом монолитных микроволновых интегральных микросхем (M^3IS) в настоящее время является арсенид галлия. Однако поскольку технология арсенида галлия и транзисторов на его основе не была в достаточной степени отработанной, первые разработки в области твердотельной интегральной электроники СВЧ представляли собой гибридные интегральные схемы, толсто пленочные или тонко пленочные, с кремниевыми биполярными транзисторами, чаще всего в бескорпусном исполнении.

Толсто пленочная технология в технике СВЧ используется только в цифровой технике для обеспечения соединений на плате между установленными интегральными схемами (БИС или СБИС). Одновременно с этим нельзя не отметить негативного влияния паразитных параметров таких соединений на быстродействие.

В СВЧ-технике толсто пленочная технология может быть использована для относительно низкочастотных и узкополосных (до 20%) устройств.

Большинство СВЧ ИС являются широкополосными (с отношением граничных частот в диапазоне 1:2 или 67% и более), поэтому в гибридном варианте более предпочтительными оказываются тонко пленочные СВЧ-устройства.

Толщина тонких пленок определяется тремя-пятью толщинами скин-слоев на нижней частоте диапазона (толщина скин-слоя определяется по спаданию плотности тока от поверхности вглубь объема в 2,7 раза). В результате такие пленки имеют толщину до 7,5 мкм.

Если сравнивать тонкопленочную гибридную и монокристалльную технологии, то последняя более предпочтительна. Однако монокристалльная технология является более трудоемкой.

Одним из ожидаемых преимуществ должна являться относительно низкая стоимость монокристалльных ИС по сравнению с гибридными. Этот фактор определяется значительными удельными объектами производства, с одной стороны, и значительным удельным весом трудоемких и дорогостоящих "индивидуальных" сборочных операций в гибридной технологии, с другой стороны.

Серьезной проблемой является и воспроизводимость результатов. Так, отклонение в величине емкостей оказывает существенное влияние на величину коэффициента усиления. В гибридной технологии эта проблема может решаться за счет достаточно трудоемких процессов подстройки. В монокристалльной технологии жесткий контроль технологических процессов позволяет, например, довести отклонения емкостей от номинала до величины, не превышающей 3%.

Проблема воспроизводимости и повторяемости результатов является весьма серьезной. Эту проблему можно решить только тесно увязывая схемотехнические решения с конструкцией и технологией.

14.2. Элементная база электроники СВЧ

На первых этапах развития радиотехнических систем связи широко использовались электровакуумные приборы: лампы, магнетроны, клистроны и т. д. Они позволили освоить СВЧ-диапазон, однако не всегда удовлетворяли по таким параметрам, как миниатюрность, надежность, оптимальное энергопотребление. Взгляды радиоинженеров все чаще обращались к микроэлектронике. Именно микроэлектронные устройства позволяли получить высокую надежность при малом энергопотреблении, малые габариты и низкую цену обработки одного бита информации.

Известно, что в любой электронной аппаратуре различают пассивные и активные элементы, линии межсоединений. В традиционной микроэлектронике линии межсоединения выполняются в виде алюминиевых полосок, и проблем их создания в интегральных схемах не возникает вплоть до высокой степени интеграции. Иное дело — межсоединения в СВЧ-диапазоне. В микроэлектронной аппаратуре различают иерархические уровни соединений.

Нулевой конструктивно-технологический уровень составляют *межэлементные соединения*. Они связывают в схеме с определенными функциями пассивные и активные элементы.

Первый уровень составляют соединения в гибридных микросхемах СВЧ-диапазона, связывающие на плате бескорпусные микросхемы, навесные активные и пассивные электроадиоэлементы.

Во втором уровню межсоединений относятся проводники, соединяющие гибридные микросхемы, корпусированные микросхемы, дискретные электрорадиоэлементы в ячейки или микросборки. В свою очередь, как правило, межсоединения этих уровней представляют собой микрополосковые перемычки или полосково-коаксиальные переходы.

Ячейки, или микросборки, а также электрорадиоэлементы коммутируются в блоки СВЧ помощью межсоединений третьего уровня, выполненных в виде микрополосковых перемычек или полужестких кабелей.

На следующих уровнях межсоединений используют СВЧ-кабели, не представляющие интереса для микроэлектроники.

Микрополосные линии (МПЛ) весьма интересны с точки зрения физической электроники. МПЛ представляет собой проводник ленточного типа шириной W , прямоугольного сечения, расположенный на подложке толщиной h с высокой диэлектрической проницаемостью ϵ . Обратная сторона подложки металлизирована и заземлена (рис. 14.1, а). Микрополосковая линия такой конструкции обладает волновым сопротивлением, зависящим от соотношения W/h и величины ϵ , а также от коэффициента потерь, дисперсии и предельной передаваемой мощности. При конструировании устройств СВЧ-диапазона появляется необходимость изменения геометрических размеров МПЛ, что получило название неоднородности МПЛ.

К пассивным элементам СВЧ-диапазона относят резисторы, конденсаторы и индуктивности.

Эффект электрического сопротивления прохождения тока в СВЧ-диапазоне возникает в неоднородности микрополосковых линий в емкостях, образующихся в воздушных промежутках, диэлектрических материалах, окисных пленках между кристаллами (рис. 14.1, б).

Конденсаторы микросхем СВЧ-диапазона также изготавливаются на основе МПЛ. Малые номиналы (несколько пФ) можно получить на разрывах МПЛ (рис. 14.1, в), а большие реализуются в виде конструкции типа гребенчатого конденсатора. Для получения конденсаторов емкостью более 10 пФ используют многослойные структуры.

Индуктивность как элемент СВЧ-схем может быть реализована в виде прямоугольного отрезка МПЛ со скачком по ширине или в форме круглой и квадратной спирали.

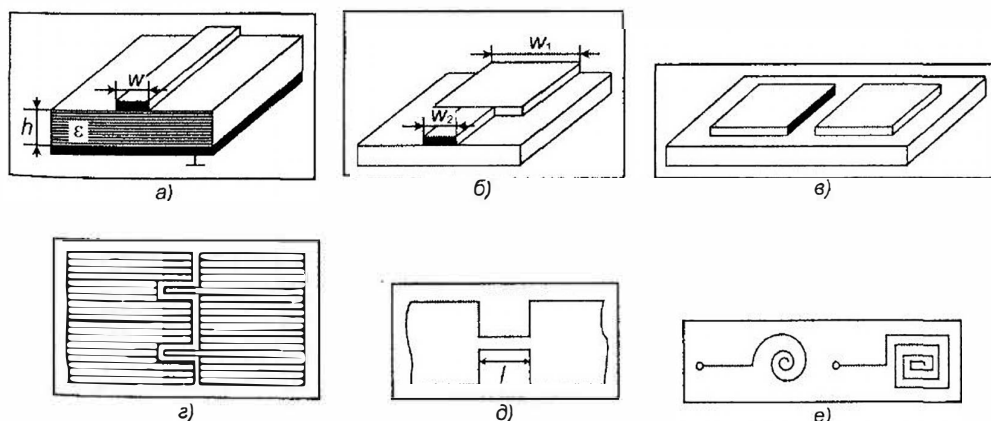


Рис. 14.1. Элементная база СВЧ-микросхем. Микрополосковая линия (а) и пассивные элементы на ее основе: резистор (б), конденсатор (в, г), индуктивность (д, е)

К пассивным элементам можно условно отнести диоды СВЧ-диапазона, которые не генерируют колебаний. Существуют конструкции диодов, обладающие S- (рис. 14.2, а) или N-образными (рис. 14.2, б) вольтамперными характеристиками. Такие диоды на определенных участках ВАХ имеют отрицательное дифференциальное сопротивление,

и, соответственно, способны генерировать электромагнитные колебания. Эти диоды и триодные структуры относятся к активным элементам СВЧ-микросхем.

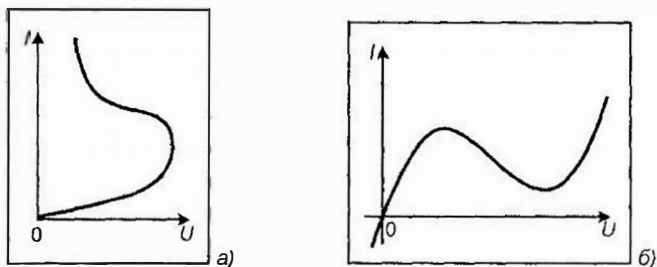
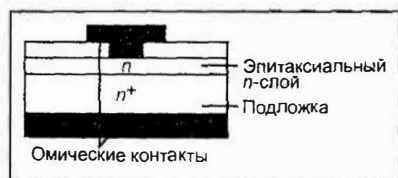


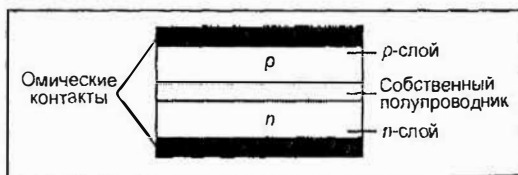
Рис. 14.2. S-образная ВАХ (лавинно-пролетный диод) (а) и N-образная ВАХ (туннельный диод) (б)

Диод с барьером Шоттки представляет собой выпрямляющий контакт "металл—полупроводник" (рис. 14.3, а). Он работает на основных носителях заряда, при этом неосновные не накапливаются. Время восстановления обратного сопротивления составляет $\sim 10^{-8}$ с, что позволяет использовать такие подложки до частот 300 ГГц.

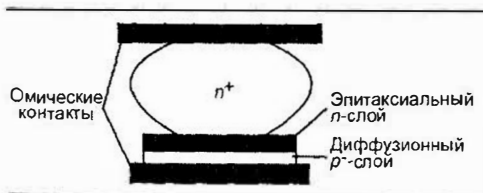
Диод р—i—n-структуры формируется на основе обедненного i-слоя между р- и n-областями. Он обладает высоким пробивным напряжением (рис. 14.3, б) и способен работать при напряжениях > 1 кВ и импульсной мощности ~ 10 кВт.



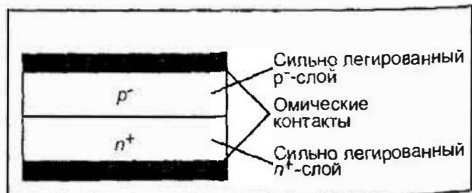
а)



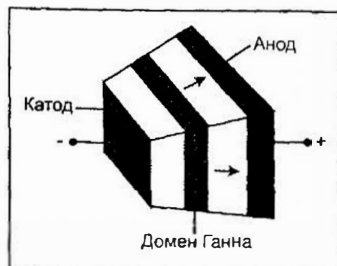
б)



в)



г)



д)

Рис. 14.3. Диоды и триоды СВЧ-диапазона:
а — диод с барьером Шоттки; б — р—i—n-диод;
в — лавинно-пролетный диод; г — туннельный диод;
д — диод Ганна

Лавинно-пролетный диод работает на основе лавинного пробоя р—п-перехода при высоких обратных напряжениях (рис. 14.3, в). На его основе можно создать достаточно мощные диоды, работающие в гигагерцевом диапазоне частот.

Туннельные диоды представляют собой р—п-переходы с туннельным эффектом (рис. 14.3, г). Они обладают широкополосностью, низким уровнем шума, высокой температурной стойкостью.

Диод Ганна в основе своей конструкции имеет невыпрямляющий контакт "металл—полупроводник" (рис. 14.3, д). Он работает в гигагерцевом диапазоне частот при значительных мощностях импульсов.

Однако наибольший интерес, как активные элементы, представляют полевые и биполярные транзисторы СВЧ-диапазона.

Главное их отличие от традиционных транзисторных структур микроэлектроники — материал. Если в традиционной микроэлектронике все структуры выполняются на кремниевых подложках того или иного типа проводимости, то в микроэлектронике СВЧ используются полупроводниковые соединения типа A^3B^5 или A^2B^6 .

14.3. Интегральные транзисторы СВЧ-диапазона

Основной тип СВЧ-транзисторов — это полевые транзисторы с барьером Шоттки в качестве затвора, выполненные по арсенид-галлиевой технологии. На рис. 14.4 представлена структура полевого транзистора на арсениде галлия с каналом n-типа. Между затвором и слоем канала введен слой нелегированного GaAlAs толщиной 300 Å, назначение которого заключается в увеличении пробивного напряжения "затвор—сток".

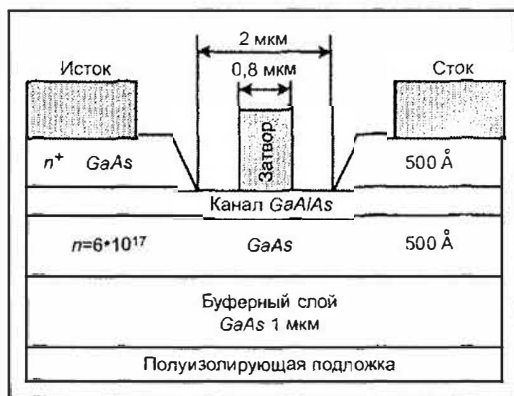


Рис. 14.4. Структура транзистора с барьером Шоттки

Разнообразные типы полевых транзисторов можно классифицировать по механизму переноса носителей. При дрейфе электронов от истока к стоку они испытывают большое количество соударений. Напряженность поля в канале полевых транзисторов обычно превышает 10 кВ/см, а среднее значение энергии электронов в установившемся режиме — 0,3 эВ. Частота соударений при этом намного превышает 10^{13} с^{-1} . Поскольку пролетное время составляет 10^{-12} с , носители испытывают за время пролета десятки или сотни соударений. За время 10^{-13} с , проходящее между двумя соударениями, носители проходят расстояние, не превышающее 400 Å.

При конструировании полевых транзисторов приходилось сталкиваться с проблемой падения подвижности при повышении концентрации носителей в канале, необходимой при малой длине канала. Поскольку рост концентрации носителей связан с повышением степени легирования, то возрастание концентрации доноров увеличивает вероятность столкновения носителей с ионами доноров и снижает подвижность. Использование гетеропереходов позволило разрешить это противоречие: двумерный электронный газ обеспечивает возможность получения слоя с повышенной концентрацией носителей без увеличения концентрации доноров и свободных электронов, что дает возможность получать высокие концентрации.

Изменение степени легирования в данных транзисторных структурах нашло отражение в их названии — "модуляционно легированный" или "селективно легированный".

Возможны и другие варианты транзисторной структуры с высокой подвижностью электронов (HEMT, High Electron Mobility Transistor), например, с каналом в слое на основе узкозонного полупроводника GaAs и слоем "поставщиком электронов" — широкозонным полупроводником AlInAs.

Подвижность в канале GaInAs при 300 К достигает $10\,000\text{ см}^2/(\text{В}\cdot\text{с})$. Обеспечивается высокая плотность заряда в слое двумерного электронного газа $(3\text{--}4,5)\times 10^{12}\text{ см}^{-2}$.

При этом необходимо отметить, что все эти качества в транзисторе с высокой подвижностью электронов в значительно большей степени проявляются при пониженных температурах.

Одним из серьезных препятствий на пути реализации возможностей транзисторов с высокой подвижностью электронов является наличие глубоких ловушек для электронов при высоком уровне содержания алюминия в AlGaAs. Для получения слоя с двумерным электронным газом содержание алюминия в AlGaAs должно превышать $x > 0,2$, но при этих значениях глубокие ловушки приводят к срыву стоковых ВАХ, повышению уровня генерационно-рекомбинационных шумов и даже к появлению эффекта фоточувствительности.

В качестве меры противодействия предлагается формировать слой двумерного электронного газа на границе раздела AlGaAs/InGaAs. Другими словами, эта модификация транзистора отличается введением между слоем AlGaAs (30—40 Å) и нелегированным GaAs (1 мкм) слоя InGaAs толщиной в 200 Å. Эта модификация получила название *псевдоморфного транзистора с высокой подвижностью электронов*. Действие слоя InGaAs объясняется тем, что этот материал имеет меньшую по сравнению с арсенидом галлия ширину запрещенной зоны. Именно это дает возможность успешно использовать в паре с ним AlGaAs с низким содержанием алюминия ($x = 0,15$).

Ряд транзисторов СВЧ-диапазона был рассмотрен ранее.

Отметим некоторые особенности изготовления полевых транзисторов.

- В технологическом процессе используется молекулярно-лучевая эпитаксия. В качестве подложек применяются пластины арсенида галлия.
- Вольфрамовые электроды затворов изготавливаются с помощью электронно-лучевой литографии и реактивного ионного травления. Возможно получение электродов шириной менее 0,25 мкм и высотой более 1 мкм.
- Активные области стоков легируются кремнием с помощью ионной имплантации.
- Совмещение элементов прибора с двух сторон подложки (затвор и исток) проводится в инфракрасном свете.

- Возможны два варианта транзистора этой конструкции: на однородной активной области канала и с гетеропереходом и слоем двумерного электронного газа (по типу HEMT).

Особо следует отметить двухзатворный вариант полевых транзисторов. Эта конструкция особенно удобна для использования в схемах автоматической регулировки усиления или в каскадах с управляемым усилением, как это требуется, например, в модулях активной фазированной антенной решетки.

14.4. Монолитные арсенид-галлиевые ИС

Монолитные арсенид-галлиевые интегральные схемы (МИС) перекрывают диапазон частот от 1 до 100 ГГц. Это позволяет их широко использовать в радиолокационных станциях, спутниковых системах навигации, средствах связи и т. п.

рост спроса на GaAs МИС стимулируется стремительно развивающимся рынком беспроводных систем связи.

В арсенид-галлиевых интегральных схемах в основном используются следующие транзисторные структуры:

- полевой транзистор с барьером Шоттки (MESFET);
- транзисторы на горячих электронах (HEMT);
- биполярные гетеротранзисторы (HBT).

На рис. 14.5 приведена частотная зависимость выходной мощности различных типов арсенид-галлиевых транзисторов.

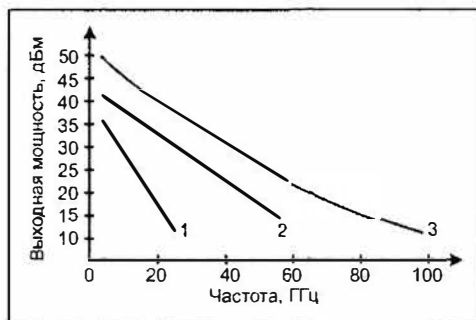


Рис. 14.5. Частотная зависимость выходной мощности:
1 — GaAs MESFET;
2 — AlGaAs/GaAs HBT;
3 — GaAs HEMT

Выбор типа транзисторов для арсенид-галлиевых ИС зависит от фундаментальных механизмов работы, от степени совершенства технологии. Основным активным элементом современных GaAs ИС являются MESFET-структуры. Однако высокочастотные характеристики этого типа транзисторов ограничены подвижностью электронов и временем пролета канала. Уменьшая длину затвора можно увеличить быстродействие MESFET-транзисторов, что, однако, не очень эффективно.

Создаются сложные структуры, в которых стараются увеличить подвижность электронов. Например, в HEMT-структуре создаются гетеропереходы с квантовыми колодцами, в которых формируется двумерный электронный газ, в котором существенно увеличивается подвижность электронов.

Транзисторные структуры типа HBT по конструкции во многом похожи на кремниевые биполярные транзисторы, а по принципу действия аналогичны транзисторам на горячих электронах. Активной областью HBT транзисторных структур служит гетеропереход типов GaAs-GaAlAs и GaAs-InGaAs. В этой области носители движутся в нелегированном канале без рассеяния на примесных ионах. Это приводит к увеличению быстродействия.

Успехи в области разработки конструкции и создания технологии арсенид-галлиевых транзисторных структур позволяет надеяться на расширение их потребности в ВЧ-схемах, малошумящих усилителях и усилителях мощности.

Если первые промышленные арсенид-галлиевые интегральные схемы предназначались только для военных систем связи, то в настоящее время ожидается их широкое применение в системах гражданской коммуникации следующего поколения.

Контрольные вопросы

1. Что собой представляет элементная база ИС СВЧ-диапазона?
2. Какими характеристиками должен обладать генерирующий диод СВЧ-диапазона?
3. Какую конструкцию СВЧ-транзистора вы знаете?

Рекомендуемая литература

1. Филатов И. Н., Бакрунов О. А., Панасенко П. В. Микроэлектронные СВЧ-устройства. Микроэлектроника: Учебное пособие для вузов под ред. Л. А. Коледова, кн. 7. — М.: Высшая школа, 1987.
2. Шур М. Современные приборы на основе арсенида галлия. — М.: Мир, 1991.
3. Щука А. А. Микроэлектроника СВЧ-диапазона волн. — М.: МИРЭА, 1998.

15. От микро- к наноэлектронике

15.1. Барьеры на пути перехода от микро- к наноэлектронике

В основе физики полупроводников и всей остальной микроэлектроники лежит фундаментальное понятие — энергетический барьер. Это, прежде всего, традиционный р—п-переход, а также граница полупроводника либо с другим веществом, либо с окружающей средой (воздух, вакуум, твердое тело). Основной тенденцией развития микроэлектроники является увеличение степени интеграции. В соответствии с пресловутым законом Мура число элементов, входящих в состав одной только микросхемы, удваивается каждые полтора (два?) года. Однако на пути продолжения этой тенденции встают барьеры: технологический, физический, энергетический... Так, фотолитографическая технология, лежащая в основе технологии производства интегральных схем, достигла своего физического совершенства. Сейчас на подходе рентгеновская и лазерная литография. Лазерная литография позволяет получить разрешение элементов схемы лучше 10 нм. Процесс печати схемы занимает всего 250 нс. На преодоление технологического барьера направлена американская программа "The National Technology Roadmap for Semiconductors". В соответствии с этой программой к 2015 году будут проектироваться транзисторы с шириной затвора 20 нм при технологической норме 30 нм. Будут увеличены площади кристаллов интегральных схем до 10 см². Это позволит разместить на кристалле 10⁹ вентилях. При этом рабочая частота составит 30 Гц—30 ГГц.

Однако до сих пор множество вопросов, касающихся преодоления физических барьеров, остаются нерешенными. Среди них известная проблема межсоединений: в современных интегральных схемах лишь 10% площади занято транзисторами, а 90% межсоединениями. Масштабирование элементов транзисторных структур с целью перехода от микро- к нанометровым масштабам является весьма деликатным процессом. Создать транзисторные структуры в нанометровом масштабе сегодня уже под силу серьезным фирмам. Однако так и не решены вопросы их интеграции, вопросы создания групповой технологии производства интегральных схем в нанометровом масштабе, создание новых или традиционных компонентов интегральных схем.

В традиционных схемах микроэлектронной схемотехники устройства всегда имеют "вход" и "выход", которые пространственно разделены и локализованы в электрической схеме, а также в определенных контактах интегральной схемы. Все связи в интегральных схемах реализованы с помощью гальванических или емкостных связей. Реализация таких связей осуществляется путем изменения типа проводимости исходной подложки и создания различных энергетических барьеров на пути потока носителей (электроны, дырки).

Информация обрабатывается и хранится в виде отдельных битов (логические нуль и единица), которые физически реализуются в виде тока, напряжения или заряда в определенной точке интегральной схемы.

Совсем иная физическая картина рисуется при рассмотрении наноэлектронных устройств. Уменьшение размеров на несколько порядков практически меняет физические основы

работы наноэлементов. Так, в наноэлементе используются уже не электроны, как частицы, переносящие электрический заряд, а их волновые функции. Процессы дрейфа и диффузии, характерные для микроэлектронных элементов, и вовсе отсутствуют в наноэлектронных элементах. В основе наноэлементов лежат полевые связи, сформированные потенциальные барьеры. "Вход" и "выход" элемента локализованы не в пространстве, а во времени. В наноэлектронных структурах определение "входа" или "выхода" возможно только в определенные промежутки времени, когда существует определенный порог внешних воздействий, соответствующий "входу" или "выходу". Этот промежуток времени получил название *рефракторного периода* и с его помощью обеспечивается распространение сигнала в определенном направлении. Переход от одного устойчивого состояния наноструктуры в другое происходит через возбужденные неустойчивые состояния.

Как правило, наноэлектронный элемент состоит из набора квантовых ям и потенциальных барьеров. Энергетический спектр таких элементов зависит от размеров, а добавление лишь одного электрона существенно меняет энергетическую диаграмму. В микроэлектронике функциональный логический элемент представляет собой совокупность структурных компонентов — резисторов, конденсаторов, диодов и транзисторов. В наноэлектронике структурированные компоненты обладают свойством многофункциональности, способны выполнять сложные динамические функции.

В качестве материалов изделий микроэлектроники используются легированные полупроводники. В наноэлектронике применяются гетероструктуры, наноструктурированные материалы, кластеры, органические материалы. Технология формирования наноструктур основана на процессах направленного роста, методах, связанных со сканирующим туннельным и атомным силовым микроскопами. Если плотность размещения активных элементов в интегральных схемах достигает 10^8 см^{-2} , то в устройствах наноэлектроники она может достигать 10^9 — 10^{10} элементов на квадратный сантиметр. Наноэлементы дают возможность получать излучатель и приемник в диапазонах частот 10^9 — 10^{13} Гц.

15.2. Начала наноэлектроники

Под *наноэлектроникой* понимают направление электроники, в котором изучаются физические явления и процессы взаимодействия электронов с электромагнитными полями, а также разработка нанотехнологии создания приборов и устройств, в которых данное взаимодействие используется для передачи, обработки и хранения информации.

Под нанотехнологией будем понимать совокупность способов и приемов создания элементов и приборов нанометровых размеров, в том числе из отдельных молекул и атомов.

Исключительно малая инерционность электронов позволяет эффективно использовать взаимодействие электронов с микрополями внутри атома, молекулы или кристаллические решетки для создания приборов и устройств нового поколения, отличающихся высокой производительностью, ничтожным потреблением энергии, сверхминиатюрными размерами.

Наноэлектроника является логическим развитием микроэлектроники. Твердотельные информационные приборы уменьшились от микро- (10^{-6} м) до нанометрового (10^{-9} м) размера.

По мере приближения характерного размера твердотельной структуры электронного прибора к нанометровой области, соизмеримой с размерами атомов, проявляются квантовые свойства электронов. Если в микроэлектронных приборах поведение электрона определяется поведением элементарной частицы, имеющей массу и заряд, то в наноэлектронных приборах поведение электрона определяется его волновыми свойствами.

Де-бройлевская длина волны электрона равна

$$\lambda_e = \frac{h}{p}$$

или

$$\lambda_e = 2\pi\hbar / m_e \cdot V,$$

где $p = m_e \cdot V$ — импульс электрона, m_e — масса электрона, V — скорость его движения, \hbar — постоянная Планка.

Импульс электрона p связан с ускоряющим напряжением соотношением

$$p = \sqrt{2m_e E},$$

где E — энергия электрона. Для определения длины волны электрона удобно пользоваться практической формулой

$$\lambda \cong \sqrt{\frac{150}{V}},$$

где V — ускоряющее напряжение, В; λ — длина волны, Å.

Движение электрона как волны описывается уравнением Шредингера для нерелятивистских электронов и уравнением Дирака для релятивистских. В наноэлектронике обычно используются нерелятивистские электроны. Для электрона массой m_e , движущегося под действием силы, порождаемой потенциалом $V(x, y, z, t)$, уравнение Шредингера имеет вид:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m_e} \nabla^2 \psi + V(r) \equiv H\psi \quad (15.1)$$

или

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(r) \right] \psi = E\psi, \quad (15.2)$$

где r — функция положения электрона в пространстве, $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ — оператор

Лапласа, H — оператор Гамильтона, E — полная энергия электрона, ψ — волновая функция (собственная функция уравнения). Смысл функции ψ состоит в том, что величина $\psi^* \psi dx dy dz$ является вероятностью нахождения электрона в объеме $dx dy dz$, а ψ^* — комплексно сопряженная величина ψ .

Движение электрона в свободном пространстве $V(r) = C$ описывается уравнением

$$-\frac{\hbar^2}{2m} \nabla^2 \psi = E\psi.$$

Решение этого уравнения пишется в виде

$$\psi \sim e^{ik_{ij}},$$

где \vec{k} — волновой вектор, определяемый как $|\vec{k}| = k = \frac{2\pi}{\lambda}$.

На рис. 15.1 приведена уникальная фотография, экспериментально подтверждающая наличие де-бройлеровской волны. В лаборатории фирмы IBM с помощью туннельного микроскопа удалось рассадить атомы кобальта вдоль периметра кольца диаметром 20 нм на поверхности меди.

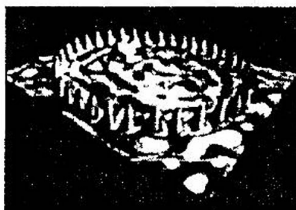


Рис. 15.1. Атомы кобальта на медной подложке

Если внутри кольца поместить еще один атом кобальта, то возникнет его изображение, инверсное относительно центра кольца. Эффект инверсности объясняется интерференцией де-бройлеровской волны атома кобальта на периодической решетке других атомов. Возникновение или отсутствие изображения зависит от положения вновь имплантированного атома. Если де-бройлеровские волны складываются в фазе в процессе конструктивной интерференции, то изображение появляется. При деструктивной интерференции оно исчезает.

Изображение интерференции удалось получить именно на туннельном микроскопе. Эта картинка — одно из доказательств волновой природы отдельного атома или электронов внешних его орбит.

Рассмотрим квантовые эффекты, лежащие в основе функционирования элементов, предназначенные для кодирования, обработки и хранения информационных сигналов. Распространение волновой функции электрона в твердотельных наноразмерных структурах связано с эффектом квантового ограничения, интерференцией, дифракцией волновой функции электрона, туннелированием через потенциальные барьеры.

Рассмотрим квантовые эффекты, наноразмерные структуры и явления взаимодействия между ними.

В приборах квантовой нанoeлектроники роль носителя информации отводится волновой функции электрона. Такой подход в обработке, предоставлении и хранении информации сводится к контролируемой передислокации волновой функции в наноструктуре. Фундаментальным явлением в квантовых структурах является связанность дискретной составляющей системы с дискретностью квантово-размерного энергетического спектра.

15.3. Квантовые ограничения

Переход системы из одного дискретно-квантового состояния в другое представляет собой предельно быстрый безынерционный квантово-механический процесс.

Электроны проводимости в твердом теле являются свободными и могут беспрепятственно распространяться в любом направлении. Волна Де-Бройля, соответствующая свободным электронам, обладает теми же свойствами. Волновые функции этих электронов необходимо искать в виде

$$\psi \sim e^{ikx}$$



рис. 15.2. Увеличение подвижности носителей соответственно быстродействию электронных приборов при уменьшении размерности

Эти функции должны отвечать следующим условиям:

- Они должны быть нормированы так, чтобы интеграл от плотности вероятности по всей рассматриваемой области пространства был равен единице, т. е. $\int \psi^* \cdot \psi d\vec{r} = 1$.

Это возможно, если $\psi = \frac{1}{\sqrt{V}} e^{ikr}$, где V — объем тела, в котором движется электрон.

- Волновая функция должна удовлетворять надлежащим граничным условиям на поверхности твердого тела, а именно обращалась бы в нуль на его поверхности. Например, движение электрона на прямолинейном отрезке L , для которого $\psi(0) = \psi(L) = 0$, описывается функцией $\psi_n = \left(\frac{2}{L}\right)^{1/2} \sin \frac{n\pi x}{L}$, а собственно энергия: $E_n = \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 n^2$, где

$n = 1, 2, 3, \dots$ Другими словами, допустимые уровни энергии квантуются.

Дрейфовая скорость электронов в малых электрических полях определяет их подвижность μ , которая пропорциональна среднему времени свободного пробега электрона. Ясно, что подвижность носителей в трехмерной среде меньше, чем в двумерном или одномерном случае (рис. 15.2).

Рассмотрим электрон, движение которого ограничено структурой длиной L . В этом случае $\Psi(0) = \Psi(L) = 0$ и решение волнового уравнения в одномерной интерпретации будет иметь вид:

$$\psi_n = \left(\frac{2}{L}\right)^{1/2} \sin \frac{n\pi x}{L}.$$

Значение для энергии примет вид:

$$E_n = \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 n^2,$$

где $n = 1, 2, 3, \dots$

Ясно, что допустимые уровни энергии квантуются, а все прочие значения энергии не имеют места, ввиду того, что функции ψ не удовлетворяют граничным условиям.

На рис. 15.3 приведены графики первых двух волновых функций свободного электрона, находящегося в твердотельной структуре ограниченного размера или области твердого тела, ограниченной непроницаемыми потенциальными барьерами. Разрешенные значения для волнового вектора для одного из направлений задаются соотношением:

$$k = \frac{2\pi}{\lambda_n} = \frac{\pi n}{L}.$$

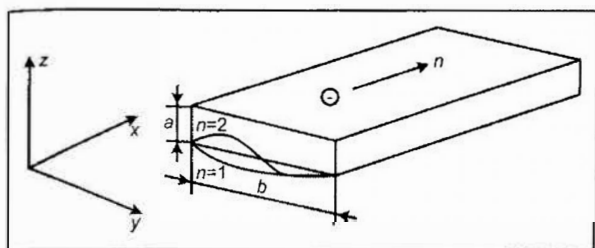


Рис. 15.3. Траектории движения электрона в квантово-ограниченной наноразмерной структуре и виды соответствующих функций

Явление наличия фиксированных значений энергии электронов в квантово-ограниченной структуре получило название *квантового ограничения*.

В соответствии с принципом неопределенности $\Delta p \cdot \Delta y \geq \hbar$, запирающая электрона с эффективной массой m^* , например, в направлении y , приводит к увеличению его импульса на величину \hbar / L . Соответственно увеличивается и их кинетическая энергия на величину

$$\Delta E = \frac{\hbar^2 k^2}{2m} = \left(\frac{\hbar^2}{2m} \right) \cdot \left(\frac{\pi^2}{L^2} \right). \quad (15.3)$$

Другими словами, квантовое ограничение сопровождается изменением энергии и формированием энергетического дискретного спектра

$$E_n = \frac{\hbar^2 \pi^2}{2mL^2} \cdot n^2,$$

$$k = \frac{2\pi}{\lambda},$$

где $n = 1, 2, 3, \dots$

Набор параллельных систем квантово-ограниченных наноразмерных структур получил название *сверхрешетки*. Другими словами, сверхрешетка представляет собой твердотельную периодическую структуру, в которой на электроны действуют потенциал кристаллической решетки и одномерный потенциал $\bar{V}(r)$, период которого d меньше длины свободного пробега электронов, но больше периода кристаллической решетки a .

Потенциал $\bar{V}(r)$ обычно создается путем чередования тонких полупроводниковых слоев, отличающихся типом легирования. На рис. 15.4 представлен пример структуры сверхрешетки на основе GaAs и AlGaAs.

Если длина свободного пробега носителя заряда существенно превосходит период потенциала $\bar{V}(r)$, то энергетический спектр вдоль оси z представляет собой систему узких полос — *минизон*. В направлении оси x носители ведут себя как свободные частицы с соответствующей эффективной массой.

Энергетический спектр характеризуется плотностью состояний $g(E)$, которая представляет собой число возможных физически неэквивалентных энергетических состояний в малом интервале энергии E , отнесенной к ширине интервала ΔE . Другими словами

$$g(E) = \lim \frac{\Delta j(E)}{\Delta E},$$

где Δj — число состояний с энергиями в интервале E и $E + \Delta E$.

На рис. 15.5 приведены плотности состояний для двумерной структуры типа сверхрешетки (СР), одномерной структуры типа квантовых нитей, проволок (КП) и точечной структуры типа квантовых точек (КТ). Парабола на первом рисунке относится к разделенной плотности состояний, соответствующей трехмерному случаю.

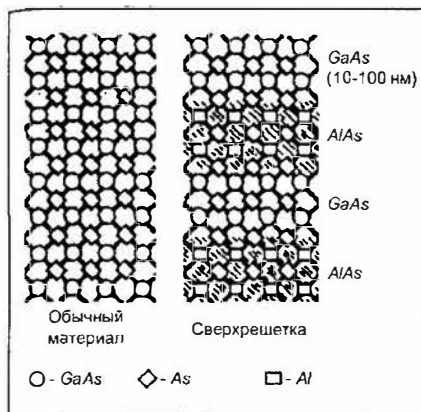


Рис. 15.4. Пример структуры сверхрешетки на основе слоев GaAs и AlGaAs. Толщина слоев GaAs и AlGaAs составляет 100 нм

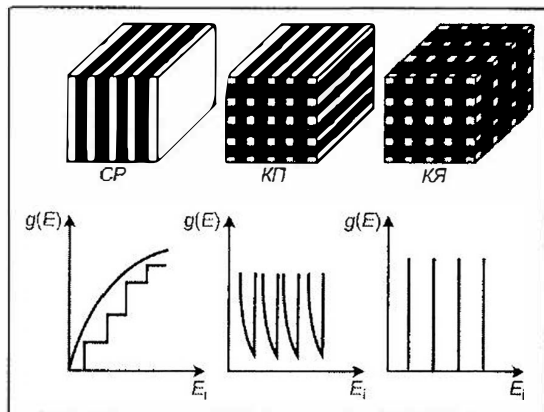


Рис. 15.5. Плотности состояний $g(E)$ для сверхрешетки (СР), квантовых проволок (КП) и квантовых точек (КТ)

15.4. Туннелирование электронов

Процесс *туннелирования* электронов заключается в квантовом переходе системы через потенциальный барьер в случае, когда энергия электронов меньше высоты потенциального барьера.

Процесс туннелирования квантовой механикой разрешен благодаря соотношению неопределенности между импульсом частицы и ее координатами. Поэтому имеется отличная от нуля вероятность обнаружения электрона в области, запрещенной классической механикой, например за потенциальным барьером. Соответственно появляется определенная вероятность прохождения частицы сквозь потенциальный барьер и возникновение процесса туннелирования через барьер. Эта вероятность тем больше, чем меньше масса частицы, уже потенциальный барьер и чем меньше разность между энергией частицы и высотой потенциального барьера. На рис. 15.6 показан процесс взаимодействия электрона с барьером в классическом и квантовом приближениях.

В многослойной структуре с ограниченным числом слоев механизм туннелирования носит резонансный характер. Другими словами, туннельно просочиться через структуру с множеством потенциальных барьеров могут только электроны с определенной энергией. Рассмотрим многослойную структуру — сверхрешетку, сформированную из слоев GaAs ($n^+ \sim 10^{17} \text{ см}^{-3}$, $E_g = 0,005 \text{ эВ}$) и $\text{Al}_{0,5}\text{Ga}_{0,5}\text{As}$ ($n^+ \sim 10^{17} \text{ см}^{-3}$), которые создают барьеры $\sim 0,5 \text{ эВ}$ (рис. 15.7).

Процесс туннелирования в многобарьерной структуре может быть описан уравнением Шредингера для периодического потенциала $U(x)$, взятого вдоль сверхрешетки.

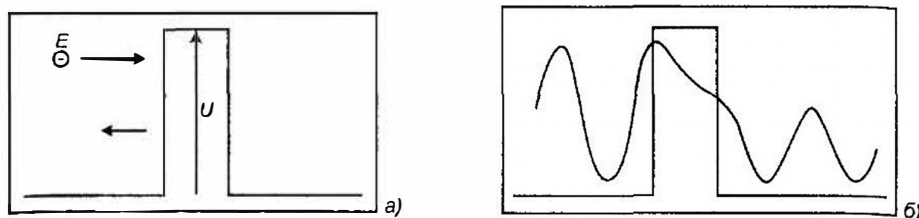


Рис. 15.6. Процесс взаимодействия электрона с энергией E с барьером в классической интерпретации (а) и взаимодействие волновой функции электрона с потенциальным барьером высотой $U > E$ (б)

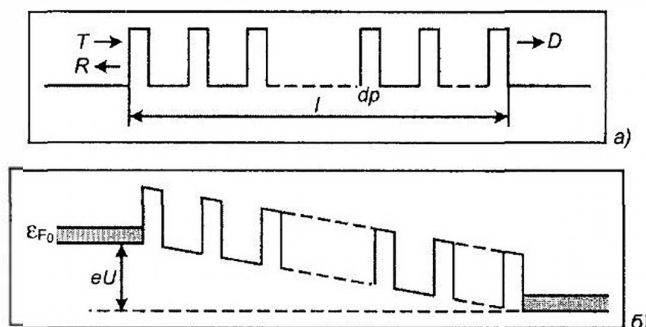


Рис. 15.7. Энергетическая диаграмма сверхрешетки длиной l в равновесии (а) и при приложении внешнего напряжения (б). Здесь T — падающий поток электронов, R — отраженный поток электронов, D — прошедший поток электронов

Общую энергию электронов E запишем как сумму продольных и поперечных составляющих:

$$E = E_1(U) + \frac{\hbar^2 k^2}{2m^*}. \quad (15.4)$$

Аналогично волновую функцию представим как $\psi = \psi_l \psi_r$. Длина свободного пробега электронов в полупроводнике составляет несколько периодов l сверхрешетки (~ 5 барьеров). На рис. 15.8 приведена зависимость рассчитанной плотности тока от приложенного напряжения при $T = 0$ К для структур с двумя (J_1) и тремя барьерами (J_2). Выражение для плотности тока имеет вид:

$$J = \frac{qm^*}{2\pi^2 \hbar^2} \int_0^{E_F} (E_F - E_x) D_e dE_x. \quad (15.5)$$

Вольтамперные характеристики носят резонансный характер и имеют максимумы. Соответственно, существуют области структуры с отрицательным дифференциальным сопротивлением (ОДС). Первый пик ВАХ приходится на 0,082 В, что соответствует первому квантово-размерному уровню в яме. Максимальный ток возникает при напряжениях, при которых резонансные энергии кратны энергии Ферми E_F . Вследствие того, что средняя длина свободного пробега электронов имеет конечное значение, электроны могут протуннелировать через несколько слоев прежде, чем они столкнутся. Большинство электронов вернутся на низкие уровни. Под влиянием приложенного напряжения эти электроны

будут туннелировать через следующие слои. Многократное некогерентное туннелирование приводит к расширению пиков и увеличению напряжений на пиках на ВАХ.

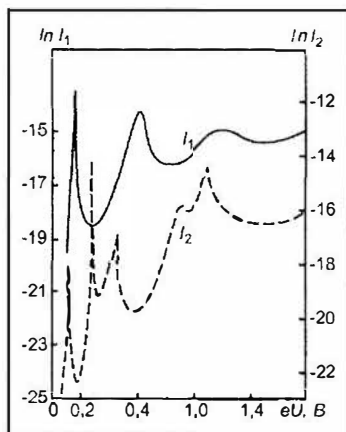


Рис. 15.8. Зависимость тока от напряжения для структуры с двумя (I_1) и тремя барьерами (I_2)

15.5. Квантовые точки

При запираании электрона в полупроводниковой среде квантовые эффекты начинают возникать при характерных размерах области запираания $\sim 100 \text{ \AA}$. Технология изготовления квантовых точек предполагает многократное использование электронолитографических шаблонов и многократное реактивное травление, которые используются в технологии приготовления интегральных схем.

После ряда этапов технологического процесса резист заменяется тонким слоем металла в тех областях поверхностей, которые были просканированы электронным лучом при повышенной интенсивности. Последующая обработка потоком реактивной плазмы позволяет стравить материал, незакрытый слоем металла. При этом металлами остаются столбики с характерным размером примерно 1000 \AA . Сверху и снизу от материала, образующего квантовую яму расположены тонкие слои диэлектрика, называемые туннельными барьерами, за которыми следуют проводящие контактные слои (рис. 15.9).

Изолирующие слои запирают электроны в яме. Квантово-механическое туннелирование электронов может происходить как в яму, так и из ямы. Как только напряжение на контактах ямы окажется соответствующим энергии одного из ее резонансных состояний, туннельный ток возрастет.

Приповерхностный слой, из которого электроны уходят вглубь, образует изолирующую оболочку вокруг столбика. Из столбика с характерным размером 1000 \AA получается квантовая точка диаметром 100 \AA .

Конструктивно квантовая точка в виде столбика не совсем удобна. В альтернативной конструкции поверх и ниже лежащего слоя материала, образующего квантовую яму, находятся электроды определенной формы. Если на эти электроды подать напряжение, то электростатическое поле вытолкнет электроны из нижележащего слоя везде, за исключением малых островков — квантовых точек. Квантовым запираением электронов в островках можно управлять напряжением электродов.

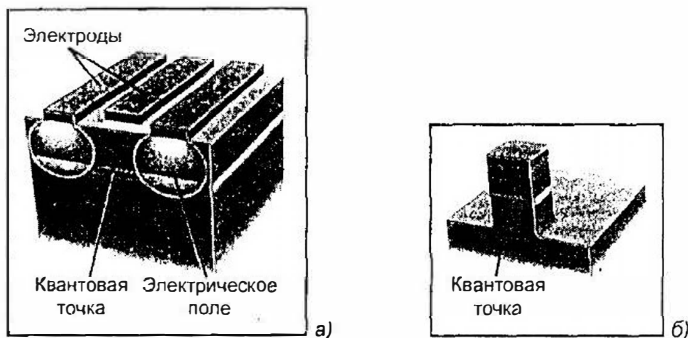


Рис. 15.9. Квантовая точка, управляемая электродами (а), и квантовая точка в виде столбика (б)

Исследования показали, что в каждой точке можно запирает по одному электрону, добавляя или вычитая по одному электрону за один процесс, что аналогично биту в цифровой технике.

15.6. Квантовые эффекты в наноструктурах

15.6.1. Одноэлектронное туннелирование

Одноэлектронное туннелирование впервые было предсказано профессором МГУ К. К. Лихаревым в 1986 году.

Рассмотрим туннельный переход между двумя металлическими контактами и тонким слоем диэлектрика между ними. По сути дела это плоский конденсатор емкостью C , на обкладках которого находится заряд Q . Энергия, запасенная в таком конденсаторе, равна

$$\epsilon = \frac{Q^2}{2C}. \quad (15.6)$$

Изменение емкости конденсатора происходит дискретно, и минимальное значение изменения энергии определяется так:

$$\Delta\epsilon = \frac{q^2}{2C},$$

где q — заряд электрона.

С одной стороны, для наблюдения эффекта туннелирования необходимо, чтобы $\Delta\epsilon \gg kT$, где k — постоянная Больцмана, T — температура. Тогда для таких температур $\beta = q^2 / 2CkT \rightarrow \infty$, а квантовая единица сопротивления $R \gg R_Q = h / 4q^2 \sim 6,45$ кОм.

С другой стороны, необходимо, чтобы

$$\Delta\epsilon \gg \frac{\hbar G}{C},$$

где $G = \max(G_s, G_t)$, G_s — проводимость шунтирующего перехода, G_t — проводимость туннельного перехода. В этом случае $G \ll R_Q^{-1}$ и частоты проходящих процессов определяются из соотношения $\tau \approx RC$, а $\omega\tau \ll 1$.

Важнейшим предположением теории одноэлектронного туннелирования является тот факт, что начальный заряд Q_0 может быть отличен от нуля и принимает значения, кратные целому числу электронов.

Для обеспечения процесса туннелирования через переход необходимо преодолеть силу кулоновского отталкивания электронов. Это так называемый эффект кулоновской блокады.

Кулоновская блокада представляет собой явление отсутствия тока при приложении напряжения к туннельному переходу из-за невозможности туннелирования электронов вследствие их кулоновского отталкивания. Для преодоления кулоновской блокады необходимо приложить к переходу напряжение $U_{ки} = q / 2C$.

Процесс протекания тока через одиночный туннельный переход происходит в несколько стадий.

На первой стадии в начальный момент времени граница между металлом и диэлектриком является электрически нейтральной. Электрический ток является величиной непрерывной. Для его поддержания необходимо на одной стороне туннельного перехода накопить определенный заряд. На рис. 15.10 представлен процесс одноэлектронного туннелирования в условиях кулоновской блокады. Здесь можно провести аналогию с каплей воды, отрывающейся от края трубки, предложенную К. К. Лихаревым.

На второй стадии к металлическим обкладкам прикладывается электрический потенциал и на границе раздела начинает накапливаться заряд. На параллельной схеме начинает формироваться капля.

В ходе третьей стадии происходит накопление заряда до тех пор, пока его величины не будет достаточно для возникновения туннелирования одного единственного электрона через диэлектрик.

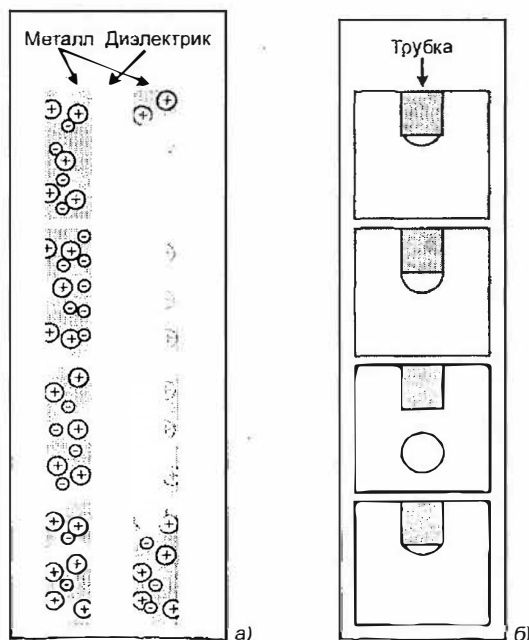


Рис. 15.10. Стадии процесса одноэлектронного туннелирования (а) и аналог образования капли в трубке (б)

На четвертой стадии после акта туннелирования система возвращается в первоначальное состояние. При сохранении внешнего приложенного напряжения цикл повторяется вновь. Перенос заряда в такой структуре осуществляется порциями, по одному электрону. Частота перехода определяется величиной

$$f = I / q,$$

где I — ток через переход. Такие осцилляции называются *одноэлектронными туннелированными осцилляциями* (SET, Single Electron Tunneling).

На конструкцию одноэлектронных приборов накладываются жесткие условия на емкость прибора:

$$C \ll q^2 / 2kT.$$

Подставив численные значения, получим, что при температуре $T = 4,2$ К необходимо создать емкость $C \ll 2 \times 10^{-16}$ Ф, для $T = 77$ К — $C \ll 10^{-17}$ Ф, а для $T = 300$ К — 3×10^{-18} Ф или 0,1—1,0 аФ (аттоФарада).

В системе с одним переходом технологически невозможно осуществить процесс одноэлектронного туннелирования.

15.6.2. Транспортный эффект Ааронова — Бомы

Транспортный эффект был открыт в 1959 году Якиром Аароновым и Дэвидом Бомом. Суть эффекта Ааронова — Бомы (ЭАБ) заключается в том, что внешнее электромагнитное поле, сосредоточенное в недоступной для заряженной частицы области, влияет на квантовое состояние этой частицы. Этот эффект проявляется в сдвиге фаз вдоль различных возможных путей распространения частицы. Эффект существует как для скалярного, так и векторного потенциала электромагнитного поля.

Эффект можно наблюдать, используя следующую экспериментальную схему (рис. 15.11).

Поток электронов в плоскости " $a - a$ " расщепляется таким образом, чтобы его составляющие огибали магнитный поток с двух сторон. В плоскости " $b - b$ " пучки соединяются, и происходит их взаимная интерференция. Относительная фаза электронов в обоих электронных потоках определяется магнитным потоком соленоида Φ . Магнитный поток может менять интерференционную структуру суммарного потока. Соответственно меняется электронный ток и проводимость структуры.

Сгенерируем и пропустим через плоскость " $a - a$ " два электрона, волновые функции которых имеют одинаковые начальные фазы, а также начальные ($A_1(0)$ и $A_2(0)$) и конечные ($A_1(L)$ и $A_2(L)$) амплитуды. В плоскости " $b - b$ " коэффициент прохождения имеет вид:

$$D = \left| \frac{A_1(0) \exp(jK_1 L) + A_2(0) \exp(jK_2 L)}{A_1(0) + A_2(0)} \right|^2.$$

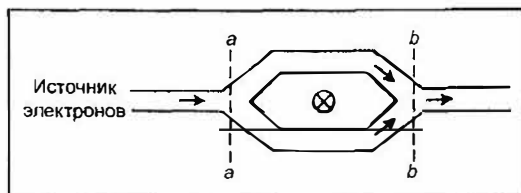


Рис. 15.11. Схема наблюдения эффекта Ааронова — Бомы

Если в плоскости "а — а" электроны находились в одинаковых состояниях, то

$$A_1(0) = A_2(0)$$

$$D = \cos^2 \left[(k_1 - k_2) \frac{L}{2} \right].$$

Приложенное к структуре магнитное поле с индукцией

$$B = \frac{\Phi}{S_{эф.}}$$

изменит электронные потоки так, что

$$(k_1 - k_2)L = \frac{q\Phi}{h},$$

где L — длина колена. Тогда коэффициент прохождения D примет вид:

$$D = \cos^2 \left(\frac{q\Phi}{2h} \right).$$

В этом случае электронный ток и проводимость структуры Ааронова — Бома (АБ) начнут периодически осциллировать с периодом h/q при изменении магнитного потока Φ .

При этом имеет место и электростатический эффект Ааронова — Бома, в котором на поток электронов воздействуют электростатические поля от конденсаторных электродов. Постоянное напряжение, прикладываемое к конденсаторным электродам, приводит к накоплению фазы в каналах петли на величину

$$\frac{q}{h} \int V dt.$$

Изменение напряжения ΔV будет вызывать осцилляцию коэффициента прохождения

$$D = \cos^2 \left(\frac{q\bar{\tau}\Delta V}{2h} \right),$$

где $\bar{\tau}$ — среднее время пролета электронов через канал длиной L .

Изменение фазы происходит из-за изменения длины волны электронов и определяется временем пролета участка, находящегося в электрическом поле. Таким образом, в магнитном эффекте АБ изменение фазы пропорционально площади, охваченной магнитным потоком. В электростатическом эффекте определяющим фактором является длина канала, на котором действует электрическое поле.

15.7. Новые транзисторные структуры

Нанoeлектроника является, с одной стороны, неким продолжением развития традиционных идей микроэлектроники, связанных со схемотехническими методами обработки информации. С другой стороны, совершенно новые идеи, касающиеся в квантово-размерных явлений и эффектов, открывают широкие возможности создания приборов нового поколения. Американская программа развития технологии "The National Technology Roadmap For Semiconductors" предусматривает, в частности, создание к 2015 году транзисторов с шириной затвора 200 Å. Это позволит разместить 10^9 транзисторов на 1 см^2 интегральной схемы при частотах 30—40 ГГц.

15.7.1. Полевые транзисторы

Традиционные полевые транзисторы, выполненные по субмикронной технологии, приобретают новые свойства. Уменьшение толщины окисла, длины канала приводит к квантованию поперечного движения в канале. В результате образуется квазидвумерный газ носителей заряда, увеличивается их подвижность и туннельный ток. При длинах затворов транзисторов до 20 нм они становятся сравнимы с длиной когерентности и длиной волны Де-Бройля при комнатной температуре.

Используя кремниво-германиевую технологию, были получены двух- и трехзатворные транзисторы, которые обладали более эффективным управлением. Фирма IBM создала кремниво-германиевый транзистор, способный работать 350 млрд циклов в секунду. На основе этой конструкции транзисторов можно получить специализированный микропроцессор для телекоммуникационных приборов, работающих на частоте 150 ГГц. Американская фирма AMD создала двухзатворный транзистор с длиной затвора 10 нм, изготовленный по технологии Fin Field Effect Transistor (Fin FET), при которой используется кремниевый плавник (FIN), позволяющий предотвращать утечки токов в отключенном состоянии, при этом резко возрастает производительность информационной системы. "Плавниковая" технология представляет собой островок кремния, высота которого больше, чем ширина, на котором и изготавливается транзистор. Короткий участок такого плавника и становится каналом. Правда, это уже не планарный прибор.

Весьма перспективна технология кремния на диэлектрике. Толщина кремния в 20 нм позволяет сформировать планарные транзисторы практически малых паразитных емкостей.

Полевой транзистор на гетероструктурах (ГСПТ) с использованием арсенида галлия представляет собой эпитаксиальную пленку, нанесенную на полупроводниковую подложку (рис. 15.12). Образуется гетеропереход с потенциальной ямой со стороны подложки, в которой формируется проводящий канал, подобный каналу в сильно инвертированном выраженном поверхностном слое.

Когда квазиуровень Ферми подвижных носителей, представляемый потенциалами истока и стока, пересечет дно ямы, то яма деформируется. Она станет узкой и состояние носителей в ней приобретет квантовый характер. Образуется квазидвумерный газ, подвижность носителей в котором увеличится благодаря уменьшению рассеивания носителя на примесях.

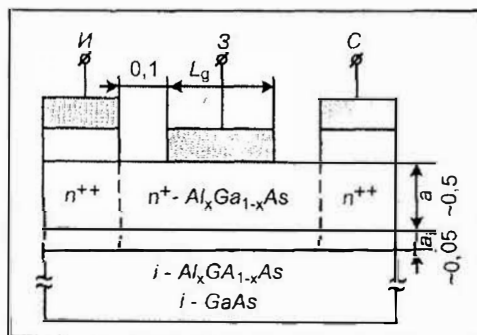


Рис. 15.12. Структура арсенид галлиевого транзистора:
И — исток; З — затвор; С — сток

Причиной тому является малая концентрация остаточных заряженных центров в буферном слое полупроводниковой подложки вблизи поверхности ($\sim 10^{14} \text{ см}^{-3}$) и высокое качество

во поверхности границы гетероперехода. Формирование со стороны эпитаксиальной пленки тонкого, порядка нанометра, нелегированного пограничного слоя, который определяет канал от сильнолегированной области в эпитаксиальной пленке также способствует увеличению подвижности носителей. Наиболее эффективны и каналные ГСПТ, выполненные на основе гетероструктуры типа $n^+-Al_xGa_{1-x}As / i-GaAs$ со значением $x \sim 0,3$. При нормальной работе ГСПТ эпитаксиальная пленка обедняется носителями, а пространственный заряд в ней создается ионизированными донорами с концентрацией $\sim 10^{18} \text{ см}^{-3}$. При ширине затвора $L_g \sim 0,25$ и температуре 300 К обнаруживается квазибаллистический характер переноса носителей.

15.7.2. Транзисторы с резонансным туннелированием

Технология молекулярно-лучевой эпитаксии позволяет получать различные структуры с квантовыми ямами, создавать туннельные полупроводниковые гетероструктуры с определенными потенциальными барьерами, сверхрешетки с периодической структурой туннельных барьеров.

На рис. 15.13 приведена типовая структура транзистора на горячих электронах с резонансным туннелированием (Resonant Tunneling Hot Electron Transistors).

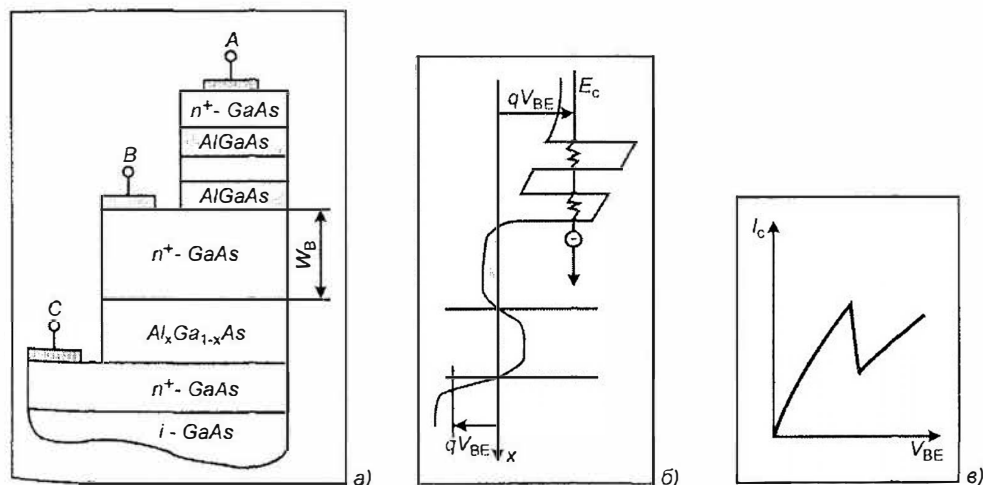


Рис. 15.13. Структура транзистора на горячих электронах с резонансным туннелированием (а), его энергетическая диаграмма (б) и выходная ВАХ (в)

Эмиттер имеет структуру, состоящую из двух потенциальных барьеров с заключенной между ними квантовой ямой шириной ~ 5 нм. За счет резонансного туннелирования в базу поступает монохроматический электронный поток. Характерной особенностью ВАХ такого транзистора является ее N -образный характер с участком отрицательной дифференциальной проводимости.

Научный и практический интерес представляют полупроводниковые гетероструктуры с квантовыми точками. Методом гетероэпитаксии слоя на несогласованной по параметру решетки подложке получены транзисторы, в которых квантовые точки встраиваются в "токовый" канал (рис. 15.14). Средний размер квантовых точек составляет ~ 40 нм, а кон-

центрация — $3 \times 10^{10} \text{ см}^{-2}$. В сильных полях была выявлена аномальная двухступенчатая форма ВАХ, что обусловлено вкладом квантовых точек в процесс протекания тока. Это объясняется двумя типами состояния электронов: подвижные двумерные электроны и электроны, локализованные на квантовых точках. Это принципиально новый тип транзистора на горячих электронах, который имеет хорошие перспективы использования в высокочастотных системах.

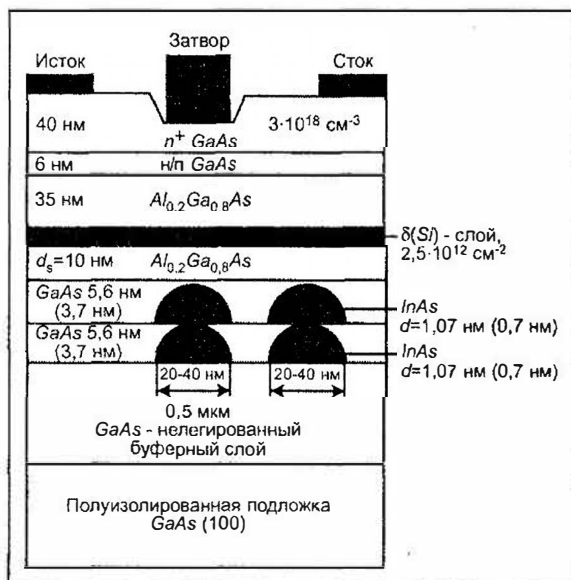


Рис. 15.14. Структура $n\text{-AlGaAs} / \text{GaAs} / \text{InAs} / \text{GaAs} / \text{InAs} / \text{GaAs}$ с квантовыми точками InAs, встроенными в приборный канал

15.8. Квантовые приборы наноэлектроники

В полупроводниковых квантовых приборах обработка, хранение и преобразование информации происходят путем контролируемой передислокации волновой функции в полупроводниковой наноструктуре.

Если в традиционных полупроводниковых приборах различные дискретные состояния, которые соответствовали логическим уровням и определялись инерционными кинетическими процессами диффузии и дрейфа, то в квантовых структурах дискретность состояний системы определяется дискретностью безынерционного квантово-размерного энергетического спектра. Это является фундаментальным свойством квантовых приборов.

В квантовых приборах обработка информации связана с управлением перестройки волновой функции электрона в квантовой структуре. Эта перестройка должна соответствовать заданному логическому закону преобразования информации. Один из способов перестройки волновой функции заключается в передислокации максимума вероятности нахождения электрона из одной части квантовой системы в другую. Эта передислокация осуществляется под действием внешнего напряжения, например, с помощью квантовых точек, разделенных туннельно-прозрачными барьерами.

Исходное распределение волновых функций должно быть несимметрично, чтобы электронная плотность была сосредоточена в одной квантовой яме. При приложении внешнего напряжения, повышающего уровень системы, происходит передислокация максимума амплитуды волновой функции из одной квантовой точки в другую.

Таким образом, передислокация волновой функции происходит путем межямной релаксации параллельно с процессом испускания фотона. Время переключения определяется временем релаксации, характерное значение которого составляет ~ 1 нс. С уменьшением туннельной связи между квантовыми точками это время экспоненциально растет. Считывание информации, закодированной в виде зарядного распределения волновой функции по квантовым ямам, осуществляется с помощью специальных контактов к каждой квантовой яме.

15.8.1. Квантовый вентиль

Две квантовые ямы могут использоваться как *квантовый логический элемент*. При малом напряжении на входе, соответствующем напряжению логического нуля, максимум амплитуды волновой функции находится в первой квантовой яме W_1 . Сопротивление первой ямы будет мало по сравнению с сопротивлением второй ямы (рис. 15.15, а).

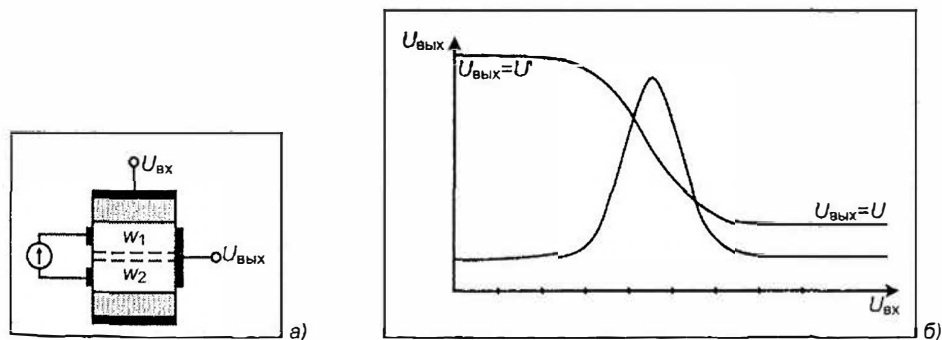


Рис. 15.15. Схема квантового инвертора (а) и его характеристика (б)

На выходе напряжение будет близко к напряжению источника питания — это напряжение логической единицы (рис. 15.15, б). При напряжении логической единицы на входном электроде волновая функция локализуется во второй яме W_2 . Соотношение сопротивлений изменится на противоположное по сравнению с ранее рассмотренным случаем. На выходе напряжение будет соответствовать напряжению логического нуля (рис. 15.15, б). Сквозной ток через структуру между электродами протекает только во время переходного процесса или в период передислокации максимального нуля волновой функции.

Это является полным аналогом КМОП-инвертора с той лишь разницей, что имеет близкую к нулю составляющую динамической мощности.

Быстродействие квантового инвертора определяется временем туннелирования электронов через барьер и составляет ~ 1 нс. Возможны и другие схемы включения квантовых ям для создания логических квантовых структур.

Например, резонансные туннельные диоды, получившие название квантовых точечных структур (точек). Квантовая точка представляет собой потенциальную яму нулевой раз-

мерности. В таких структурах возможно туннелирование между отдельными точками. На их основе можно создать матрицу квантовых точек с шагом 1 мкм, не имеющих между собой обычных межсоединений. Каждая квантовая точка имеет два состояния: наличие или отсутствие в ней электрона.

Электроны можно переключить в любом направлении между потенциальными ямами. Режим туннельного резонанса достигается путем выбора полярности напряжения смещения. Между двумя точками образуется зарядовая связь — кулоновское взаимодействие электронов соседних потенциальных ям. Моделирование таких матриц показало, что они могут выполнять логические функции без отказов и сбоев.

15.8.2. Квантовый интерферометр

В основе работы *квантовых интерферометров* лежит явление интерференции волновых пакетов. На вход прибора подается волновой пакет или отдельный цуг (рис. 15.16).

В плечах интерферометра предусматривается ввод цуга определенной фазы, позволяющей на выходе получить цуг, соответствующий, например, логической единице (рис. 15.16, а). Если в плечи интерферометра вводятся цуги соответствующей противофазы, то на выходе сигнала не будет (рис. 15.16, б). Такое положение может соответствовать логическому нулю. Предполагается, что приборы этого типа найдут широкое применение в однокристалльных компьютерах, схемах оптической связи.

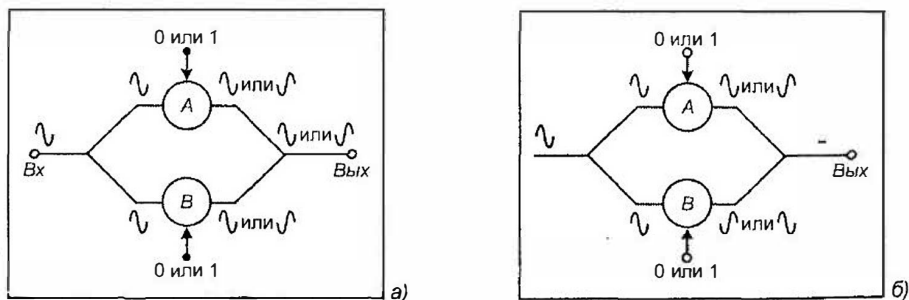


Рис. 15.16. Квантовый интерферометр на волновых пакетах

15.8.3. Квантовый каскадный лазер

В новом классе лазеров используются оптические переходы между минизонами сверхрешеток. Такой класс лазеров получил название *квантовых каскадных лазеров*.

В сверхрешетках с постоянным периодом отдельные уровни в квантовых ямах из-за туннелирования электронов сливаются в мини-зоны. При приложении к сверхрешетке внешнего электрического поля эти уровни могут разделиться по энергии. Мини-зоны в этом случае разрушаются.

Если же создать сверхрешетку с постепенно меняющимся периодом (*chirped*), то при приложении внешнего электрического поля наблюдается эффект схождения энергетических уровней в отдельных квантовых ямах в энергетическую мини-зону (рис. 15.17).

Таким образом, в неэквидистантной сверхрешетке под действием внешнего электрического поля формируются энергетические зоны — мини-полосы. В такой структуре появ-

дается возможность стимулировать вынужденное излучение, подобно тому, как оно возникает в активной среде оптического лазера.

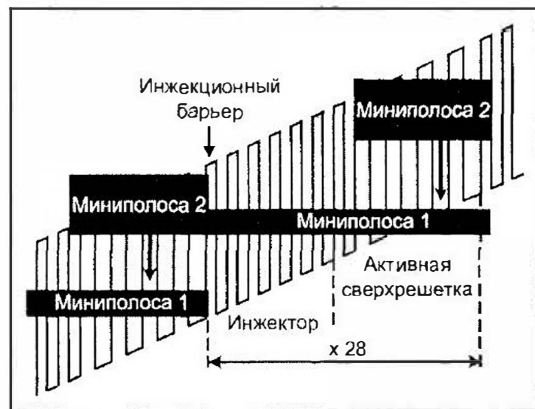


Рис. 15.17. Формирование энергетических уровней в структуре сверхрешетки

Для примера в подобной структуре удалось осуществить генерацию лазерного излучения на длине волны 7,6 мкм с пиковой мощностью 0,5 Вт при пороговой плотности тока $5 \text{ кА} / \text{см}^2$. Температура, при которой осуществлялась генерация, составляла 160 К (Applied Physics Letter, v. 73, p. 2101, 1998).

15.9. Одноэлектронные приборы

В туннельных переходах малой площади при низких температурах наблюдается дискретное туннелирование одиночных электронов. Это возможно в переходах с малой собственной емкостью C , при которой напряжение изменяется на величину $\Delta V = E / S$, причем это изменение напряжения должно быть больше термических флуктуаций $\Delta V \geq kT / E$, где k — постоянная Больцмана. Эта смелая идея легла в основу нового перспективного направления в нанoeлектронике — *одноэлектроники*.

Фундаментальным одноэлектронным устройством является *одноэлектронный транзистор*. Он содержит область проводимости, соединенную с истоковыми и стоковыми электродами — туннельными барьерами, которые имеют емкостную связь с затвором.

Изменяя напряжение на затворе, можно моделировать ток через область "исток—сток". Если создать несколько областей с взаимной емкостной связью, то возможно перемещение через эти области дискретных электронов. В этом случае можно реализовать логические функции. Одной из перспективных конструкций является *T-образная транзисторная структура* (рис. 15.18).

Транзисторы изготавливаются на кремниевой подложке по технологии изоляции имплантированным кислородом (SIMOX, Separation by IMplantation of OXygen). Такая технология позволяет получить тонкий слой кремния, изолированный от кремниевой подложки. С помощью процесса электронно-лучевой литографии возможно получение *T-образного* соединения из слоя кремния толщиной 30 нм, шириной плеча 40—50 нм и длиной плеча 50—80 нм. Кремниевые области проводимости сделаны так, что сужения кремниевых проводников образуют туннельные барьеры.

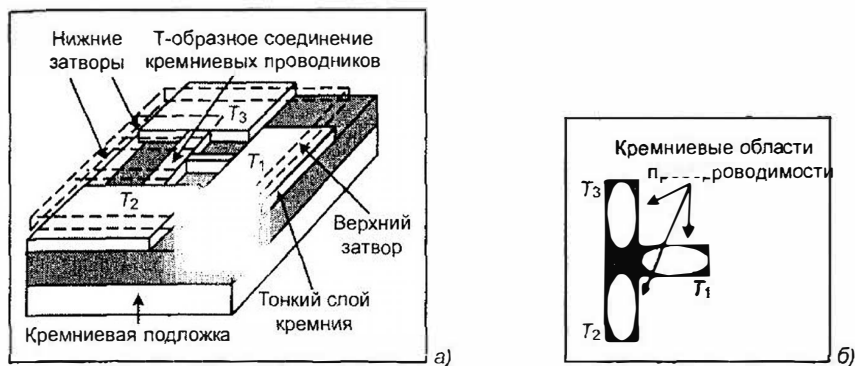


Рис. 15.18. Структура кремниевого T -образного транзистора (а) и схема расположения кремниевых областей проводимости (б)

Такое конструктивное решение стало возможным благодаря использованию эффекта самопроизвольного формирования сужения кремниевого проводника в процессе окисления. Этот эффект основан на двух противоположных механизмах окисления: снижении степени окисления в связи с ростом внутренних напряжений и рост степени окисления вблизи края проводника. T -образная конструкция позволяет снизить степень окисления в плечах из-за наличия внутренних напряжений. Одновременно можно увеличить степень окисления в точке соединения областей, где из-за относительно большой площади снижается внутреннее напряжение.

Тонкие поликремниевые затворы над каждым плечом T -образного соединения были сформированы с помощью электронно-лучевой литографии и процесса их легирования фосфором. Точность перекрытия в 20 нм позволяет использовать их для управления электростатическим потенциалом областей проводимости каждого плеча.

После изготовления тонких затворов был сформирован промежуточный изолирующий слой оксида кремния и верхний поликремниевый затвор. И наконец, используя верхний затвор как маску, формировались выводы истока и стока путем имплантации ионов фосфора. Фотография общего вида T -образного транзистора, сделанная с помощью электронного микроскопа, приведена на рис. 15.19.

Сложную транзисторную структуру можно разложить на отдельные части и для удобства расчета составить ряд эквивалентных схем.

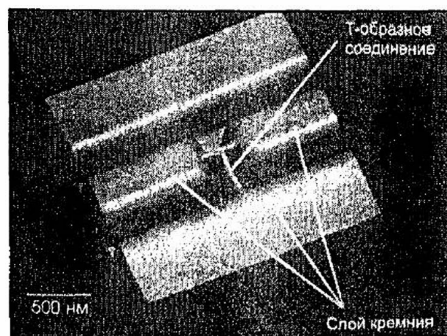


Рис. 15.19. Фото общего вида T -образной транзисторной структуры

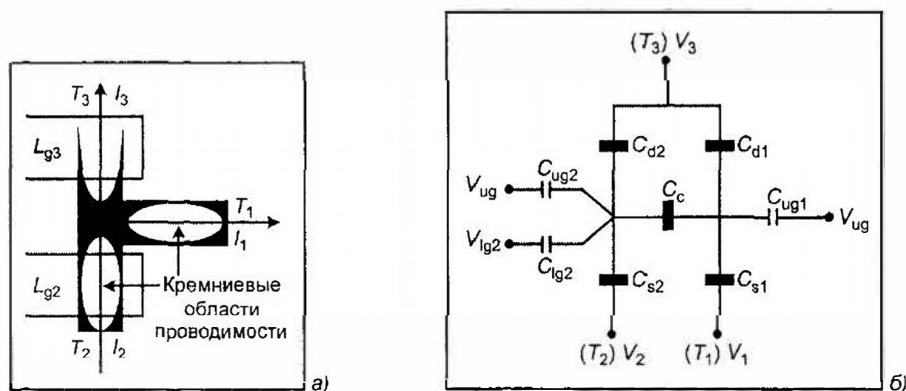


Рис. 15.20. Структурная схема (а) и эквивалентная схема устройства для двух областей проводимости (б)

Устройство состоит из двух одноэлектронных транзисторов, связанных между собой через туннельный конденсатор. Область в плече T_1 не имеет над собой затвора, и ее электростатический потенциал управляется верхним затвором L_{g3} (рис. 15.20).

Потенциал кремниевой области в плече T_2 управляется верхним и нижним затворами L_{g2} и L_{g3} . Это возможно из-за неполной экранировки области проводимости каждой из областей. Верхний электрод играет роль общего затвора, управляющего всеми кремниевыми областями проводимости.

Управление напряжением нижнего затвора позволяет получать различные конфигурации транзисторных структур. Если включен канал J_3 , то получаются два параллельно включенных одноэлектродных транзистора (рис. 15.21, а).

Если считать напряжение в плече T_3 , то формируется аналог инвертора КМОП-типа (рис. 15.21, б). Если канал J_3 выключен, то формируется схема из двух последовательно включенных транзисторов, работающих в одноэлектронном устройстве передачи. Эта схема также называется *одноэлектронным насосом* (рис. 15.21, в).

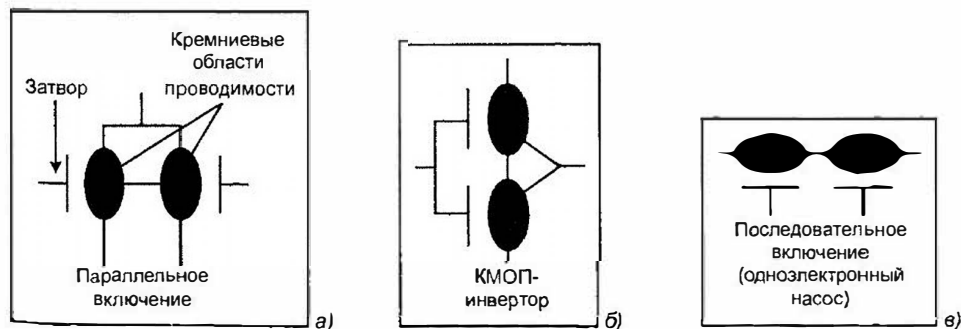
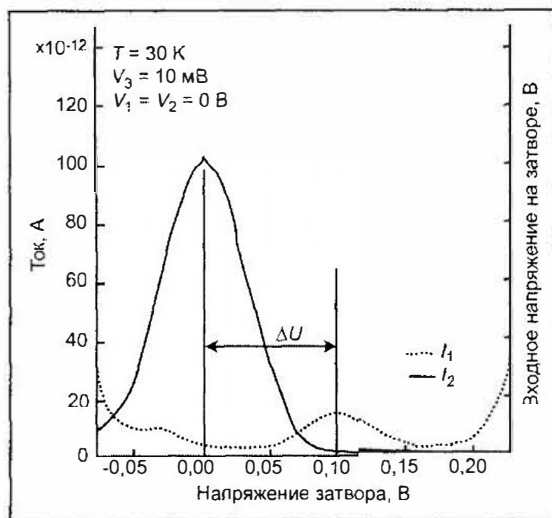


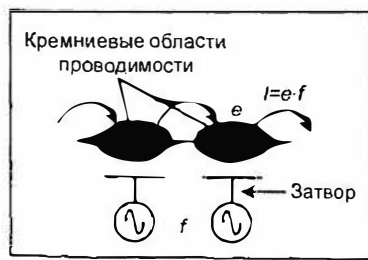
Рис. 15.21. Формирование различных структур при включенном (а, б) и выключенном канале J_3 (в)

Устройство может быть использовано для коммутации тока между двумя кремниевыми областями. Если подать напряжение $V_3 = 10$ мВ, а $V_1 = V_2 = 0$, то вольтамперная характеристика будет иметь вид, представленный на рис. 15.22.

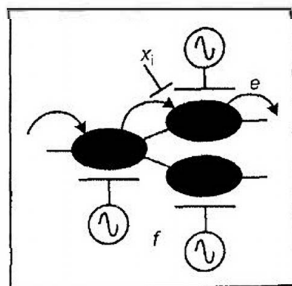
Рис. 15.22. ВАХ одноэлектронного насоса ($T = 30 \text{ К}$)

На ВАХ видны колебания, связанные с эффектом кулоновской блокады. Различные периоды колебаний связаны с размерами областей. Противоположность колебаний позволяет осуществить коммутацию тока между плечами T -образного соединения. Отметим, что переключение токов с соотношением $1 : 20$ реализовано в области $200 \times 300 \text{ нм}^2$.

На основе T -образных транзисторных структур может быть реализован ряд уникальных схем. Одноэлектронный коммутатор или одноэлектронный насос представлен на рис. 15.23, а. Гармонический сигнал управления с частотой f подается на затворы и позволяет прокачивать одиночные электроны. Ток в цепи I определяется величиной $I = e \cdot x f$. На рис. 15.23, б представлен одноэлектронный коммутатор, использующий три области проводимости.



а)



б)

Рис. 15.23. Схема устройств перемещения одиночных электронов: одноэлектронный насос (а) и коммутатор одноэлектронного перемещения (б)

В этом устройстве соединены два одноэлектронных насоса. Каждый одиночный электрон может перемещаться по одному из двух путей — плат T -образного соединения. Выбор конкретного пути перемещения определяется входным управляющим сигналом X . Такие устройства позволяют осуществить построение логических структур с использованием

отдельных электронов. Эти одноэлектронные структуры могут иметь сверхвысокую степень интеграции и сверхмалое потребление мощности при минимальных габаритах.

15.10. Новые материалы наноэлектроники

С 60-х годов прошлого века стала известна третья модификация углерода, структура которого состоит из одномерных линейных цепочек. В 1960 году были найдены стабильные долгоживущие кластеры C_n , где $n = 32, 44, 50, 58, 60, 70, 72, 78, 80, 82$ и т. п. Наиболее стабильные из них C_{60} и C_{70} назвали *фуллеренами* в честь американского архитектора и изобретателя Ричарда Фуллера, предложившего строительные конструкции в виде полусфер или полусфероид, состоящих из фрагментов многогранных поверхностей.

Аллотропная модификация углерода C_{60} получила название *фуллерит*.

Фуллерены, которые образуют соединения и комплексы с простыми элементами, получили название *фуллериды*.

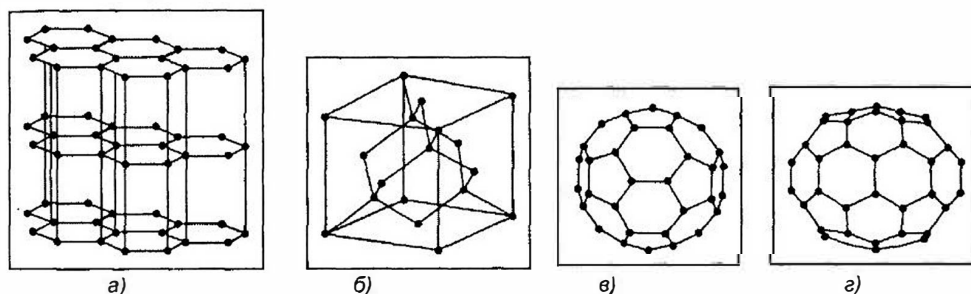


Рис. 15.24. Аллотропные модификации углерода: а — структура графита; б — структура алмаза; в — структура бакминстер фуллерена (C_{60}); г — структура фуллерена C_{70}

В отличие от графита и алмаза, представляющих собой периодическую решетку атомов, фуллерены представляют собой молекулы. Минимальный элемент фуллерена состоит не из атомов, а из молекул. Каркас молекулы C_{60} , например, состоит из 12 правильных пятиугольников (пентагонов) и 20 неравносторонних шестиугольников (гексагонов).

Впервые фуллерены были обнаружены при анализе продуктов абляции в результате воздействия лазерного излучения на графит.

Фуллерены нашли применение в наноэлектронике в качестве механически прочных резистов с высоким разрешением. Добавление фуллеренов позволяет повысить стойкость резистов к плазменному травлению и уменьшить оптическое поглощение на длине волны 193 нм. Сейчас резисты, содержащие фуллерен, выходят на уровень эффективного коммерческого использования.

В 1991 году японский исследователь Сумио Инджима изучал осадок, образующийся на катоде при распылении графита в электрической дуге в вакууме или инертном газе. Цепочки осажденных атомов углерода завиваются в трубочки диаметром несколько нанометров, длиной до нескольких микрон. Концы трубок закрыты одно- или многослойными полусферическими молекулами, напоминающими по структуре фуллерены.

Протяженные структуры, состоящие из свернутых гексагональных сеток с атомами углерода в узлах, получили название *нанотрубок*. На рис. 15.25 представлены нанотрубки, полученные в исследовательском центре компании NEC.

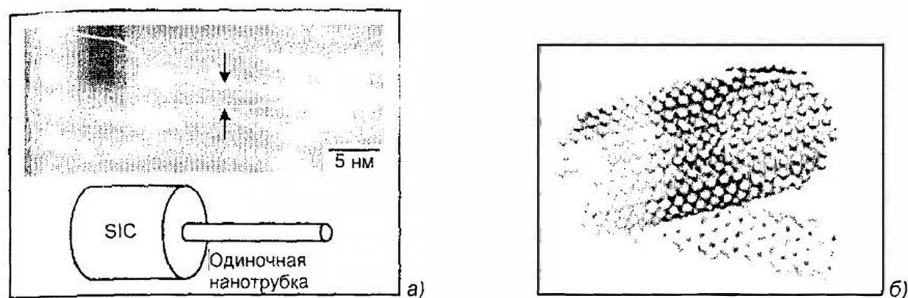


Рис. 15.25. Углеродные нанотрубки: общий вид в электронном микроскопе (а) и их структура (б)

Нанотрубки обладают рядом уникальных свойств. Однослойные нанотрубки могут удлиняться на 16% длины. Высокая механическая прочность, в сотню раз превосходящая прочность стали, обеспечивает возможность их применения в качестве зондов в сканирующем туннельном микроскопе.

Из нанотрубок можно получить сверхпрочную и эластичную ткань. Нанотрубки с открытыми концами проявляют капиллярный эффект и способны втягивать жидкости, а также расплавленный металл. Это позволяет на основе нанотрубок изготавливать сверхпроводящие нити диаметром в несколько нанометров. Высокая проводимость нанотрубок сравнима со сверхпроводимостью.

Электрические свойства нанотрубок определяются их хиральностью. Под хиральностью понимается угол ориентации графитовой плоскости относительно оси трубки. В зависимости от хиральности однослойная трубка может проявлять свойства графитополуметалла, не имеющего запрещенной зоны. Нанотрубка может обладать и свойствами полупроводника с шириной запрещенной зоны в пределах 0,01—0,70 эВ.

Если состыковать две нанотрубки, имеющие разную хиральность, то возможно создание $p-n$ -перехода. Размер такого перехода составляет несколько нанометров и открывается возможность компоновки электронных устройств.

В современных чипах с помощью нанотрубок возможна реализация ультраплотной мезинтеграции. Для этого необходимо научиться приделывать механические контакты к нанотрубкам различной хиральности. Эти контакты не должны содержать дефектов на атомарном уровне.

Электронные свойства нанотрубок можно целенаправленно менять путем их заполнения ионами различных веществ.

Нанотрубки обладают высокими эмиссионными характеристиками. При напряжении 500 В при комнатной температуре плотность автоэлектронной эмиссии достигает величины 0,1 А/см².

На основе углеродных нанотрубок разрабатываются ультратонкие дисплейные панели, по своим параметрам превосходящие жидкокристаллические.

Одной из особенностей углеродных нанотрубок является высокое значение диамагнитной восприимчивости. Большая отрицательная магнитная восприимчивость нанотрубок, видимо, обусловлена протеканием электронных токов по окружности нанотрубок. Диамагнетизм усиливается при низких температурах. Однако при применении нанотрубок возникают следующие проблемы. Во-первых, существующие методы получения нанотрубок имеют эмпирический характер. Технологически они не обработаны и поэтому не позво-

ляют получить нанотрубки заданного диаметра и длины заданной хиральности. Во-вторых, для практического использования нанотрубок необходимо также решить ряд проблем, связанных с квантованием электрического и магнитного полей, воплотить идеи создания гетероструктур в однослойной нанотрубке.

Контрольные вопросы

1. Что такое наноэлектроника?
2. Что такое нанотехнология?
3. Что является носителем информационного сигнала в приборах квантовой наноэлектроники?
4. В чем смысл квантового ограничения для электронов проводимости?
5. Что представляет собой явление резонансного туннелирования электронов?
6. Что представляет собой явление одноэлектронного туннелирования электронов?
7. Опишите конструкцию транзистора с резонансным туннелированием.
8. Опишите конструкцию инвертора на квантовых вентилях.
9. Опишите конструкцию квантового интерферометра.
10. Что представляют собой одноэлектронные приборы?
11. Что представляют собой фуллерены?
12. Что такое нанотрубки?

Рекомендуемая литература

1. Алексенко А. Г. Основы микросхемотехники. — 3-е изд. — М.: ЮНИМЕДИАСТАЙЛ, 2003.
2. Алексин А. П. Физико-химические основы субмикронной технологии. — М.: МИФИ, 1996.
3. Драгунов В. П. Неизвестный И. Г., Гридчин В. А. Основы наноэлектроники. — Новосибирск, Изд. ИГУ, 2000.
4. Кульбачинский В. А. Структуры малой размерности в полупроводниках. — М.: Изд. МГУ, 1998.
5. Ландау Л. Д., Лифшиц Е. М. Теоретическая физика. т. 3. — М.: Наука, 1995.
6. Херман М. Полупроводниковые сверхрешетки. — М.: Мир, 1989.
7. www.semiconductorline.com.
8. www.research.com/pics/nanotech.

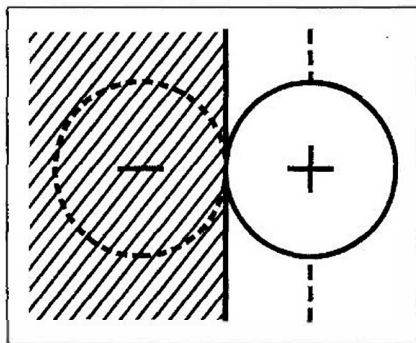
Заключение

Начало XXI века знаменует логический переход от микроэлектроники к наноэлектронике. Из этого вовсе не следует, что микроэлектроника с ее развитой микроэлектронной индустрией не станет. Микроэлектроника останется навсегда, потому как мир без нее уже не может существовать.

В наноэлектронике еще предстоит научиться создавать совершенные технологические линии, создавать эффективные приборы.

Параллельно развиваются новые отрасли знаний, такие как наномеханика, нанооптика, нанотехнология, нанометрология, нанолитография, наноэлектродинамика...

По мнению экспертов, ключевым фактором успешного развития страны является интенсивное развитие высоких технологий. Именно микроэлектронная технология и нанотехнология являются самыми передовыми технологиями современности.

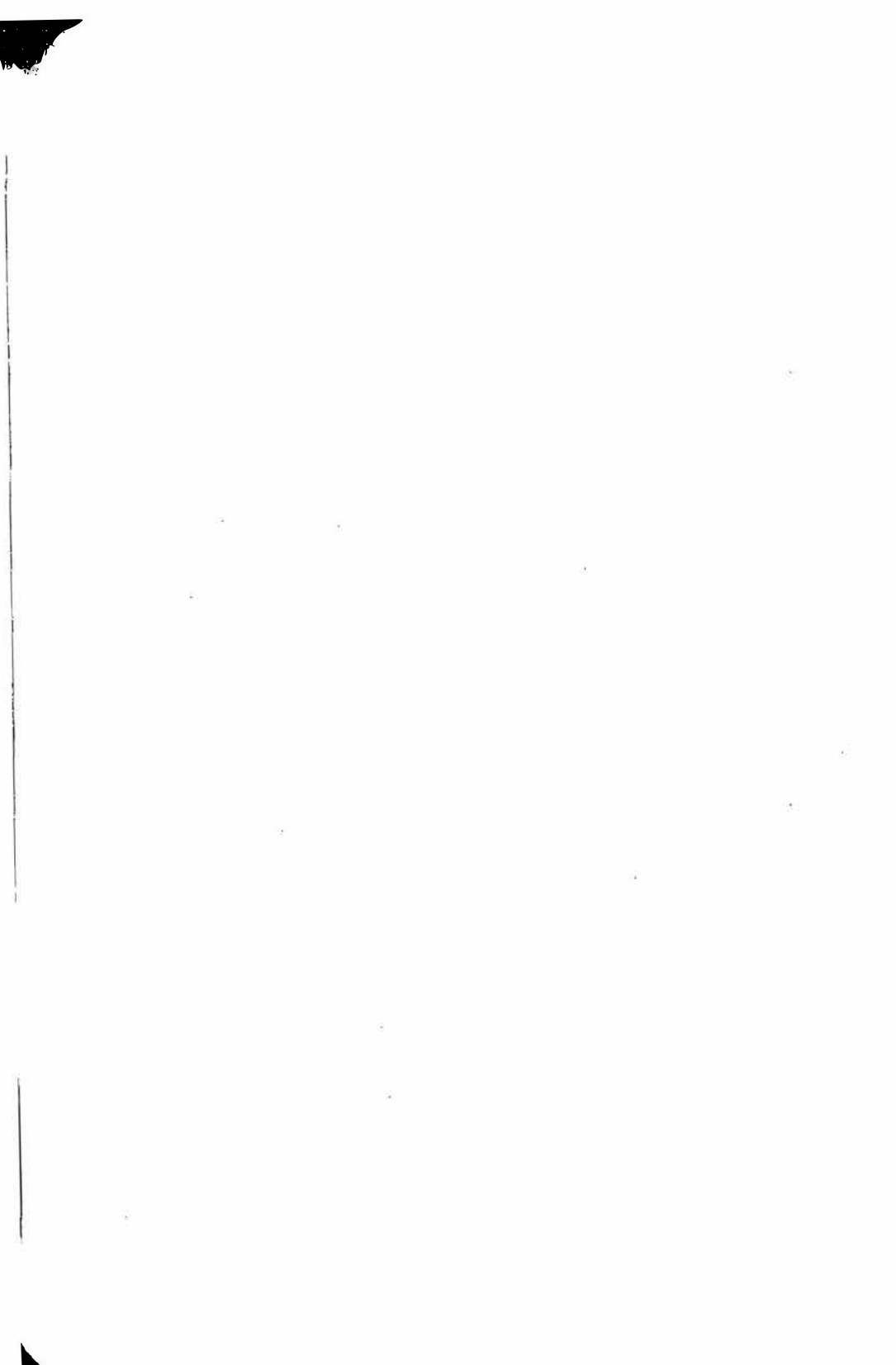


ЧАСТЬ III

Квантовая и оптическая электроника

Краткая историческая справка

1. Физические основы квантовой электроники
2. Принципы работы лазера
3. Типы лазеров
4. Оптические волноводы
5. Устройства управления световыми потоками
6. Приемники излучения
7. Введение в интегральную оптику
8. Введение в оптоэлектронику
9. Оптические методы обработки информации



Краткая историческая справка

Квантовая электроника возникла в середине 50-х годов на стыке квантовой механики и теории излучения, радиофизики и радиоскопии физики твердого тела и физики низких температур, электроники СВЧ и оптики, полупроводниковой и вакуумной электроники.

Истоки квантовой электроники восходят к исследованиям А. Эйнштейна, теоретически предсказавшего в 1917 году существование явления индуцированного излучения. Явления спонтанного и индуцированного излучения он постулировал как вероятностный закон. Эйнштейн объединял эти два процесса общим термином "изменение состояний под действием облучения" и считал, что:

- индуцированное излучение пропорционально плотности излучения, воздействующего на молекулу;
- частота излучения молекулы в точности равна частоте воздействующего поля;
- в условиях термодинамического равновесия между энергетическим состоянием молекул и полем излучения процессы индуцированного поглощения и индуцированного излучения имеют одинаковую вероятность;
- индуцированные и вынуждающие излучение кванты света абсолютно тождественны (когерентны).

Все это свидетельствует о безошибочной интуиции Эйнштейна и слабости существующей тогда теории излучения.

Строгая квантово-механическая теория излучения была развита трудами Пауля Дирака лишь в 1927—1930 гг. Впервые было доказано, что в элементарном акте излучения индуцированное и вынуждающее излучения обладают одной и той же частотой, одним и тем же состоянием поляризации и распространяются в одном и том же направлении. Другими словами, имеет место абсолютная когерентность вынуждающих и индуцированных квантов света.

Дальнейшие усилия исследователей были направлены на поиск в природе явлений индуцированного излучения. Для этого необходимо было получить среду с инверсной населенностью уровней или, как ее называют, среду с отрицательной температурой. Такая активная среда позволила бы создать условия преобладания индуцированного излучения над поглощением. При этом квантовые системы (атомы, молекулы, ионы) находились бы в термодинамически неравновесном состоянии, характеризующимся тем, что число частиц, находящихся на более высоком энергетическом уровне, превосходит число частиц, находящихся на более низком уровне.

Мысль о существовании сред с нарушением термодинамического равновесия или с отрицательной температурой долгое время была спорной. Одним из серьезных экспериментальных вкладов в исследования, связанные с обнаружением индуцированного излучения, следует считать работы советского физика Валентина Александровича Фабриканта. Изучая оптические свойства газового разряда, он впервые предложил использовать молеку-

лярные примеси для избирательного разрушения нижних энергетических уровней путем соударений второго рода. Эта работа лишь в 1965 году была признана открытием и была удостоена золотой медали им. С. И. Вавилова.

Практически одновременно, в 1950 году, парижская группа физиков во главе с Альфредом Кастлером разработала метод оптической накачки для целей исследования структуры атомов. Десятилетие спустя именно этот метод был использован для образования инверсной населенности активной среды лазеров. А в то время ни Кастлер, ни кто-либо другой не связывали его со способом обнаружения и формирования индуцированного излучения. Индуцированное излучение впервые наблюдали американские физики Эдвард Милс Парсел и Роберт Вивиан Паунд, которые осуществили инверсию в кристалле фтористого лития путем быстрого изменения направления приложенного статистического магнитного поля.

Сказалось бы, все эти работы должны были привести к созданию приборов квантовой электроники. Однако этого не произошло. Квантовая электроника не родилась в оптике, где все источники света по своей природе являются квантовыми. До появления квантовой электроники все оптические источники излучали некогерентные и немонохроматические колебания. Отсутствовали в оптике и когерентные усилители, монохроматические генераторы.

В то же время в радиофизике существовали методы и концепции, позволяющие генерировать монохроматические колебания, усиливать их. В пятидесятые годы возникла *радиоспектроскопия* — отрасль науки, изучающая спектры молекул, атомов, ионов, попадающих в диапазон сверхвысоких частот. Бурное развитие радиоспектроскопии было связано интенсивным развитием техники сантиметровых волн. В радиодиапазоне возбужденных атомов гораздо больше, спонтанное излучение слабее, чем в оптике, а индуцированное излучение непосредственно определяет величину эффектов, наблюдаемых в радиоспектроскопических экспериментах.

Существующие методы генерирования и усиления монохроматических колебаний в радиодиапазоне не позволили существенно продвинуться в миллиметровый и инфракрасный диапазоны частот. Причиной тому были слишком малые размеры резонаторов генерирующих систем, конструирование которых было весьма затруднительно. Качественный скачок от старой оптической и инфракрасной спектроскопии к радиоспектроскопии привел к постановке задачи по созданию генераторов когерентного излучения в оптическом диапазоне, связанной с новым физическим явлением — *вынужденным излучением*.

В начале 50-х годов прошлого века в Физическом институте им. П. Н. Лебедева Академии наук СССР проводились работы по усовершенствованию радиоспектрометра. Занимаясь исследованиями в области микроволновой радиоспектроскопии, советские ученые Александр Михайлович Прохоров и Николай Геннадьевич Басов стремились расширить возможности использования индуцированного излучения молекул для генерации и усиления микроволнового диапазона, сделать этот метод пригодным для наблюдения тонкой и сверхтонкой структуры спектров молекул. Для этого нужно было повысить разрешающую способность и чувствительность спектрометров. Однако для повышения разрешающей способности необходимо было уменьшить ширину линии поглощения молекулы, которой свойственно уширение за счет эффекта Доплера. А. М. Прохоров предложил вместо однородного газа применять молекулярные пучки, движущиеся перпендикулярно направлению распространения электромагнитных волн. Но оказалось, что возможности чистых радиоспектрометров сильно ограничены из-за малых интенсивностей наблю-

даемых линий. Причиной тому служила незначительная разница населенностей уровней, обуславливающих квантовый переход в микроволновом диапазоне. Тогда и возникли революционные идеи: во-первых, искусственно изменить населенности уровней и, во-вторых, использовать эффект индуцированного излучения молекул (высокого, энергетического уровня) вместо эффекта поглощения. Эти идеи были впервые высказаны А. М. Прохоровым и Н. Г. Басовым (рис. 1).



Рис. 1. Академики, лауреаты Нобелевской премии Николай Геннадьевич Басов и Александр Михайлович Прохоров в лаборатории ФИАН НА СССР (1966 г.)

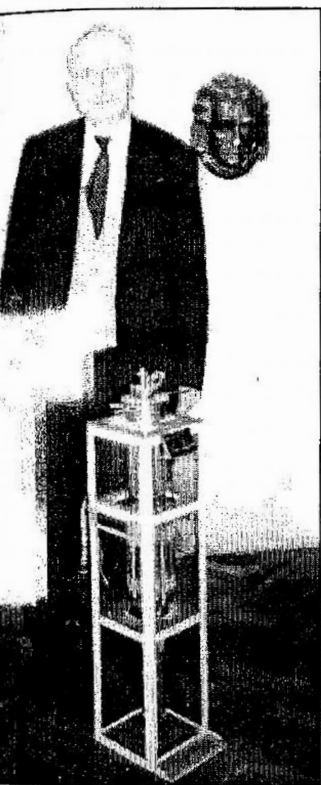
В течение 1952—1955 годов в ФИАН СССР был успешно осуществлен запуск первого квантового прибора — молекулярного генератора на пучке молекул аммиака. В дальнейшем на основе молекулярного генератора промышленностью были разработаны и выпускались стандарты частоты широкого применения. Созданные генераторы на пучке атомарного водорода, парах изотопа рубидия-87 с оптической накачкой нашли применение в космической технике, авиации, навигации, доплеровской локации, метрологии, службе времени и т. п.

В дальнейшем был проведен цикл исследований по разработке атомно-лучевых трубок для стандартов частоты, обеспечивающих стабильность частоты генерируемых колебаний не хуже 10^{-11} .

Практически одновременно предложение об использовании индивидуального излучения для генерирования миллиметровых волн было высказано американским физиком Чарльзом Хардом Таунсом с сотрудниками, работающими в области газовой радиоспектроскопии.

В 1955 году был предложен новый термин для обозначения молекулярного генератора — *лазер* (сокращенное от Microwave Amplification by Stimulated Emission of Radiation, усилитель микроволнового излучения за счет индуцированного излучения). В дальнейшем

мазерами стали называть и другие устройства, работающие на принципах индуцированного излучения (рис. 2).



Н. Г. Басов и А. М. Прохоров для получения активного состояния вещества предложили метод накачки, сущность которого состоит в использовании эффекта насыщения одного из переходов под действием вспомогательного излучения в многоуровневой квантовой системе. Метод накачки в системе с тремя уровнями оказался очень плодотворным и получил дальнейшее развитие при создании квантовых приборов других типов — *парамагнитных усилителей*.

Рис. 2. Академик А. М. Прохоров демонстрирует первый мазер (ФИАН, 1999 г.)

Этот метод был применен американским физиком Николасом Бломбергеном в предложенном им трехуровневом парамагнитном усилителе непрерывного действия. Разработка квантовых парамагнитных усилителей стала реальной после предложения А. М. Прохорова в 1957 году использовать в них монокристалл рубина. В парамагнитных усилителях на рубине впервые был получен неведомый для классической электроники уровень собственных шумов. Области применения квантовых парамагнитных усилителей явились радиоастрономия, радиолокация космических объектов, связь и телевидение через искусственные спутники Земли.

Таким образом, к шестидесятым годам в радиодиапазоне физики обладали всеми элементами квантовых генераторов: активной и трехуровневой средой, объемным резонатором, методами получения инверсной населенности.

Первые успехи в создании мазеров поставили на повестку дня вопрос о движении в сторону более коротких длин волн. И проблема перехода от радиоволн к инфракрасному, а затем и видимому диапазону решалась уже не радиометодами, а методами только что одившейся квантовой электроники.

Важно отметить, что при продвижении по шкале электромагнитных волн в сторону коротких волн существенную трудность представляет вопрос о резонаторах. В радиодиапа-

зоне резонаторами служили колебательные контуры, размеры которых были намного меньше длины волны, а в СВЧ-диапазоне размеры резонатора были уже сравнимы с длиной волны. В оптическом диапазоне такие резонаторы вообще невозможно было изготовить. Поэтому был выбран путь создания резонаторов, размеры которых были намного больше длины волны — так называемых открытых резонаторов. Это, прежде всего, известный в оптике *интерферометр Фабри-Перо*. Именно радиофизический подход позволил использовать эту конструкцию в качестве резонатора для лазеров. Впервые резонатор открытого типа был предложен А. М. Прохоровым в 1958 году.

Квантовая электроника как наука родилась именно в тот момент, когда возбужденная квантовая система была помещена в резонатор. И это сделал американский физик Теодор Гарольд Мейман (рис. 3, а). В августе 1960 года ему удалось получить на рубине генерацию в оптическом диапазоне волн. Им был создан первый лазер, работающий на рубине (рис. 3, б). Он также обнаружил и проанализировал два характерных явления, свидетельствующих о возникновении генерации путем индуцированного излучения: сокращение времени жизни возбужденного уровня в трехуровневой системе и резкое сужение линии излучения. Именно Т. Г. Мейман предложил термин *лазер* (Light Amplification Stimulated Emission of Radiation, усилитель светового излучения за счет индуцированного излучения).



а)



б)

Рис. 3. Изобретатель рубинового лазера Т. Г. Мейман (а) и первый рубиновый лазер (б)

Дальнейшие исследования выявили все основные особенности излучения лазера, а именно: возрастание интенсивности линии, пиковую структуру импульса, временную и пространственную когерентности, поляризацию, высокую спектральную яркость и направленность излучения.

Основным достоинством твердотельных лазеров явилась возможность получения с помощью больших значений и мощности в импульсе. Этот факт послужил толчком к созданию установок для обработки различных материалов. Уже в 1963 году были начаты работы по созданию лазерной технологической установки и с ее помощью проводились поиски и исследования в области технологии размерной обработки, закалки, сварки различных материалов. Одновременно проводились поисковые работы по использованию твердотельных лазеров в масс-спектрометрии, системах оптической локации, сверхбыстрой передачи информации, медицине.

На основе успехов в области твердотельных лазеров быстро развивалась новая отрасль оптики — нелинейная оптика.

Исследования возможности получения инверсной населенности в газах параллельно проводились в ФИАНе и в лабораториях ряда американских фирм. Николай Геннадьевич

Басов и Олег Николаевич Крохин провели теоретический анализ возможности создания состояния отрицательной температуры в газовом разряде путем столкновений второго рода в смеси газов. В конце 60-х годов американские физики Али Джаван и Уилард Харрисон Беннет сообщили о получении непрерывной генерации на сконструированном им гелий-неоновом лазере на пяти длинах волн в инфракрасном диапазоне электромагнитного спектра.

Отличие работы газового лазера от рубинового заключается, прежде всего, в методе создания инверсной населенности и накачке активного вещества. Отмечались высокая стабильность излучения первого газового лазера, временная и пространственная когерентности. Уже год спустя была получена генерация на десяти различных газовых смесях примерно на сорока различных оптических переходах, перекрывающих видимую оптическую и инфракрасную части спектра. Широко были поставлены исследования по изысканию оптимальных конструкций газоразрядных трубок, оптических резонаторов, простых и надежных источников накачки.

Первые отечественные промышленные газовые лазеры были созданы в 1963 году и сразу нашли широкое применение в научных исследованиях и отдельных областях народного хозяйства. На базе первого газового лазера ОКГ-11 был создан лазерный визир ЛВ-1 для управления горнопроходческим щитом под землей. В строительстве Останкинской телебашни применили лазерную осевую вертикаль, газовые лазеры использовались в первой экспериментальной оптической системе передачи телевизионных сигналов, а также в линии телефонной связи.

Помимо лазеров непрерывного действия был создан ряд импульсных лазеров, мощных азотных одночастотных и высокостабильных лазеров. Возник новый класс квантовых приборов — *квантовые оптические гироскопы*, обладающие высокой чувствительностью, быстродействием и устойчивостью к механическим нагрузкам.

При всех своих достоинствах твердотельные и газовые лазеры имеют один существенный недостаток — коэффициент их полезного действия (КПД) не превышает нескольких процентов. И в этом плане весьма перспективным представлялись полупроводниковые лазеры, КПД которых теоретически мог быть близок к ста процентам.

Первое теоретическое обоснование возможности создания полупроводникового лазера было сделано в 1958 году советскими физиками Н. Г. Басовым, Б. М. Вулом и Э. М. Поповым. Однако новый класс лазеров был создан лишь в 1962 году. В первых работах в качестве рабочего вещества был применен арсенид галлия, а инверсная населенность была достигнута путем инжекции электронов и дырок через $p-n$ -переход.

В 1963 года началось создание лазеров на $p-n$ -переходе на ряде полупроводниковых материалов, что значительно расширило спектральный диапазон излучения такого типа лазеров. Возможность прямой модуляции излучения путем изменения величины тока накачки, большое число методов накачки, высокий КПД — все это позволило широко использовать полупроводниковые лазеры. Потенциальные области применения полупроводниковых лазеров лежат, прежде всего, в создании устройств дальнометрии и высотомерии, для автоматической посадки транспортных средств, устройств стыковки космических объектов, в оптических линиях связи, в том числе и волоконных линиях связи, вычислительной техники. Полупроводниковые лазеры нашли широкое применение в оптоэлектронных устройствах.

Исследования в области создания лазеров на новых принципах накачки, новых активных веществах разворачивались широким фронтом. Появились сообщения о создании химиче-

ского лазера, энергия излучения которого была получена за счет неравновесного распределения химической энергии среди продуктов реакции непосредственно в газовом реакторе.

В Советском Союзе в 1962 году Постановлением Совета Министров был создан Государственный союзный НИИ № 333, основной задачей которого было лазерное приборостроение для нужд народного хозяйства и обороны. Институт возглавил Митрофан Федорович Стельмах — видный специалист в области СВЧ-электроники. Учеными и инженерами НИИ "Полус" (НИИ № 333) были выпущены десятки тысяч лазеров различных типов, разработаны лазерные технологии, опубликовано тысячи научных работ по квантовой электронике (рис. 4).



Рис. 4. Президент Академии наук СССР академик А. П. Александров, директор НИИ "Полус" М. Ф. Стельмах, академик А. М. Прохоров у стенда (1977 г.)

В течение последующих десятилетий появилась широкая гамма лазеров на нейтральных атомах и ионизированных газах, молекулярные лазеры, импульсные CO_2 -лазеры высокого давления, газодинамические лазеры на парах металлов и др.

Внимание исследователей было обращено на жидкие активные среды, которые в определенной мере сочетают преимущества твердотельных и газоразрядных, в которых может быть достигнута высокая концентрация активных центров. Это позволило надеяться на получение большой выходной энергии и мощности излучения с единицы объема рабочего вещества. Было предложено два класса жидких материалов — металлоорганические, или хелатные, и неорганические жидкости, определившие конструкции и характеристики жидкостных лазеров. В 1963 году было опубликовано первое сообщение о работе жидкостного лазера. Жидкостные лазеры породили появление лазеров на красителях.

Появились мощные газодинамические лазеры, лазеры на парах металлов и, наконец, лазеры на свободных электронах. Существующие лазеры работают в диапазоне длин волн от

вакуумного ультрафиолета до субмиллиметрового диапазона. Трудно назвать отрасль науки и техники, где бы не применялись лазеры и мазеры.

К числу наиболее важных достижений квантовой электроники является рождение в ее недрах оптоэлектроники.

Оптическая электроника или *оптоэлектроника* является направлением в электронике, в котором исследуются вопросы генерации, обработки и хранения информации на основе преобразований электрических сигналов в оптические (и наоборот), а также разработка приборов и устройств на основе этих исследований.

В основе оптоэлектроники лежат достижения квантовой электроники, такие как создание различных типов полупроводниковых лазеров и неогерентных светодиодов, генерирующих излучение в разных диапазонах оптического спектра. Одновременно в оптоэлектронике получило развитие направление, связанное с созданием фотоприемников излучения. Благодаря достижениям квантовой и оптической электроники появились волоконно-оптические системы связи, сотовая телефония, приборы и устройства интегральной оптики. На повестке дня современной оптоэлектроники стоит задача создания экономичных полупроводниковых источников света, способных заменить лампочки накаливания.

Данный перечень можно продолжить и указать заметные результаты, полученные благодаря использованию лазеров и оптоэлектронных приборов в научных исследованиях, технологиях, системах связи и телекоммуникации, производстве, учебном процессе...

1. Физические основы квантовой электроники

1.1. Спонтанное и вынужденное излучение

Квантовая электроника является областью электроники, в которой исследуются явления генерации и усиления электромагнитных колебаний на основе эффекта вынужденного излучения, явления нелинейного взаимодействия мощного излучения с веществом, а также возможность создания квантовых электронных приборов и устройств — молекулярных генераторов (мазеров), квантовых генераторов (лазеров), усилителей, устройств нелинейного преобразования частот лазерного излучения.

В квантовой электронике используются физические явления, в которых основное участие принимают связанные электроны. Эти электроны входят в состав систем из атомов, молекул, отдельных кристаллов, континуальных сред всех агрегатных состояний вещества.

В соответствии с законами квантовой механики, энергия электрона, связанного в атоме, имеет ряд дискретных значений $E_0, E_1, E_2, E_3, \dots, E_n, \dots$. Эти дискретные значения называются *уровнями энергии*. Весь набор разрешенных квантовой механикой уровней образует энергетический спектр атома. Основным уровнем E_0 назовем наименьший уровень. Все остальные уровни называются возбужденными.

Переход связанных электронов с одного уровня на другой сопряжен с излучением или поглощением электромагнитной энергии, частота которой определяется соотношением:

$$h\nu_{ij} = |E_i - E_j|, \quad (1.1)$$

где h — постоянная Планка, ν_{ij} — частота излучения (поглощения) при квантовом переходе с уровня E_i на уровень E_j .

Излучение и поглощение происходит отдельными порциями, квантами — фотонами: при поглощении фотона энергия атома увеличивается, при испускании фотона — уменьшается. При поглощении электрон переходит вверх на более высокий уровень, а при испускании фотона электрон совершает обратный переход вниз с уменьшением энергии атома. Такие скачкообразные переходы называют *квантовыми переходами*.

Различаются спонтанные (самопроизвольные) и вынужденные переходы.

При *спонтанном* квантовом переходе испускание фотона происходит вне зависимости от внешних факторов и воздействий на квантовую систему. При этом направление излучения и поляризация фотонов могут быть любыми.

Вынужденный квантовый переход происходит под воздействием внешнего излучения частоты ν , удовлетворяющего соотношению (1.1). При этом в процессе вынужденного излучения фотон имеет направление излучения и поляризацию, соответствующую этим же параметрам фотона стимулирующего излучения. Частота испущенного фотона в точности совпадает с частотой вынужденного излучения (рис. 1.1).

Атом находится в возбужденном состоянии некоторое время τ и скачкообразно переходит в невозбужденное состояние. Если предположить, что время жизни на возбужденном

уровне составляет τ , то при большом числе возбужденных частиц N , это время убывает по закону

$$N_t = N_0 \exp(-t/\tau), \quad (1.2)$$

где N_0 — общее число частиц, t — текущее время, τ — время, за которое число возбужденных частиц N убывает в e раз.

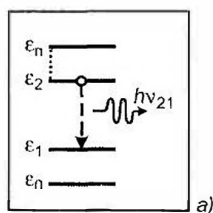
Частицы, потерявшие энергию и испустив фотон, могут вновь возбуждаться и переходить на уровень E_i . Чем меньше τ_i , тем чаще будут испускаться фотоны.

Величина $A = 1/\tau_i$ называется *вероятностью спонтанного испускания* с уровня E_i и определяет среднее число фотонов, испускаемых одной частицей за 1 с.

При спонтанном переходе с уровня E_2 на E_1 величина A_{21} называется *вероятностью перехода*. Полная вероятность A_i спонтанного испускания с уровня E_i на любой другой уровень равна сумме вероятностей отдельных спонтанных переходов:

$$A_i = \sum A_{i,k}, \quad (1.3)$$

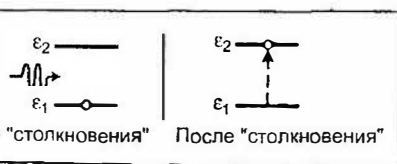
где $A_{i,k}$ — коэффициент Эйнштейна для спонтанного испускания. Типичное время жизни возбужденных атомов составляет $\sim 10^{-8}$ с.



а)



б)



в)

Рис. 1.1. Квантовые переходы:
а — спонтанный переход между уровнями E_1 и E_2 с генерацией фотона; б — вынужденный переход вниз с увеличением числа фотонов; в — вынужденный переход вверх с поглощением фотона

Вынужденные квантовые переходы генерируют фотоны, которые являются копией фотонов, стимулирующих усиление электромагнитного излучения. Число вынужденных квантовых переходов пропорционально плотности излучения ρ_ν на частоте $\nu = \frac{E_2 - E_1}{h}$.

Число фотонов $N_{k,l}$, поглощенных 1 см^3 за 1 с, пропорционально населенности n_k нижнего уровня E_k и плотности излучения ρ_ν :

$$N_{k,l} \approx B_{k,l} n_k \rho_\nu, \quad (1.4)$$

где $B_{k,l}$ — коэффициент Эйнштейна для поглощения в квантовой системе. Этот коэффициент характеризует вероятность поглощения и равен числу фотонов, поглощаемых одной частицей за 1 с, при приведенной плотности излучения ρ_ν . Вынужденное излучение характеризуется числом фотонов N_{ik} , испущенных в 1 см^3 за 1 с:

$$N_{ik} = B_{ik} n_i \rho_\nu, \quad (1.5)$$

где B_{ik} — коэффициент Эйнштейна для вынужденного испускания, который характеризует вероятность вынужденного испускания. Этот коэффициент определяется числом фотонов, испускаемых в среднем одной частицей под действием излучения плотности $\rho_\nu = 1$ за 1 с.

Эйнштейн рассмотрел равновесную систему, в которой число фотонов, испускаемых в переходе $E_i \rightarrow E_k$ с частотой $\nu = |E_i - E_k| / h$ равно числу фотонов той же частоты ν , поглощаемых при обратном квантовом переходе $E_k \rightarrow E_i$. Тогда:

$$N_{i,k} + N_{i,k} = N_{k,i},$$

$$A_{i,k} n_i + B_{i,k} n_i \rho_\nu.$$

Равновесная плотность излучения ρ_ν связана с коэффициентами Эйнштейна следующими соотношениями:

$$\begin{aligned} g_k B_{k,i} &= g_i B_{i,k} \\ A_{i,k} &= \frac{8\pi h \nu^3}{c^3} \frac{g_k}{g_i}, \\ B_{k,i} &= \frac{8\pi h \nu^3}{c^3} B_{i,k} \end{aligned} \quad (1.6)$$

где g_i и g_k — степени вырождения уровней E_i и E_k , соответственно, c — скорость света, коэффициент $\frac{8\pi \nu^3}{c^3}$ входит в формулу Планка для плотности энергии равновесного излучения.

Вероятность излучаемых переходов зависит от свойств уровней E_i и E_k , между которыми осуществляются квантовые переходы.

В квантовой системе имеет место резонансное поглощение фотонов. Суть этого явления заключается в том, что если атом находится на нижнем уровне E_i и электромагнитное излучение содержит n фотонов частоты ν_{ik} , то возможен переход атома на возбужденный уровень E_k . При этом поглощается фотон и число фотонов уменьшается и становится равным $(n - 1)$. Другими словами, вынужденные переходы вверх приводят к поглощению электромагнитной энергии вещества.

Различают излучательные (дипольные, магнитные и квадрупольные переходы) и безызлучательные квантовые переходы. Излучательные переходы сопровождаются изменением дипольного момента P_{ik} , магнитного момента M_{ik} , квадрупольного момента Q_{ik} , которые связаны с коэффициентами Эйнштейна.

При безызлучательных переходах изменение энергии квантовой системы связано с ее взаимодействием с другими квантовыми системами.

Среди квантовых переходов различают также *разрешенные* и *запрещенные* переходы. Если в какой-либо момент вероятность перехода отлична от нуля, то квантовый переход возможен. Если же в момент перехода его вероятность равна нулю, то квантовый переход невозможен и такой переход называется запрещенным.

В случае, когда квантовый переход с некоторого возбужденного уровня, называемого метастабильным, на более низкие уровни запрещен правилами отбора, то возбужденные состояния на таком уровне могут существовать довольно длительное время. Например, в рубине время жизни возбужденных атомов хрома на метастабильном уровне составляет

$\sim 10^{-3}$ с. Наличие долгоживущего метастабильного уровня позволяет создать на нем высокую населенность возбужденных состояний.

1.2. Спектральные линии

В идеальном случае из тонких уровней энергии возбужденные атомы должны излучать строго монохроматическое излучение одной частоты. Однако на практике излучение образует спектральную линию определенной ширины и формы. Даже для изолированных от внешних воздействий атомов линии излучения уширяются за счет фундаментального закона квантовой механики. Согласно принципу неопределенности, если Δt — время жизни атома в возбужденном состоянии, а ΔE — значение его энергии состояния, то они связаны соотношением Гейзенберга

$$\Delta E \cdot \Delta t \approx \hbar.$$

Неопределенность или "размытие" уровня обратно пропорционально времени жизни частицы τ_0 в начальном состоянии. Вследствие того, что время жизни свободной частицы на энергетическом уровне всегда конечно, то существует определенная естественная ширина спектральной линии.

Спектральное распределение квантов спонтанного излучения определяет ширину уровня

$$\Delta E = \hbar / \tau_0.$$

Контур линии спонтанного излучения имеет лоренцеву форму с шириной линии (рис. 1.2)

$$\Delta \nu_L = \Delta E / \hbar = 1 / 2\pi \tau_0.$$

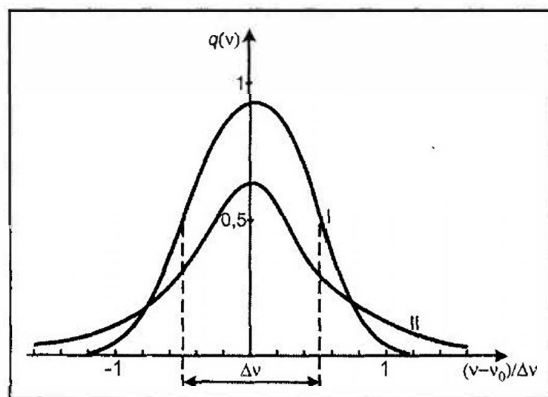


Рис. 1.2. Гауссова (I) и лоренцева (II) формы линий

Шириной линии называют интервал частот между точками, для которых интенсивность излучения (или поглощения) падает в два раза.

Лоренцева форма линии имеет вид резонансной кривой с максимумом на частоте ν_0 и описывается так называемым форм-фактором

$$q(\nu) = \frac{1}{2\pi (\nu - \nu_0) + \Delta \nu_L^2 / 4}.$$

В реальных условиях спектральные линии несколько размыты и представляют собой полосы излучения и поглощения. Причиной этому служат различные физические явления. Уширение линии происходит вследствие эффекта Доплера, вызывающего смещение частоты движущихся частиц.

Доплеровски уширенная линия описывается функцией Гаусса и симметрична относительно частоты ν_0 (рис. 1.2).

Форм-фактор доплеровски уширенной линии имеет вид

$$q(\nu) = \frac{1}{\nu_T \sqrt{\pi}} \exp \left[-\frac{\nu - \nu_0}{\Delta \nu_T} \right]^2,$$

где $\nu_T = \nu_0 u_0 / c$ — доплеровский сдвиг частоты при средней тепловой скорости движения излучающей частицы, u_0 — средняя тепловая скорость, c — скорость света. С увеличением частоты роль доплеровского уширения линии возрастает.

В твердых телах уширение спектральной линии и даже их расщепление возможно вследствие влияния электрических и магнитных полей (эффект Штарка, эффект Зеемана).

1.3. Поглощение и усиление

В естественных условиях при равновесии между средой и веществом нижние уровни энергии заселены более плотно, чем верхние. Существует фундаментальный закон распределения частиц по энергии

$$N_n = C \exp \left(-\frac{E_n}{kT} \right), \quad (1.7)$$

где C — константа, зависящая от полного числа частиц в единице объема, k — постоянная Больцмана. Это закон *Больцмана*.

Основная проблема, возникающая при создании квантовых усилителей и генераторов, состоит в поиске способов нарушения теплового равновесия рабочего вещества так, чтобы населенность верхних уровней была существенно выше населенности нижних уровней.

Система квантовых частиц, в которых хотя бы для двух уровней энергии более высоко расположенный уровень населен значительно больше нижнего, называют *системой с инверсной населенностью*.

Процесс инверсии населенностей уровней получил название *накачка*.

В соответствии с уравнением (1.7) отношение населенностей верхнего уровня N_2 к нижнему N_1 определяется соотношением:

$$\frac{N_2}{N_1} = \exp \left(-\frac{E_2 - E_1}{kT} \right) = \exp \left(-\frac{h\nu_{21}}{kT} \right), \quad (1.8)$$

где $\nu_{21} > 0$ — частота перехода.

Температуру вещества можно определить как

$$T = \frac{h\nu_{21}}{k \ln \frac{N_1}{N_2}}. \quad (1.9)$$

При $N_2 < N_1$ — обычный тепловой режим $T > 0$.

При инверсной населенности, когда $N_2 > N_1$ температура перехода становится отрицательной величиной ($T < 0$). Следует особо отметить, что отрицательная температура является условной математической величиной, характеризующей физический процесс инверсии населенностей для данной пары уровней энергии в квантовой системе.

Таким образом, условие генерации квантовой системы может быть осуществлено при условии создания инверсной населенности уровней или создания квантовых переходов при отрицательной температуре.

Совокупность квантовых частиц с отрицательными потерями энергии (усилением) распространяющегося в этой совокупности частиц излучения называется активной средой. В этой совокупности квантовых частиц имеет место инверсия населенностей уровней, которая необходима для усиления сигнала в активной среде.

Поглощаемое квантовой системой частиц излучение нарушает в ней тепловое равновесие. В случае, когда вероятность переходов под влиянием поля накачки становится сравнимой с вероятностью релаксационных переходов, равновесное распределение населенностей уровней изменяется. В этом случае доля энергии, поглощаемой квантовой системой, уменьшается и возникает эффект насыщения. В этой ситуации коэффициент поглощения α падает

$$\alpha = (-1/I) \cdot (dI/dx) = (-1/c\rho) \cdot (d\rho/dt).$$

Здесь I — плотность мощности излучения, Вт/см²; x — направление распространения волны; ρ — плотность излучения, c — скорость света.

Для двухуровневой квантовой системы существует закон сохранения

$$n_1 + n_2 = n,$$

где n_1 — число частиц на нижнем уровне, n_2 — число частиц на верхнем уровне, n — общее число частиц.

Динамика изменения числа частиц на верхнем уровне определяется уравнением

$$dn_2/dt = -(w_{12} + 1/\tau_0)n_2 + w_{12}n_1 - W_{21}n_2 + W_{12}n_1, \quad (1.10)$$

где первый член соответствует спонтанному переходу и вероятности релаксации w_{12} частиц с верхнего уровня на нижний; второй член соответствует релаксационному заселению второго уровня с вероятностью w_{12} ; третий и четвертый члены описывают индуцированные переходы соответственно на уровень 1 и на уровень 2.

Известно, что

$$W_{21} = B_{21} \rho$$

$$W_{12} = B_{12} \rho,$$

где $\rho = 8\pi\nu^2/c^3 [h\nu/\exp(h\nu/kT) - 1]$ — плотность энергии поля излучения. Известно также, что

$$g_1 B_{12} = g_2 B_{21}.$$

Тогда получаем

$$dn_2/dt = [1/\tau + (g_1 + g_2)/g_2 B_{21} \rho / \pi \Delta\nu_{12}] n_2 + [w_{12} + 2B_{12} \rho / \pi \Delta\nu_{12}] n_1,$$

где g_1, g_2 — кратность вырождения соответствующих уровней, а

$$\tau = 1/(w_{21} + w_{12} + 1/\tau_0).$$

В стационарных условиях $dn_2/dt = 0$ и при $\rho \rightarrow \infty$ имеем

$$n_2 = g_2 n / (g_1 + g_2)$$

$$n_1 = g_1 n / (g_1 + g_2)$$

Другими словами, населенности верхнего и нижнего уровней выравниваются, и происходит полное насыщение.

Величину резонансного усиления (поглощения) определим как

$$\alpha = (n_1 / g_1 - n_2 / g_2) \cdot (g_1 \cdot 2B_{12} \cdot h\nu / c \Delta\nu_{ij})$$

или

$$z = z_0 / [1 + (g_1 + g_2) \cdot I / 2g_2 I_s], \quad (1.11)$$

где I_s — уровень насыщения, z_0 — разность населенностей в отсутствие внешнего поля.

Условие, при котором активная среда позволяет усилить проходящее через нее электромагнитное усиление, можно записать в следующем виде:

$$I = I_0 \cdot \exp[(\alpha - \beta_n)], \quad (1.12)$$

где β_n — коэффициент потерь, служащий количественной характеристикой потерь в активной среде.

Усилительные свойства активной среды можно повысить. С этой целью используют положительную обратную связь. Конструктивно положительная обратная связь реализуется путем помещения активной среды в резонатор с высокой добротностью. В этом случае усиление может превысить суммарные потери усилителя в цепи обратной связи. Произойдет самовозбуждение усилителя, и он превратится в генератор. Условием самовозбуждения лазера с резонатором типа Фабри — Перо, состоящего из двух зеркал с коэффициентами отражения r_1 и r_2 , и расстоянием между ними L будет уравнение

$$\alpha \geq \beta_n + 1 / 2L \cdot \ln[1 / (r_1 r_2)]. \quad (1.13)$$

Другими словами, коэффициент усиления равен сумме коэффициента потерь активной среды и коэффициента потерь на зеркалах.

Задачи и упражнения

1.1. Накачка четырехуровневой квантовой системы

Накачка квантовой системы может происходить по четырехуровневой схеме, например, в газовом лазере. Возбуждение осуществляется с уровня 1 на уровень 4. Инверсия населенностей образуется между средними уровнями 3 и 2. Вероятности переходов с уровня 4 на уровень 3 и с уровня 2 на уровень 1 достаточно велики, так что уровни 4 и 2 можно всегда считать пустыми (рис. 1.3). Кинетические уравнения для такой системы обычно записываются в виде

$$\frac{dn_3}{dt} = W - B_c n_3 N - \frac{n_3}{\tau};$$

$$\frac{dN}{dt} = B_c n_3 N - \frac{N}{\tau_p} + \frac{n_3}{\tau} \epsilon_c,$$

где W — число частиц, попадающих на уровень 3 за счет накачки в единицу времени; n_3 — число частиц на уровне 3 (напомним, что число частиц на уровне 2 равно $n_2 = 0$);

B_0 — константа, зависящая от частоты; N — число фотонов в типе колебаний резонатора лазера; τ — время спонтанного излучения; τ_p — время жизни фотонов в резонаторе; ε_c — вероятность того, что фотоны, образующиеся в результате спонтанного распада, попадают в рассматриваемый тип колебаний.

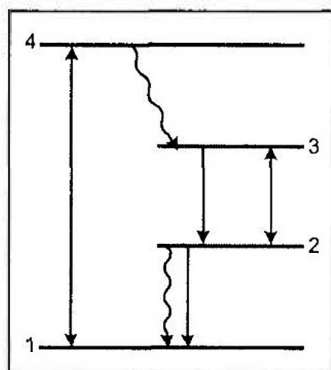


Рис. 1.3. Четырехуровневая система. Накачка происходит с уровня 1 на уровень 4. Рабочий переход осуществляется между уровнями 3 и 2. Прямые стрелками показаны излучательные переходы, а волнистыми — безызлучательные

1. Определить B_c и ε_c и показать, что $B_c = \varepsilon_c / \tau$, т. е. показать, что кинетическое уравнение для числа фотонов можно записать в виде $dN/dt = B_c n(N+1) - N/\tau_p$.
2. Написать явный вид выражений B_c и ε_c для лоренцевой формы линии люминесценции активного вещества.

Решение

Для вычисления величины B_c рассмотрим изменение числа фотонов в типе колебаний за счет индуцированных переходов. Если ρ_v — спектральная плотность энергии поля, V — объем кристалла, то изменение числа фотонов в типе колебаний в единицу времени за счет индуцированных переходов будет

$$dN/dt = \rho_v B_{32}(\nu) n_3 - \rho_v B_{23}(\nu) n_2.$$

Учтем, что $N = \frac{\rho_v V}{h\nu}$, а также соотношения между коэффициентами $B_{32}(\nu)$ и $B_{23}(\nu)$.

Тогда

$$\frac{dN}{dt} = \frac{h\nu N}{V} B_{32}(\nu) \left[n_3 - \frac{B_{23}(\nu)}{B_{32}(\nu)} n_2 \right] = \frac{h\nu N B_{32} g(\nu)}{V} \left[\frac{n_3}{\tilde{g}_3} - \frac{n_2}{\tilde{g}_2} \right] \tilde{g}_3,$$

где \tilde{g}_i — кратности вырождения уровней, а $g(\nu)$ — форм-фактор линии.

Учитывая, что

$$B_{32} \tilde{g}_3 = \frac{c^3}{8\pi h\nu^3} \tilde{g}_2 A_2$$

и вводя обозначение

$$n = \frac{n_3}{\tilde{g}_3} - \frac{n_2}{\tilde{g}_2},$$

получим

$$dN/dt = Nn \frac{c^3}{8\pi\nu^2 V} \tilde{g}_3 A_{32} g(\nu). \quad (1.1.1)$$

Для лоренцевой формы линии имеем

$$g(\nu) = \frac{1}{\pi} \cdot \frac{\Delta\nu}{(\nu_0 - \nu)^2 + (\Delta\nu)^2},$$

где $\Delta\nu$ — полуширина линии. Тогда

$$\frac{dN}{dt} = Nn \frac{c\lambda^2 \tilde{g}_3 A_{32}}{8\pi^2 \Delta\nu V} \cdot \frac{(\Delta\nu)^2}{(\nu_0 - \nu)^2 + (\Delta\nu)^2}. \quad (1.1.2)$$

Из равенства (1.1.2) видно, что

$$B_c = \frac{c\lambda^2 \tilde{g}_3 A_{32}}{8\pi^2 \Delta\nu V} \cdot \frac{(\Delta\nu)^2}{(\nu_0 - \nu)^2 + (\Delta\nu)^2} \quad (1.1.3)$$

и, в частности, для вершины линии

$$B_c = \frac{c\lambda^2 \tilde{g}_3 A_{32}}{8\pi^2 \Delta\nu V}. \quad (1.1.4)$$

Выражения (1.1.3) и (1.1.4) для константы B_c получены в предположении, что на обоих рабочих уровнях число активных частиц отлично от нуля. Согласно же условиям задачи $n_2 = 0$, т. е. $n = n_3 \tilde{g}_3$. Подставляя это выражение в (1.1.1), получаем вместо (1.1.3) и (1.1.4)

$$B_c = \frac{c\lambda^2 A_{32}}{8\pi^2 \Delta\nu V} \cdot \frac{(\Delta\nu)^2}{(\nu_0 - \nu)^2 + (\Delta\nu)^2}; \quad (1.1.3a)$$

$$B_c = \frac{c\lambda^2 A_{32}}{8\pi^2 \Delta\nu V}. \quad (1.1.4a)$$

Отметим, что в формулы (1.1.3) и (1.1.4) входит длина волны излучения в веществе. Если n_r — показатель преломления кристалла, то

$$\lambda = \lambda_0 / n_r,$$

где λ_0 — длина волны излучения в вакууме.

Вычислим теперь независимым образом ϵ_c . Предположим, что все поверхности лазерного кристалла идеально отражающие, т. е. сам кристалл представляет собой резонатор без потерь. Тогда в частотном интервале от ν до $\nu + d\nu$ в кристалле с показателем преломления n_r возбуждается число типов колебаний:

$$\rho_T(\nu) = \frac{8\pi\nu^2 V n_r^3}{c^3} d\nu,$$

где V — объем кристалла.

Число спонтанных распадов уровня 3 в единицу времени составляет n_3/τ , причем частное распределение излучаемых фотонов определяется формой линии люминесценции активного вещества, и в частотный интервал от ν до $\nu + d\nu$ попадает $n_3 g(\nu) d\nu / \tau$ фотонов.

Очевидно, число фотонов, попадающих в один тип колебаний, равно

$$\frac{\frac{n_3}{\tau} g(\nu) d\nu}{\rho_f(\nu) d\nu} = \frac{n_3}{\tau} \varepsilon_c.$$

Следовательно,

$$\varepsilon_c = g(\nu) \frac{c^3}{8\pi\nu^2 V n_r^2}.$$

Для лоренцевой формы линии люминесценции

$$\varepsilon_c = \frac{c^3}{\Delta\nu 8\pi^2 \nu^2 V n_r^2} \cdot \frac{(\Delta\nu)^2}{(\nu - \nu_0)^2 + (\Delta\nu)^2}.$$

Учитывая, что $\frac{c^2}{\nu^2 n_r^2} = \lambda^2$ и то, что для рассматриваемой четырехуровневой системы

$A_{32} = 1/\tau$, получаем

$$\frac{\varepsilon_c}{\tau} = \frac{c\lambda^2 A_{32}}{8\pi^2 \Delta\nu V} \frac{(\Delta\nu)^2}{(\nu_0 - \nu)^2 + (\Delta\nu)^2},$$

т. е. выражение ε_c/τ совпадает с выражением (1.1.3а). Это позволяет записать кинетическое уравнение для числа фотонов в виде

$$\frac{dN}{dt} = B_c n(N+1) - \frac{N}{\tau_p}. \quad (1.1.5)$$

1.2. Накачка лазерного кристалла

Рассмотреть накачку лазерного кристалла источником интенсивностью $I_0(\lambda_n)$ в единичном интервале длин волн. Кристалл выполнен в виде прямоугольного параллелепипеда длиной L с квадратным сечением, длина которого равна a (рис. 1.4). Зеркала нанесены прямо на грани кристалла (штриховка).

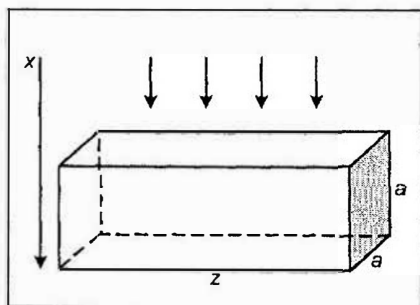


Рис. 1.4. Схема накачки лазерного кристалла

Излучение источника накачки падает перпендикулярно к грани кристалла в направлении x . Коэффициент поглощения излучения в кристалле $k(\lambda_n)$, квантовая эффективность $\eta(\lambda_n)$. Источник накачки считать достаточно узкополосным, т. е. в пределах ширины спек-

тра источника величины $k(\lambda_n)$, $\eta(\lambda_n)$ и λ_n можно считать постоянным: $k(\lambda_n) = k_p$; $\eta(\lambda_n) = \eta_p$; $\lambda_n = \lambda_p$.

1. Оценить пороговую мощность источника накачки, считая, что поглощение в кристалле невелико и не влияет на интенсивность проходящего через кристалл излучения.
2. Учтя изменение интенсивности сигнала накачки за счет поглощения в кристалле, определить предельную толщину кристалла, для которой накачка еще эффективна.

В этом случае считать, что $I_0 = \int I(\lambda_n) d\lambda_n = 26,8 \text{ Вт/см}^2$, где интегрирование происходит по всему спектру источника накачки.

3. Кроме того, при оценках использовать следующие данные: $n_r = 1,4$, $n = 3 \times 10^{18} \text{ см}^{-3}$, $k_p = 1 \text{ см}^{-1}$, $\lambda_p = 0,64 \text{ мкм}$, $\eta_p \approx 1$. Ширина линии люминесценции (при $T = 290 \text{ К}$) составляет $2\Delta\nu = 1,1 \text{ см}^{-1}$, длина параллелепипеда $l = 1,2 \text{ см}$, а сторона квадрата $a = 0,1 \text{ см}$, коэффициент отражения зеркал $r_{\text{отр}} = 0,9$, длина волны излучения $\lambda = 0,7 \text{ мкм}$.

Решение

Оценим мощность источника накачки, поглощаемую в единице объема кристалла.

За счет поглощения в кристалле интенсивность падающего на грань кристалла излучения источника накачки уменьшается по закону

$$I(x, \lambda_n) = I_0(0, \lambda_n) \exp[-k(\lambda_n)x]. \quad (1.2.1)$$

Изменение интенсивности на элементе пути dx в плоскости x_1 кристалла равно

$$\begin{aligned} I(x_1, \lambda_n) - I(x_1 + dx, \lambda_n) &= I_0(0, \lambda_n) \cdot \{ \exp[-k(\lambda_n)x_1] - \exp[-k(\lambda_n)(x_1 + dx)] \} = \\ &= I(x_1, \lambda_n) k(\lambda_n) dx. \end{aligned}$$

Таким образом, мощность, поглощаемая в единице объема и интервале длин волн от λ до $\lambda + d\lambda$, равна

$$I(x_1, \lambda_n) k(\lambda_n) d\lambda_n \approx I(\lambda_n) k(\lambda_n) d\lambda_n.$$

Из этого количества только η -я часть идет на создание инверсной населенности, т. е. полезная мощность равна

$$I(x_1, \lambda_n) k(\lambda_n) \eta(\lambda_n) d\lambda_n.$$

Энергия поглощенного фотона $h\nu_n = hc / \lambda_n$.

Считая, что один поглощенный фотон приводит к появлению одной возбужденной частицы на верхнем рабочем уровне, получаем, что число n' частиц, возбуждаемых в единице объема в единицу времени за счет поглощения излучения во всей ширине источника, равно

$$n' = \int d\lambda_n \frac{I(\lambda_n) k(\lambda_n) \eta(\lambda_n)}{hc / \lambda_n}.$$

Поскольку $k(\lambda_n)$, $\eta(\lambda_n)$, λ_n в пределах спектра излучения источника мало изменяются, выносим их из-под знака интеграла, получая

$$n' = \frac{k_p \lambda_p \eta_p}{hc} \int I(\lambda_n) d\lambda_n = \frac{k_p \lambda_p \eta_p}{hc} I,$$

где I — интегральная интенсивность по всему спектру источника накачки:

$$I = \int I(\lambda_n) d\lambda_n.$$

Число частиц в единице объема, уходящих с уровня в единицу времени за счет спонтанных переходов на пороге генерации, равно $n_{\text{пор}} / \tau_{\text{сп}}$ (рассматривается пороговое значение накачки). Очевидно, должно выполняться равенство $n' = n_{\text{пор}} / \tau_{\text{сп}}$ и, следовательно,

$$\frac{n_{\text{пор}}}{\tau_{\text{сп}}} = \frac{k_p \lambda_p \eta_p}{hc} I_{\text{пор}}.$$

т. е. пороговая интенсивность

$$I_{\text{пор}} = \frac{n_{\text{пор}} hc}{k_p \lambda_p \eta_p \tau_{\text{сп}}}, \quad (1.2.2)$$

При отсутствии вырождения для лоренцевой формы линии

$$n_{\text{пор}} = \frac{8\pi^2 \Delta \nu (1 - r_{\text{отпр}})}{\lambda^2 A_{32} L}. \quad (1.2.3)$$

Подставляя (1.2.3) в (1.2.2) и учитывая, что $A_{32} = 1 / \tau_{\text{сп}}$, а длина волны излучения в кристалле λ связана с длиной волны излучения в вакууме λ_0 соотношением

$$\lambda = \lambda_0 / n_r,$$

получим

$$I_{\text{пор}} = \frac{8\pi^2 \Delta \nu (1 - r_{\text{отпр}}) hc n_r^2}{\lambda_0^2 k_p \lambda_p \eta_p L}. \quad (1.2.4)$$

Для исходных данных задачи это дает $I_{\text{пор}} = 13,4 \text{ Вт} / \text{см}^2$. Поверхность кристалла, на которую падает поток $S = La = 0,12 \text{ см}^2$, т. е. мощность падающего на кристалл излучения $P = S I_{\text{пор}} \approx 1,2 \text{ Вт}$.

Критическая толщина a кристалла определится из условия, что при проникновении в глубину кристалла на расстояние a интенсивность падает настолько, что инверсная населенность становится меньше пороговой $n < n_{\text{пор}}$.

Если I_0 — начальная интенсивность, то a определится из условия

$$I_{\text{пор}} = I_0 \exp(-k_p a),$$

и $I_{\text{пор}}$ находят по формуле (1.2.4). Таким образом,

$$a = \frac{1}{k_p} \ln \frac{I_0}{I_{\text{пор}}}.$$

Для $I_0 = 26,8 \text{ Вт} / \text{см}^2$, $I_{\text{пор}} = 13,4 \text{ Вт} / \text{см}^2$, используя исходные данные, получаем $a \approx 0,7 \text{ см}$.

3. Добротность резонатора

Рассчитать добротность Q_r и время жизни фронта в резонаторе Фабри — Перо с плоскими зеркалами. Расстояние между зеркалами $L = 1 \text{ м}$. В резонаторе возбуждается один основной тип колебаний TEM_{00q} , образуемый двумя бегущими навстречу друг другу плоскими волнами ($\lambda = 0,6 \text{ мкм}$). Среда, заполняющая резонатор, слабо поглощающая (коэффициент поглощения $\beta = 0,001 \text{ см}^{-1}$). Эти потери могут быть связаны с процессами рассеяния

в среде, нерезонансного поглощения и т. д. Коэффициент отражения каждого из зеркал $r_{\text{отр}} = 95\%$. Диаметр зеркал много больше диаметра светового пучка, так что дифракционными потерями можно пренебречь.

Решение

Получим формулу для добротности типа колебаний открытого резонатора, учитывающую потери в зеркалах и в среде, заполняющей резонатор.

По определению добротности, если u_{TK} — энергия типа колебаний в резонаторе, то

$$u_{TK} = u_{TK}^0 \exp\left(-\frac{\omega}{Q_p} t\right), \quad (1.3.1)$$

где u_{TK}^0 — энергия типа колебаний в момент $t = 0$.

Из (1.3.1) следует, что энергия, теряемая за время dt , равна

$$du_{TK} = -u_{TK} \frac{\omega}{Q} dt. \quad (1.3.1a)$$

Рассматриваемый тип колебаний, как две волны, бегущих навстречу друг другу, причем энергия каждой из них равна $u_{TK}/2$. Потери каждой из волн на проход за счет поглощения в среде ($\beta L \ll 1$)

$$\frac{u_{TK}}{2} \exp(-\beta L) - \frac{u_{TK}}{2} = \frac{u_{TK}}{2} (1 - \beta L) - \frac{u_{TK}}{2} = -\frac{u_{TK}}{2} \beta L. \quad (1.3.2)$$

Потери для этой же волны на проход за счет пропускания зеркала

$$-(u_{TK}/2)(1 - r_{\text{отр}}). \quad (1.3.2a)$$

Полные потери за один проход (т. е. за время $t_n = L/c$) каждой из волн определяются суммой (1.3.2) и (1.3.2a). Энергия, теряемая каждой волной в среднем в единицу времени

$$-\frac{u_{TK}}{2} [\beta L + (1 - r_{\text{отр}})] \frac{1}{L/c}, \quad (1.3.3)$$

а обими волнами за время dt

$$-u_{TK} [\beta L + (1 - r_{\text{отр}})] \frac{c}{L} dt. \quad (1.3.4)$$

С другой стороны, выражение (1.3.4) должно совпадать с выражением (1.3.1a). Приравняв их, получаем

$$Q_p = \frac{\omega L}{c [\beta L + (1 - r_{\text{отр}})]}, \quad (1.3.5)$$

$$\tau_p = Q_p / \omega.$$

Подставляя значения параметров из условий задачи, находим $Q_p = 7 \times 10^7$, $\tau_p \approx 2 \times 10^{-8}$ с.

Контрольные вопросы

1. Что такое квантовая электроника?
2. Как вы понимаете спонтанный квантовый переход?
3. Как вы понимаете вынужденный квантовый переход?
4. Что такое вероятность спонтанного испускания и как ее выразить?
5. Что такое вероятность вынужденного испускания и как ее выразить?
6. Что такое разрешенные и запрещенные квантовые переходы?
7. Как определяется Лоренцева форма линии?
8. Какие факторы влияют на уширение спектральной линии?
9. Опишите процесс оптической накачки.
10. Что собой представляет активная квантовая среда?
11. Как вы себе представляете отрицательную температуру в квантовой системе?

Рекомендуемая литература

1. Карлов Н. В. Лекции по квантовой электронике. Учебное пособие. — М.: Наука, 1983.
2. Пихтин А. Н. Физические основы квантовой электроники и оптоэлектроники. Учебное пособие. — М.: Высшая школа, 1983.
3. Страховский Г. М., Успенский А. В. Основы квантовой электроники. Учебное пособие. — М.: Высшая школа, 1979.
4. Успенский А. В. Сборник задач по квантовой электронике. Учебное пособие. — М.: Высшая школа, 1976.

2. Принципы работы лазера

Лазеры или оптические квантовые генераторы представляют собой источник когерентного электромагнитного излучения оптического или близкого к нему диапазона, действие которого основано на использовании вынужденного излучения атомов или молекул.

Слово лазер происходит от английского *laser* (Light Amplification by Stimulated Emission of Radiation, усиление света вынужденным излучением). Если во всех традиционных источниках света используется спонтанное излучение системы возбужденных атомов, то в лазерах все атомы когерентно излучают кванты света, тождественно равные между собой по частоте, направлению распространения, поляризации.

Эйнштейн открыл явление вынужденного излучения семь десятилетий назад, однако лазеры сразу не появились. Этого не случилось и половину века спустя. Их сделали только тогда, когда стало ясно, что любой усилитель в радиоэлектронике входит в режим генерации, если в системе создается положительная обратная связь.

В тот момент, когда впервые догадались поместить возбужденную квантовую систему с инверсной населенностью уровней в резонатор Фабри — Леро, создающий положительную обратную связь в системе, родился лазер. Первый лазер был создан в 1960 году. В качестве активного вещества использовался кристалл рубина.

В основе работы лазеров лежит явление вынужденного излучения под действием внешнего электромагнитного поля, усиление и формирование потока излучения.

Энергия возбуждения квантовой системы осуществляется путем накачки — импульсного или постоянного воздействия на активную среду электромагнитным излучением определенной частоты.

Возбуждение активной среды может осуществляться по трех- или четырехуровневой схеме (рис. 2.1).

Примером трехуровневой схемы является рубин $\text{Al}_2\text{O}_3 + \text{Cr}^{3+}$. Содержание атомов хрома составляет ~0,05%. Красный цвет кристалла рубина обусловлен расположением энергетических уровней атомов хрома в рубине. Возбуждение атомов хрома в рубине осуществляется за счет оптической накачки. Возбужденные атомы хрома переходят в полосу поглощения E_3 , их время жизни там составляет $\sim 10^{-3}$ с. В атомных масштабах соотношение времен 10^{-8} с и 10^{-3} с является вечностью и поэтому на метастабильном уровне накапливается большое число (больше половины всех) возбужденных атомов.

Метастабильный уровень становится населенным, и в квантовой системе возникает инверсия населенностей уровней по отношению к основному уровню E_1 . На практике накачка рубина осуществляется лампой, обеспечивающей вспышку белого света. Энергетические полосы хрома в рубине позволяют использовать до 15% света лампы-вспышки. При содержании атомов хрома в рубине 0,05% концентрация хрома составляет примерно 10^{19} атомов / см^3 . Обычно возбуждается их половина, т. е. 5×10^{18} см^{-3} . Наличие фотона с частотой $\nu = \frac{E_2 - E_1}{h}$ позволяет стимулировать лазерную генерацию, заключающуюся

вынужденном излучении при переходе ансамбля возбужденных атомов хрома с уровня E_2 на уровень E_1 . Лазерный переход $E_2 \rightarrow E_1$ позволяет генерировать излучение с частотой

$$\nu = \frac{E_2 - E_1}{h} \quad (2.1)$$

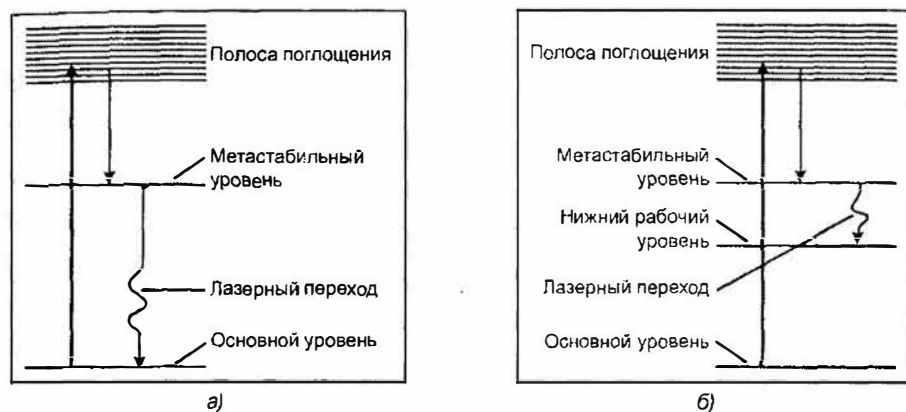


Рис. 2.1. Процессы инверсии населенности и генерации в трех- и четырехуровневых системах

существуют лазеры, работающие по четырехуровневой схеме (рис. 2.1, б). Лазер на стекле с примесью Nd^{+++} или на основе кристалла алюмоиттриевого граната с ионами Nd^{+++} является примером четырехуровневой схемы.

В энергетическом спектре такого типа лазеров между метабильным уровнем E_3 и основным уровнем E_1 имеется уровень E_2 с небольшой заселенностью возбужденными атомами. Лазерный переход между уровнями E_3 и E_2 позволяет генерировать излучение частотой $\nu = \frac{E_3 - E_2}{h}$. По четырехуровневой схеме работает также ряд лазеров на газовых средах.

Важнейшая обратная связь осуществляется с помощью двух зеркал, образующих интерферометр Фабри — Перо, который и является оптическим открытым резонатором.

В резонаторе могут возбуждаться колебания только определенной длины волны и определенной структуры. Частоты этих колебаний называются *резонансными* и соответствуют собственным частотам резонатора. Возникающие колебания имеют определенную структуру или моду, которая соответствует собственному типу колебаний резонатора (от латинского *modus* — мера, способ). *Мода* может быть интерпретирована как стационарная конфигурация электромагнитного поля, электрическая составляющая которой может быть описана в виде:

$$\vec{E}(r, t) = \vec{E}_0 U(r) \exp(i\omega t), \quad (2.2)$$

где ω — собственная частота резонатора.

Резонатор характеризуется *добротностью* Q , которая представляет собой отношение запасенной в резонаторе энергии к средней энергии за 0.5π периода колебаний. Добротность можно выразить через эффективное время жизни фотона $\tau_{\text{эфф}}$ в мод

$$Q = \omega \cdot \tau_{\text{эфф}}$$

Электромагнитные волны, распространяющиеся вдоль оптической оси резонатора, отражаются и интерферируют между собой. Коэффициент отражения зеркал достаточно высок, так что в пространстве резонатора могут существовать электромагнитные колебания строго определенной длины волны. На поверхности идеально отражающих зеркал ($\sim 100\%$) амплитуда световых колебаний должна быть равна нулю. При этом фаза отраженной волны совпадает с фазой падающей только в том случае, когда длина резонатора

$$L = n \frac{\lambda}{2},$$

где n — целое число (рис. 2.2). Такая система зеркал называется открытым резонатором, который резонирует на собственных частотах $\nu = \frac{c}{2L}$. В действительности существуют

не резонансные линии, а резонансные полосы шириной $\Delta\nu_p = \frac{c}{2L} \approx 1,5 \times 10^{10}$ Гц для резонатора длиной в 1 м.

В пределах спектральной линии активной лазерной среды укладывается от десятка до нескольких тысяч собственных колебаний резонатора (рис. 2.3, а). Такой резонатор называется многочастотным. Спектр собственных частот лазера определяется собственными частотами резонатора, лежащими вблизи максимума спектральной линии.

Одновременно излучение лазера характеризуется поперечным распределением электромагнитного поля — *поперечными модами*. Эти колебания называются трансверсальными электромагнитными колебаниями и обозначаются TEM_{mnq} , где индекс q указывает число полуволн на длине резонатора, а индексы m и n характеризуют число изменений направления поля вдоль осей x и y , соответственно. На рис. 2.3, б приведены фотографии трансверсальных мод на зеркалах лазера. Так как величина индекса q значительно больше индексов m и n , то индекс q обычно опускается. Мода TEM_{00} является аксиальной. Остальные колебания представляют собой неаксиальные моды.

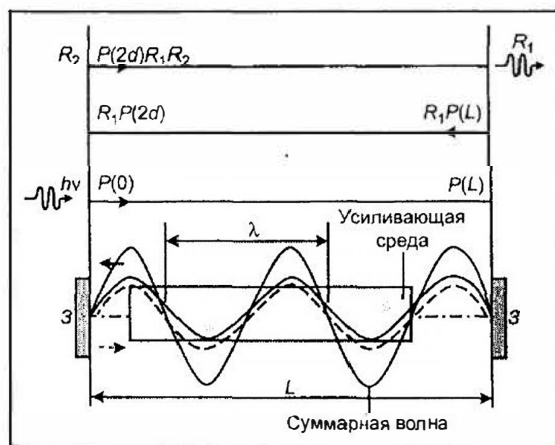


Рис. 2.2. Активное вещество в плоском резонаторе и возникновение интерференционного усиления волн: З — зеркала; L — длина резонатора; λ — длина волны усиляемого излучения (масштаб не соблюден)

Для получения сверхкоротких лазерных импульсов используют метод синхронизации мод. С этой целью в импульсных лазерах синхронизация мод осуществляется с помощью

помещаемого в резонатор нелинейного фильтра, который просветляется под воздействием излучения. В лазерах непрерывного действия синхронизация мод осуществляется путем модуляции энергетических потерь или фазы поля излучения на частоте, близкой к частоте межмодовых биений. В режиме синхронизации мод лазер излучает периодическую последовательность сверхкоротких импульсов с частотой, близкой к межмодовой частоте (100—500 мГц). Длительность импульсов излучения в этом режиме равна обратной ширине спектра генерируемых мод. Длительность импульсов в режиме синхронизации мод может достигать значений 10^{-12} — 10^{-13} с, что позволяет резко повысить пиковую мощность излучения.

Лазерные зеркала обеспечивают возможность многократного прохода плоской волны, если волновой вектор направлен по оси интерферометра. Многократное прохождение в резонаторе световой волны обеспечивает ее усиление путем многократного "опустошения" метастабильного уровня и генерации фотонов. Инвертированная активная среда при каждом проходе будет усиливать волну и повышать плотность фотонов.

Если $P(0)$ — мощность аксиальной моды первоначального спонтанного излучения, а $P(2L)$ — мощность после двойного прохода резонатора, то при коэффициентах отражения зеркал R_1 и R_2 , коэффициенте усиления среды γ и коэффициенте поглощения α можно записать:

$$P(2L) = P(0) \exp(2(\beta - \alpha)R_1R_2).$$

Самовозбуждение наступает при условии $P(2L) > P(0)$ или

$$\beta > \alpha + \frac{1}{2L} \ln \frac{1}{R_1R_2}. \quad (\text{см. формулу 1.13}).$$

Обычно одно зеркало делают глухим, т. е. $R_1 \approx 100\%$, а второе $R_2 \approx 95\%$.

Накопленное в резонаторе излучение, преодолев определенный порог интенсивности, выходит из зеркала с коэффициентом отражения R_2 .

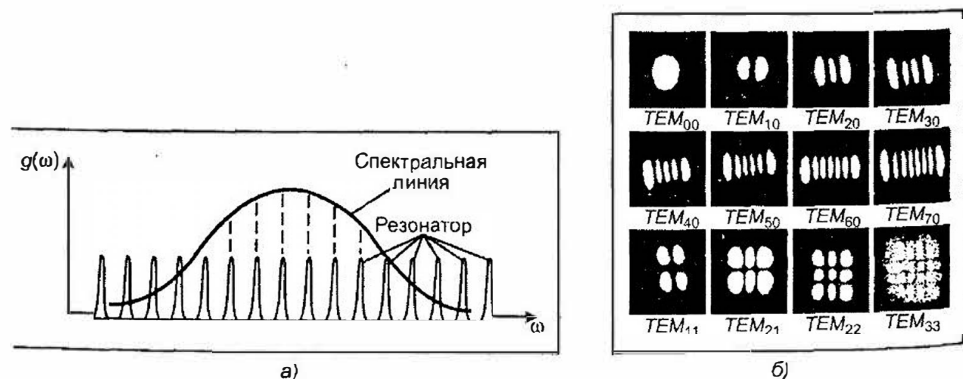


Рис. 2.3. Продольные собственные частоты на фоне спектральной линии (а) и трансверсальные моды лазера (б)

Сформированный таким образом пучок лазерного излучения является когерентным излучением с высокой спектральной плотностью излучения.

Гармоническое колебание называется *монохроматическим*, если оно может быть описано выражением

$$A(t) = A_0 \sin(2\pi\nu_0 t + \varphi_0), \quad (2.3)$$

где $A(t)$ — текущее значение амплитуды, A_0 — максимальное значение амплитуды, $2\pi\nu_0$ — круговая частота, φ_0 — начальная фаза колебаний.

Ширина спектра $\Delta\nu$ излучения определяется степенью монохроматичности излучения

$$\mu = \Delta\nu / \nu_0,$$

где ν_0 — центральная частота.

При $\mu \ll 1$ излучение называют *квазимонохроматическим*. Лазеры позволяют получить излучение со значением $\mu \approx 10^{-10}$ при достаточно большой мощности.

Понятие монохроматичности тесно связано с понятием когерентности.

Когерентность света представляет собой взаимную согласованность протекания во времени световых колебаний в разных точках пространства и (или) времени, которая характеризует их способность к интерференции.

Различают *пространственную* и *временную* когерентность. Пространственная когерентность связывается с корреляцией фазы колебаний в разных точках пространства в один и тот же момент времени. При сложении когерентных колебаний возникает устойчивая интерференционная картина.

Корреляцию колебаний в определенной точке пространства можно наблюдать только в определенном интервале времени. Этот интервал времени принято называть *временем когерентности*. Время когерентности обычно принимается за время жизни излученного колебания τ . Расстояние, проходимое светом за время когерентности, называют *длиной когерентности* L . При $\tau \approx 10^{-8}$ с длина когерентности $L = c \cdot \tau = 300$ см. Если учитывать, что ширина спектральной линии связана со временем жизни, то $\Delta\nu \approx 1/\tau$. В этом случае длина когерентности связана с шириной спектральной линии величиной $L \approx c/\Delta\nu$. Таким образом, чем уже частотный спектр излучения, тем больше время когерентности и выше степень временной когерентности, и лучше монохроматичность излучения.

Лазер представляет собой уникальный источник оптического излучения и здесь будет уместно отметить его особенности.

Пространственная когерентность характеризует форму волнового фронта излучения. Лазерное излучение имеет высокую направленность, обусловленную свойствами оптического резонатора, и высокую спектральную мощность излучения.

Описание законов распространения лазерного излучения в свободном пространстве и в оптических системах производится с помощью пространственных параметров.

К ним относятся диаметр пучка и его расходимость, диаграмма направленности, распределение плотности мощности (энергии) в поперечном сечении пучка, ось диаграммы направленности, ближняя и дальняя зоны лазерного излучения.

Под *диаметром пучка лазерного излучения* понимается диаметр поперечного сечения канала, внутри которого распространяется энергия лазерного излучения. Диаметр пучка на выходном зеркале плоского резонатора, как правило, определяется диаметром активного элемента. Если активный элемент имеет прямоугольное сечение, то размеры пучка определяются размерами сечения активного элемента. Размер поперечного сечения пучка, естественно, несколько меньше соответствующих размеров активного элемента, т. е.

электромагнитное поле спадает к его краям. Конкретное значение размера поперечного сечения пучка зависит от размеров активного элемента, модового состава излучения и выбранного уровня энергии в пучке.

Расходимость лазерного излучения — это плоский или телесный угол, характеризующий ширину диаграммы направленности лазерного излучения. Эту расходимость называют угловой. Существует также понятие энергетической расходимости, которое представляет собой телесный угол, внутри которого распространяется заданная доля энергии.

Диаграмма направленности лазерного излучения — это угловое распределение энергии или мощности лазерного излучения. Ось диаграммы направленности лазерного излучения представляет собой прямую, проходящую через максимум углового распределения энергии или мощности лазерного излучения. При симметричном распределении поля ось диаграммы направленности совпадает с энергетической осью пучка.

Дальняя зона лазерного излучения представляет собой область пространства вдоль оси лазерного пучка, которая располагается на таком расстоянии от излучателя лазера, начиная с которого диаграмма направленности остается постоянной. Диаграмма направленности носит дифракционный характер независимо от того, ограничен лазерный пучок реальной диафрагмой или нет. Известно, что при описании дифракции пользуются понятиями зоны Френеля и зоны Фраунгофера. Дальняя зона лазерного излучения соответствует зоне Фраунгофера.

Распределение плотности мощности излучения можно получить из измерений, либо могут быть рассчитаны по известным параметрам резонатора. Связь параметров пучка с параметрами резонатора определяется типом резонатора.

Задачи и упражнения

2.1. Оптимальный коэффициент пропускания зеркал

Определить оптимальный коэффициент пропускания зеркал T_p (зеркала одинаковые) резонатора лазера, при котором мощность будет максимальной. Коэффициент ненасыщенного усиления на проход g_0 , коэффициент потерь на проход β . Длина резонатора L . Дифракционными потерями пренебречь. Для численных оценок считать: $L = 10$ см, $g_0 = 0,1$ см⁻¹, $\beta = 0,01$ см⁻¹. Активная среда заполняет весь резонатор.

Решение

Стоячую волну в резонаторе лазера можно рассматривать как суперпозицию двух бегущих волн. Пусть каждая из волн характеризуется интенсивностью I . Выходная мощность P через зеркало лазерного резонатора равна

$$P = T_p I. \quad (2.1.1)$$

В дальнейшем удобнее будет характеризовать потери за счет пропускания зеркал, введя величину $g_s = T_p / L$ — потери на проход за счет пропускания зеркал. В лазере коэффициент усиления из-за насыщения имеет вид

$$g = \frac{g_0}{1 + I/I_0},$$

где I_0 — насыщающая интенсивность.

При генерации потери на проход плюс потери на зеркалах должны компенсироваться усилением на проход, т. е. должно выполняться равенство:

$$\frac{g_0}{1 + I/I_0} = g_s + \beta. \quad (2.1.2)$$

Выражая из (2.1.2) I и подставляя в (2.1.1), имеем:

$$P = I_0 L g_s \left(\frac{g_0}{g_s + \beta} - 1 \right). \quad (2.1.3)$$

Оптимальный коэффициент пропускания зеркал определится из условия нахождения экстремума выражения (2.1.3), т. е.

$$\partial P / \partial g_s = 0.$$

Отсюда

$$T_{p \text{ опт}} = L g_{\text{зонт}} = L(\sqrt{g_0 \beta} - \beta).$$

Подстановка численных значений дает $T_{p \text{ опт}} = 23\%$.

2.2. Моды колебаний

Оценить, насколько частота типа колебаний TEM_{01} отличается от частоты основного типа колебаний TEM_{00} для пустого резонатора. Резонатор образован плоским и сферическим (радиус кривизны $R \approx 100$ см) зеркалами. Длина резонатора $L = 50$ см.

Решение

Собственные частоты пустого резонатора определяются выражением

$$\nu_{mq} = \nu_0 \left[q + \frac{m+n+1}{\pi} \arccos \sqrt{g_1 g_2} \right],$$

где m, n, q — целые числа; $\nu_0 = c/2L$, $g_{1,2} = 1 - L/R_{1,2}$ (L — длина резонатора, R_1, R_2 — радиусы кривизны зеркал).

Следовательно,

$$\nu_{mn} - \nu_{00} = \nu_0(m+n) \frac{1}{\pi} \arccos \sqrt{g_1 g_2}.$$

Одно из зеркал резонатора плоское, т. е. его $R = \infty$ и $g = 1 - L/R = 1$; для другого зеркала $g = 0,5$.

Для типа колебаний с индексами $m = 0, n = 1$ имеем

$$\nu_{01} - \nu_{00} = \frac{c}{2L} \cdot \frac{1}{\pi} \arccos \sqrt{g_1 g_2} = \frac{3 \cdot 10^{10}}{2 \cdot 50 \cdot 3,14} \arccos \sqrt{0,5} = 74,5 \text{ МГц}.$$

2.3. Определение диаметра пучка

Два сферических зеркала с радиусами кривизны R_1 и R_2 расположены на расстоянии d одно от другого.

Найти минимальный размер пятна светового пучка в резонаторе, его положения и размеры пятен на зеркалах, если длина волны излучения λ . Апертурный размер зеркал доста-

точно велик, так что дифракционными потерями можно пренебречь. Для численных оценок взять: $R_1 = 106$ см, $R_2 = -109$ см, $d = 99$ см, $\lambda = 3,39$ мкм.

Решение

Радиус кривизны сферической поверхности, представляющей поверхность постоянной фазы, т. е. поверхности, определяющей возможные положения зеркал, равен

$$R(z) = -(z^2 + z_0^2)/z,$$

где z_0 — параметр, определяющий минимальный размер пятна луча в резонаторе

$$\rho_0(\rho_0 = \sqrt{\frac{\lambda z_0}{\pi}}).$$

Находим параметр z_0^2 . Имеем

$$R_1 = -z_1 - \frac{z_0^2}{z_1}, \quad (2.3.1)$$

$$R_2 = -z_2 - \frac{z_0^2}{z_2},$$

причем $z_2 - z_1 = d > 0$.

Решения для z_1 и z_2 имеют вид:

$$z_1 = -R_1/2 \pm \sqrt{(R_1/2)^2 - z_0^2}, \quad (2.3.2)$$

$$z_2 = -R_2/2 \pm \sqrt{(R_2/2)^2 - z_0^2}.$$

Вычитая из (2.3.2) равенство (2.3.1) и учитывая, что $z_1 - z_2 = d$, возводим результат вычитания дважды в квадрат. Тогда будем иметь

$$\begin{aligned} z_0^2 [-4d^2 + 4d(R_1 - R_2) - (R_1 - R_2)^2] = \\ = d(-d + R_1 - R_2) \cdot [R_1 R_2 + d(-d + R_1 - R_2)], \end{aligned}$$

или

$$z_0^2 = \frac{d(R_1 - R_2 - d)(R_1 - d)(-R_2 - d)}{[(R_1 - R_2) - 2d]^2}. \quad (2.3.3)$$

Подстановка исходных данных дает $z_0 = 52,8$ см. Минимальный размер пятна луча в резонаторе

$$\rho_0 = \sqrt{\frac{\lambda z_0}{\pi}} \approx 0,75 \text{ мм.}$$

Подставляя в формулы (2.3.2) полученное значение z_0 , находим $z_1 \approx 52,1$ см, $z_2 \approx 40,8$ см. Таким образом, минимальный размер пятна расположен в плоскости, отстоящей от одного зеркала примерно на 41, а от другого на 58 см.

Размеры пятен на зеркалах:

$$\rho_{1,2} = \rho_0 \sqrt{1 + (z_{1,2}/z_0)^2},$$

т. е. $\rho_1 = 1,11$ мм, $\rho_2 = 0,95$ мм.

2.4. Определение угла расхождения пучка

Оценить угол расхождения пучка основного типа колебаний конфокального резонатора. Для оценок принять $\lambda = 1$ мкм, расстояние между зеркалами $d = R_1 = -R_2 = 2$ м. Апертурный размер зеркал велик и дифракционные эффекты пренебрежимо малы.

Решение

Угол расхождения θ пучка основного типа колебаний определяется по формуле

$$\theta = \frac{\lambda}{\pi \rho_0}, \quad (2.4.1)$$

где ρ_0 — минимальный размер пятна луча в резонаторе.

Формула (2.3.3) задачи 2.3 дает для конфокального резонатора

$$d = R_1 = -R_2,$$

$$z_0 = d/2,$$

следовательно,

$$\rho_0 = \sqrt{\frac{\lambda z_0}{\pi}} = \sqrt{\frac{\lambda d}{2\pi}}.$$

Согласно уравнению (2.4.1)

$$\theta = \frac{2\lambda}{\pi d} \approx 5,6 \times 10^{-1} \text{ рад} \approx 1,9 \text{ угл. мин.}$$

2.5. Условие самовозбуждения

Резонатор оптического квантового генератора образован зеркалами с коэффициентами отражения $r_{\text{отр}1} = r_{\text{отр}2} = r = 0,5$, расположенными на расстоянии L друг от друга. Активная среда занимает все пространство между зеркалами.

Как нужно изменить коэффициент квантового усиления активной среды для выполнения условия самовозбуждения генератора, если в резонатор вносится поглотитель, поглощающий 50% падающего на него излучения?

В расчете не учитывать дифракционные потери на зеркалах и потери излучения в материале активной среды и зеркал. Для простоты считать, что при введении поглотителя размеры активной среды остаются неизменными. Толщиной поглотителя пренебречь.

Решение

Для получения условия самовозбуждения рассмотрим распространение волны от зеркала 1 до зеркала 2 и в обратном направлении и получим условие того, что волна остается самоподдерживающейся. Это и будет условием самовозбуждения генератора. Пусть от зеркала 1 к зеркалу 2 начинает распространяться волна с интенсивностью I_0 .

Если поглотитель расположен на расстоянии L_1 от зеркала 1, то до поглотителя дойдет волна интенсивностью

$$I_n = \exp(gL_1),$$

где g — коэффициент квантового усиления активной среды. Если α определяет долю поглощаемой поглотителем интенсивности, то после поглотителя интенсивность волны будет

$$I_0(1-\alpha)\exp(gL_1).$$

Далее волна опять усиливается в среде, и на зеркало 2 придет волна интенсивностью

$$I_0(1-\alpha)\exp(gL_1)\exp(gL_2) = I_0(1-\alpha)\exp(gL).$$

После отражения от зеркала 2 по направлению к зеркалу 1 будет распространяться волна с интенсивностью

$$rI_0(1-\alpha)\exp(gL).$$

На обратном пути к зеркалу 1 она испытывает усиление в активной среде и поглощение в поглотителе, и после отражения от зеркала 1 интенсивность волны составит

$$r^2I_0(1-\alpha)^2\exp(2gL).$$

Условие существования в резонаторе самоподдерживающейся волны получается, если приравнять интенсивность исходной волны I_0 и волны, совершившей обход резонатора:

$$r^2(1-\alpha)^2\exp(2gL) = 1,$$

откуда условие для порогового коэффициента усиления имеет вид:

$$gL = \frac{1}{2} \ln \frac{1}{r^2(1-\alpha)^2} = \ln \frac{1}{r(1-\alpha)}.$$

При отсутствии поглотителя, т. е. $\alpha = 0$

$$g_0L = \ln \frac{1}{r}.$$

Очевидно, соотношение пороговых коэффициентов усиления для среды без поглотителя и с поглотителем будет

$$g/g_0 = \ln \frac{1}{r(1-\alpha)} / \ln \frac{1}{r}.$$

Для исходных данных задачи оно равно двум. Таким образом, пороговый коэффициент усиления среды с поглотителем вдвое выше.

Контрольные вопросы

- Что такое лазер?
- Как осуществляется положительная обратная связь в лазере?
- Что такое мода лазера? Какие моды вы знаете?
- Что такое добротность лазерного резонатора?
- Что такое когерентность лазерного излучения?
- Как определяется размер пучка лазерного излучения?
- Как определяется расходимость лазерного излучения?
- Что такое диаграмма направленности лазерного излучения?

Рекомендуемая литература

1. Карлов Н. В. Лекции по квантовой электронике. Учебное пособие. — М.: Наука, 1983.
2. Пихтин А. Н. Физические основы квантовой электроники и оптоэлектроники. Учебное пособие. — М.: Высшая школа, 1983.
3. Страховский Г. М., Успенский А. В. Основы квантовой электроники. Учебное пособие. — М.: Высшая школа, 1979.
4. Успенский А. В. Сборник задач по квантовой электронике. Учебное пособие. — М.: Высшая школа, 1976.

3. Типы лазеров

Все известные типы лазеров условно можно разбить на следующие группы: лазеры на основе конденсированных сред, газовые лазеры, эксимерные лазеры и лазеры на свободных электронах.

3.1. Лазеры на основе конденсированных сред

3.1.1. Твердотельные лазеры

Под конденсированными средами будем понимать твердые тела и жидкости, атомные частицы (атомы, молекулы, ионы) которых связаны между собой. В таких телах не происходит самопроизвольный разрыв связей, и конденсированные тела сохраняют свой объем.

Твердотельный лазер представляет собой лазер, в котором активный средой служат диэлектрические кристаллы или стекла, активированные ионами редкоземельных элементов или ионами группы железа.

Первым лазером в оптическом диапазоне был лазер на кристалле рубина ($\text{Al}_2\text{O}_3 \cdot \text{Cr}^{+++}$). В нем использовались оптическая накачка и открытый резонатор.

Инверсия населенностей уровней в нем создается путем импульсного облучения мощной ртутно-кварцевой лампой. Ионы хрома Cr^{+++} поглощают излучение и переходят из основного уровня E_1 в полосу накачки E_3 . В результате безызлучательного перехода $E_3 \rightarrow E_2$ заселяется метастабильный уровень E_2 , так что число возбужденных на нем атомов превышает число атомов на уровне E_1 (см. рис. 2.1, а). При стимуляции возникает излучение в красной части спектра путем перехода возбужденных ионов хрома уровня E_2 на уровень E_1 . Наиболее благоприятные условия возникают для генерации излучения с длиной волны $\lambda = 0,6943$ мкм.

На рис. 3.1 приведена типичная функциональная схема твердотельных лазеров.

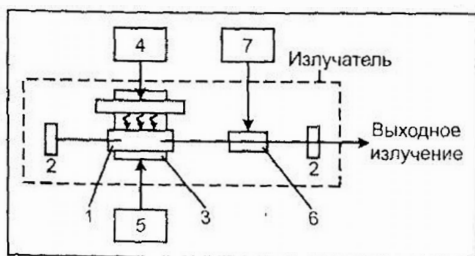


Рис. 3.1. Функциональная схема твердотельных лазеров:

- 1 — активный элемент;
- 2 — зеркала резонатора; 3 — осветитель;
- 4 — блок питания накачки;
- 5 — блок охлаждения;
- 6 — устройство управления;
- 7 — блок управления

Активный элемент 1, помещенный в оптический резонатор из двух зеркал 2, освещается осветителем. Осветитель 3 питается от блока конденсаторной батареи 4 и охлаждается

с помощью охлаждающей системы 5. Зачастую в резонатор встраивается устройство управления 6, позволяющее формировать лазерное излучение с заданными пространственно-временными характеристиками.

Блок охлаждения отводит от активного элемента и осветителя тепловую энергию, выделяемую при излучении и поглощении света накачки. КПД твердотельных лазеров составляет несколько процентов и поэтому отвод тепла крайне необходим. Выходная мощность лазера зависит от энергии накачки и имеет пороговую энергию. Выходная мощность может достигать нескольких десятков джоулей в импульсном и порядка ста милливатт в непрерывном режимах.

Обратную связь в оптическом резонаторе можно включать и выключать с помощью устройства управления. Обычно используется электрооптический затвор, например, в виде ячейки Керра или ячейки Погкельса, Фарадея. Обратная связь включается на промежутки времени $\sim 10^{-8} \div 10^{-9}$ с. Это время и определяет длительность импульса. Таким образом, накопленная на метастабильном уровне энергии возбужденных ионов хрома излучает энергию за очень короткий промежуток времени ($\sim 10^{-8}$ с), что позволяет развить гигантскую мощность. Сфокусированный поток такого излучения обладает гигантской плотностью мощности, способной разрушить любые материалы.

Такие лазерные импульсы получили название гигантских ($\sim 10^{10}$ Вт и более), а конструкции получили название лазеров с модулированной добротностью. Гигантские импульсы можно увеличить по мощности, используя каскад лазерных усилителей. Такая мощность превосходит мощность самых больших электростанций мира.

Вторым типом твердотельных квантовых генераторов был лазер на стекле, активированном ионами неодима Nd^{+++} . В настоящее время это один из наиболее распространенных типов лазеров благодаря хорошей технологичности и низкой стоимости стеклянных активных лазерных элементов. Это позволяет изготавливать активные элементы очень больших размеров ($\sim 10^2$ см) и снимать значительную энергию ($\sim 10^4$ Дж). Однако стекла имеют плохую теплопроводность и требуют применения эффективных систем охлаждения.

Лазеры на неодимовых стеклах работают по четырехуровневой схеме и излучают на основной длине волны 1,06 мкм, а также могут излучать на длине волны 1,32 мкм. Это ближний инфракрасный диапазон. В режиме свободной генерации длительность импульсов излучения составляет 0,1—10,0 мс. Мощность достигает значений 10^{13} Вт в режиме модулированной добротности и согласования мод.

Помимо неодима получили распространение лазеры, активный элемент которых активирован ионами Er^{3+} , Tu^{3+} , Ho^{3+} (эрбий, тулий, гольмий).

Среди лазерных кристаллов, легированных неодимом, наибольшее распространение получил кристалл алюмоиттриевого граната $\text{Y}_3\text{Al}_5\text{O}_{12}$ с атомарной концентрацией Nd^{+++} до 1% по отношению к иттрию.

Другим распространенным активным элементом твердотельных лазеров является ортоалюминат иттрия YAlO_3 . Он также легируется ионами неодима.

Определенный интерес для создания миниатюрных твердотельных лазеров представляют монокристаллы двойного галлий-гадолиниевого вольфрама (КГВ), легированные неодимом, а также гелий-скандий-гадолиний гранат (ГСГГ), легированные либо хромом, либо неодимом. К настоящему времени эффект стимулированного излучения обнаружен более чем у 250 кристаллов с примесью ионов переходных групп (Nd, Eu, Ho, Er, Tu, Yb).

3.1.2. Полупроводниковые лазеры

Полупроводниковый лазер представляет собой лазер на основе полупроводниковой активной среды, в которой используются квантовые переходы между разрешенными энергетическими зонами. Это отличает полупроводниковые лазеры от других типов, в которых используются квантовые переходы между дискретными уровнями энергии.

В соответствии с зонной теорией полупроводников при поглощении фотона, энергия которого больше ширины запрещенной зоны E_g , происходит переброс электрона из валентной зоны E_v в зону проводимости E_c (рис. 3.2, а). При этом в валентной зоне образуется дырка.

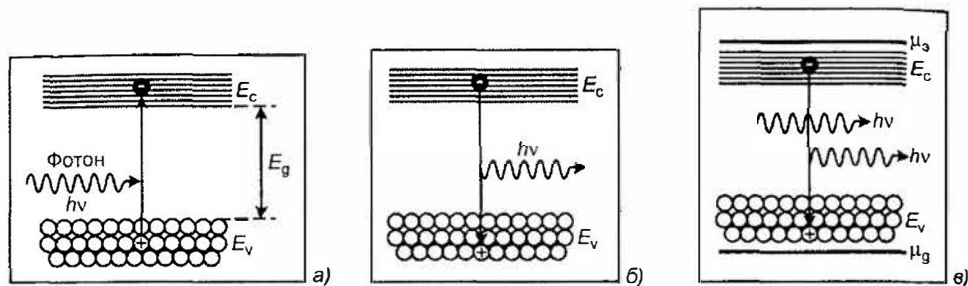


Рис. 3.2. Процесс взаимодействия полупроводника с квантами света:

а — образование электронно-дырочной пары; б — спонтанная рекомбинация, сопровождаемая излучением фотона; в — вынужденная рекомбинация электрона и дырки

Одновременно с генерацией электронно-дырочных пар в полупроводнике происходит процесс спонтанного излучения (рис. 3.2, б).

В полупроводнике имеет место вынужденная рекомбинация под действием света (рис. 3.2, в). Для создания условий усиления света необходимо создать вырожденный полупроводник, в котором нарушено тепловое равновесие. С этой целью увеличивается концентрация электронов вблизи дна зоны проводимости и дырок около потолка валентной зоны.

Обозначим наивысший уровень энергии, до которого электрон плотно заполняет зону проводимости, величиной μ_c . Чем больше электронов упаковано в дно зоны проводимости, тем выше этот уровень μ_c .

Аналогично в валентной зоне существует уровень μ_g . Если одновременно вырождены электроны и дырки, то

$$\mu_c - \mu_g \geq E_g. \quad (3.1)$$

При таких условиях электроны могут рекомбинировать только с теми дырками, которые лежат ниже μ_g . В свою очередь электроны могут быть заброшены в зону проводимости только на уровни выше μ_c , потому как остальные уровни уже плотно упакованы (рис. 3.2, в).

При этом случае возможны переходы в интервале частот

$$\mu_c - \mu_g = h\nu_{\max}$$

или

$$E_c - E_v = E_g = h\nu_{\min}.$$

В этом случае полупроводник может усиливать и генерировать свет в полосе частот

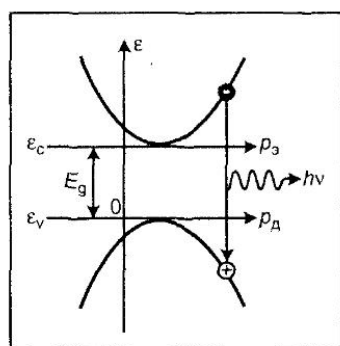
$$\Delta\nu = \nu_{\max} - \nu_{\min}.$$

Если поместить такой полупроводник в резонатор Фабри — Перо и создать положительную обратную связь, то при каждом проходе резонатора в полупроводнике полоса частот будет сужаться. Дело в том, что усиление в полосе частот $\Delta\nu$ неодинаково. Существует частота максимального усиления, которая лежит в диапазоне $\Delta\nu$ и именно на ней происходит максимальное усиление и формируется монохроматическое излучение.

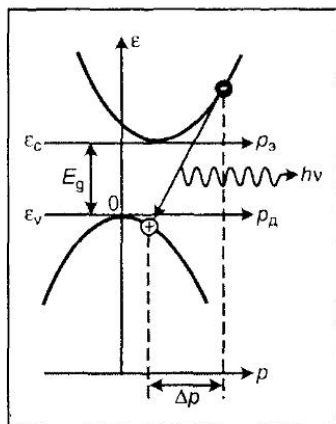
Зависимость энергии электронов проводимости вблизи дна зоны проводимости и энергии дырок вблизи потолка валентной зоны от величины импульса p имеет вид парабол (рис. 3.3), т. е., соответственно

$$E_c = \frac{p^2}{2m^*},$$

$$E_v = \frac{p^2}{2m^*}.$$



а)



б)

Рис. 3.3. Зависимость энергии электронов проводимости и дырок от их импульсов: а — для прямого перехода; б — для непрямого перехода

Величина m^* обозначает эффективную массу электронов и дырок, которая существенно зависит от структуры кристалла. Расстояние между вершинами парабол равно ширине запрещенной зоны E_g .

Различают *прямые* и *непрямые* переходы. Если рекомбинация электронно-дырочной пары не сопровождается изменением колебательного состояния решетки, то переход электрона из зоны проводимости на незаполненный уровень в валентной зоне называется *прямым*. Если рекомбинация электронно-дырочной пары сопровождается изменением колебательного состояния решетки, то переход называется *непрямым*. В этом случае колебания решетки поглощают часть импульса Δp и, соответственно, и энергию ΔE .

В этом случае уравнение инверсии населенностей для невырожденного электронно-дырочного газа записывается в виде

$$\mu_n - \mu_p = E_g + \Delta E.$$

Создать в чистых полупроводниках состояние с одновременным вырождением электронов и дырок трудно. Обычно используют два полупроводника *n*- и *p*-типов, в каждом из которых электроны и дырки вырождены. В *p*—*n*-переходе может выполняться условие инверсии населенностей $\mu_n - \mu_p \geq E_g$ только при условии прямого напряжения.

Через *p*—*n*-переход текут токи, состоящие из электронов и дырок. В тонком слое *p*—*n*-перехода они рекомбинируют, излучая фотоны

$$h\nu = E_g < \mu_n - \mu_p.$$

Встречные потоки электронов и дырок будут поддерживать в *p*—*n*-переходе концентрацию, достаточную для условия вырождения электронов и дырок (рис. 3.4).

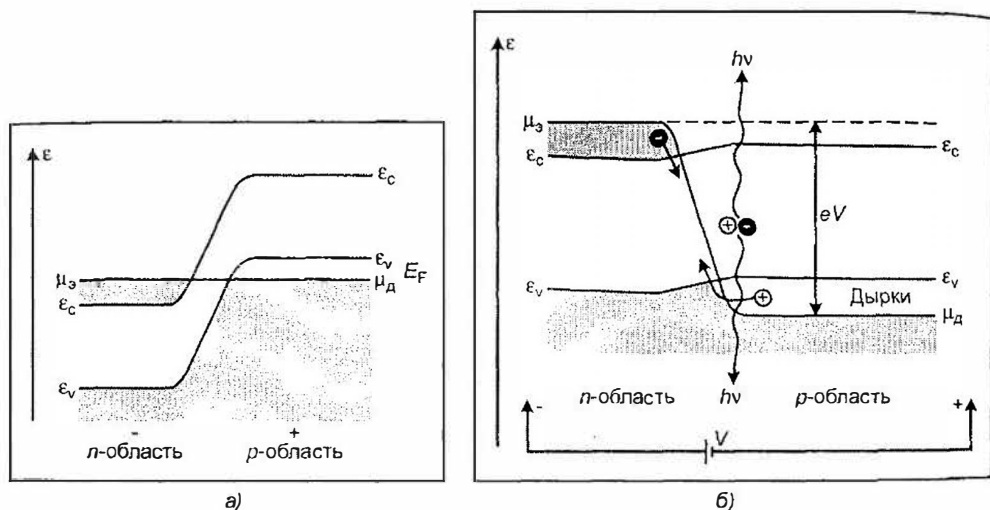


Рис. 3.4. Энергетические диаграммы *p*—*n*-перехода в положение равновесия (а) и при приложении напряжения (б)

Чем выше электрическое поле в *p*—*n*-переходе, тем больший ток протекает через *p*—*n*-переход. Минимальный ток, при котором вынужденное излучение превышает поглощение, называется *пороговым*. На пороге генерации должно выполняться условие

$$\kappa R \geq 1,$$

где κ — коэффициент усиления на длине активной среды между зеркалами, R — коэффициент отражения зеркал резонатора. При токе выше порогового *p*—*n*-переход является усиливающей средой. Введя положительную обратную связь в виде резонатора, из усилителя света можно получить генератор. В качестве резонатора используются гладкие грани полупроводникового кристалла.

Полупроводниковый лазер, в котором генерация когерентного излучения осуществляется в результате инжекции носителей заряда через электронно-дырочный переход, получил название *инжекционного лазера*.

Сравнительно малые размеры резонатора не позволяют получить высокую направленность излучения. Состояние инверсии населенности уровней может достигаться электронной накачкой, оптической накачкой, электрическим пробоем в сильном поле.

Наилучший эффект генерации получен на прямоугольных полупроводниках, среди которых ряды изоморфных твердых растворов типа A^2B^6 , A^3B^6 , A^3B^5 , A^4B^6 и т. п.

Особый интерес вызывают материалы, составляющие изопериодические пары. Это кристаллы, различающиеся по количественному составу, ширине запрещенной зоны, но имеющие одинаковый период кристаллической решетки. С их помощью методами электронно-лучевой эпитаксии выращиваются бездефектные гетеропереходы. Широкое распространение получили гетеролазеры, сформированные на основе гетероструктур.

Полупроводниковый лазер, в котором генерация когерентного излучения осуществляется в результате излучательной рекомбинации в гетероструктуре, получил название *гетеролазера*.

Наиболее эффективными оказались полупроводники типа A^3B^5 с высоким квантовым выходом излучательной рекомбинации. Наилучшими параметрами обладает гетеролазер на основе двойной гетероструктуры (ДГС) с активным слоем из узкозонного полупроводника, заключенного между слоями более широкозонного полупроводника. Двустороннее оптическое и электронное ограничение приводит к совмещению областей инверсной заселенности и светового поля. В таких лазерах уже при малом токе накачки можно получить устойчивую генерацию. Резонатором таких лазеров служат грани кристалла, но можно использовать и внешние оптические резонаторы. На рис. 3.5 приведена структура гетеролазера с резонатором в виде сколотых граней кристалла.

Твердые лазерные растворы $Ga_{1-x}As_xP_{1-y}$ позволяют получать коротковолновое излучение, а растворы типа SnP позволяют получать низкороговые инжекционные лазеры инфракрасного диапазона. Дальняя ИК область ($\lambda > 5$ мкм) осваивается с помощью твердых растворов PbS_xSe_{1-x} и $Pb_xS_{1-x}Te$.

Полупроводниковые лазеры могут быть объединены в многоэлементные излучатели — *фазированные лазерные монолитные линейки*. Мощность излучения в импульсном режиме может достигать значения 10^5 Вт. Модуляцией накачки можно получить модулированное по амплитуде лазерное излучение.

Также создаются гетеролазеры с распределенной обратной связью. В таком типе лазеров одно из зеркал резонатора заменено дифракционной решеткой. С ее помощью можно получить излучение строго на определенных частотах.

Полупроводниковые лазеры находят широкое применение в устройствах обработки информации.

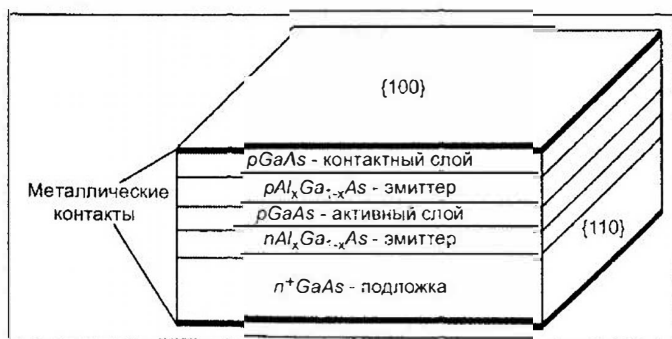


Рис. 3.5. Структура ДГС гетеролазера на основе арсенида галлия

В табл. 3.1 приведены некоторые материалы, на основе которых получен эффект генерации и сконструированы полупроводниковые лазеры.

Таблица 3.1. Материалы для полупроводниковых лазеров

Полупроводник	Рабочая температура, К	Длина волны излучения, Å	Цвет излучения
ZnS	80	3300	Ультрафиолетовый
ZnSe	80	4530	Фиолетовый
CdS	4—300	4850	Зеленый
CdSe	80	6100	Оранжевый
CdSe	80	6950	Красный
CdTe	4—80	7900—7960	Красный
GaPAs	80—300	8300—8360	Темно-красный
GaAs	4—300	8200—9000	Инфракрасный
GaSb	20	15300	» »
InAs	20	30080	» »
InSb	20	49590	» »
Te	20	38440	» »
PbS	4	42700	» »
PbTe	4	64100	» »
PbSe	4	85500	» »

3.1.3. Жидкостные лазеры

Жидкостный лазер представляет собой квантовый генератор, рабочим веществом которого является жидкость.

Жидкостные лазеры делятся на следующие типы:

- ☐ на красителях;
- ☐ на органических жидкостях;
- ☐ на неорганических жидкостях.

Жидкостные лазеры были реализованы на металлоорганических жидкостях, а именно на хелатах европия. Однако малая фотохимическая стойкость, большой коэффициент поглощения света не позволили применять их в промышленных конструкциях.

Неорганические жидкостные лазеры представляют собой раствор соединений типа TR^{+++} в неорганических растворителях сложного состава. Активными ионами в них служат, например, Nd^{+++} . Генерация происходит по четырехуровневой схеме с поглощением света накачки собственными полосами поглощения Nd^{+++} . Рабочее вещество, например, смесь хлороксида фосфора (POCl_3) с кислотой SnCl_4 и ионами Nd^{+++} находится в режиме циркуляции и позволяет получить узкий спектр генерации. В лазерах на красителях в качестве рабочего вещества используются сложные органические соединения. Эти соединения

обладают системой сопряженных связей и интенсивными полосами поглощения во всех областях спектра. Растворы красителей представляют собой красители в воде, спирте, бензоле или активированные красителями полимерные материалы типа полиметилметакрилата, полиуретана. Схема лазера на красителях представлена на рис. 3.6.

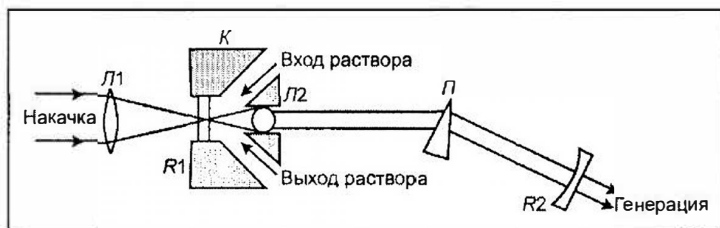


Рис. 3.6. Схема лазера на красителях

Главной особенностью лазеров на красителях является возможность перестройки длины волны генерируемого излучения в диапазоне от ультрафиолетового (~ 330 нм) до ИК-диапазона ($\sim 1,8$ мкм). Грубая перестройка производится путем замены кюветы с красителем. Для этого нужно примерно 30 кювет с различными соединениями, которых насчитывается порядка тысячи (рис. 3.7).

Тонкая настройка на заданную длину волны осуществляется с помощью спектрально-селективных элементов, вводимых в резонатор. Такими элементами могут служить дисперсионные призмы, интерференционно-поляризационные фильтры и т. п.

Накачка таких лазеров осуществляется излучением импульсных ламп, излучением лазеров других типов. В зависимости от типа накачки различают лазеры импульсного или непрерывного режимов работы.

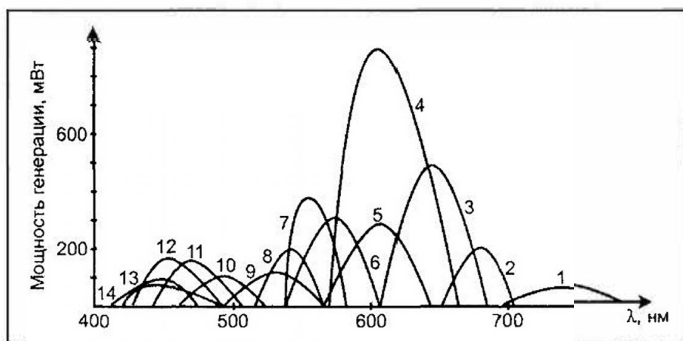


Рис. 3.7. Спектр мощности непрерывного лазера при использовании различных красителей: 1 — нильский голубой; 2 — крезил-фиолет перхлорат; 3 — родамин В; 4 — родамин 6G (мощность возбуждения 4 Вт); 5 — родамин 6G (мощность возбуждения 2 Вт); 6 — родамин 110; 7 — флуоресцеин; 8 — кумарин В; 9 — кумарин 7; 10 — кумарин 102; 11 — 7-диэтил-амино-4-метилкумарин; 12 — кумарин 2; 13 — кумарин 120; 14 — карбостирол

Особый класс составляют лазеры с *распределенной обратной связью* (РОС). В РОС-лазерах роль резонатора выполняет структура с периодическим изменением показателя преломления или усиления. Это можно осуществить, например, воздействуя на активную

среду интерферирующими пучками накачки. РОС-лазеры способны генерировать на узкой линии ($\sim 10^{-2} \text{ см}^{-1}$), которая легко перестраивается в пределах полосы усиления.

Лазеры на красителях с пассивной синхронизацией мод позволяют генерировать ультракороткие импульсы излучения (10^{-14} с). Основная область применения жидкостных лазеров — это спектроскопические исследования.

Особенностью жидкостных лазеров является высокое значение ширины линии усиления активного перехода. Это обстоятельство позволяет создавать мощные лазеры в Уф-диапазоне длин волн. Одновременно можно производить плавную перестройку длины волны в достаточно широком диапазоне.

3.2. Газовые лазеры

Газовый лазер представляет собой лазер, в котором активной средой являются газы, пары или смеси газов или паров.

Как и все типы лазеров, газовый лазер состоит из следующих основных элементов (см. рис. 3.1):

- ☐ активная среда с усилением на одной или нескольких линиях;
- ☐ оптический резонатор для создания положительной обратной связи;
- ☐ устройство накачки для создания инверсной населенности уровней.

Трубку или камеру с активной газовой средой помещают в оптический резонатор, состоящий из зеркал различной конфигурации. Плоскости зеркал должны быть перпендикулярны продольной оси трубки или камеры с газом. При создании в газе инверсной населенности уровней в результате накачки электромагнитным полем стимулируется процесс вынужденного излучения. Лазерное излучение выводится из резонатора через полупрозрачное зеркало, через края непрозрачного зеркала или через отверстие в нем.

Среди отличительных особенностей газовых лазеров по сравнению с лазерами на конденсированных средах следует отметить следующие:

- ☐ высокая оптическая однородность активной среды;
- ☐ узкие спектральные линии излучения;
- ☐ высокая степень когерентности излучения;
- ☐ острая направленность излучения;
- ☐ стабильность частоты излучения;
- ☐ широкий диапазон рабочих длин волн излучения;
- ☐ возможность использования активных газовых сред большого объема и протяженности.

По способу накачки газовые лазеры условно делятся на газоразрядные, газодинамические и химические.

3.2.1. Лазеры на нейтральных атомах

Первым газовым лазером был гелий-неоновый лазер. Он относится к газоразрядным лазерам. Инверсная населенность уровней в нем осуществляется с помощью газового разряда. Атомы возбуждаются при их соударении с быстрыми свободными электронами. Давление газов составляет величину в пределах долей мм рт. ст. Малая плотность газов позволяет

зародившемуся в результате вынужденного излучения лучу многократно проходить между зеркалами резонатора не искажаясь. На рис. 3.8 приведена схема газового He-Ne лазера. Капилляр с газом помещается в цилиндр катода газоразрядной трубки. Окошки Брюстера помогают уменьшить потери излучения за проход луча. В He-Ne лазере рабочим веществом служат нейтральные атомы неона.

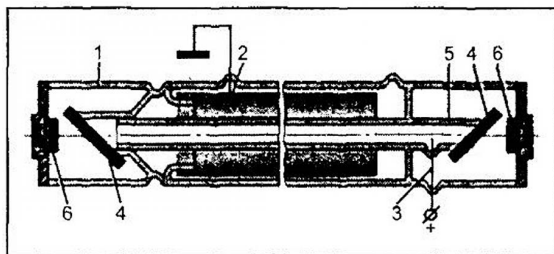


Рис. 3.8. Схема He-Ne лазера: 1 — разрядная трубка; 2 — катод; 3 — анод; 4 — окна Брюстера; 5 — капилляр со смесью газов; 6 — зеркала

Инверсия населенностей уровней осуществляется за счет первоначального возбуждения атомов гелия на уровни E_2 и E_3 . Они точно совпадают с уровнями E_4 и E_5 возбужденных атомов неона. При столкновении возбужденных атомов гелия с атомами неона происходит перекачка энергии (рис. 3.9). Механизм резонансной передачи возбуждения позволяет перевести атомы неона в возбужденное состояние. Гелий, как буферный газ, является резервуаром возбуждения неона. Неон нельзя возбудить прямым переходом на уровни E_4 и E_5 из-за долго живущего метастабильного уровня E_2 .

Инверсия населенностей достигается за счет долгоживущих уровней E_4 и E_5 . Именно переход с них на уровень E_3 позволяет получить стимулированное излучение в красном и инфракрасном диапазонах длин волн.

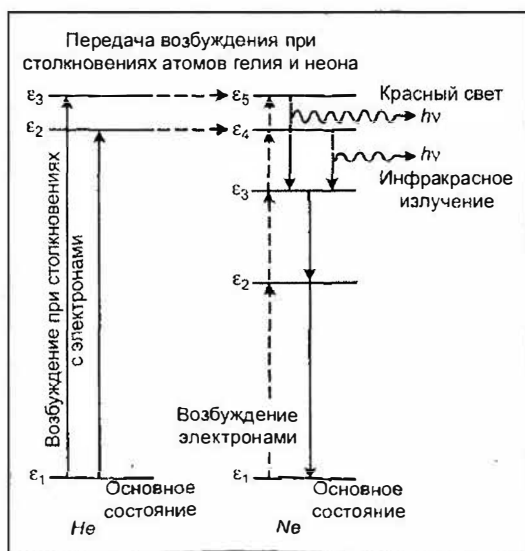


Рис. 3.9. Уровни энергии возбуждения атомов гелия и неона

Опустошение короткоживущего уровня E_3 иона происходит путем соударений атомов неона со стенками разрядной трубки, что переводит их на уровень E_2 . С этой целью подбирается оптимальный диаметр газоразрядной трубки (до 10 мм). Генерация в красном диапазоне происходит на длине волны 0,63 мкм (6328 Å). Также возможна генерация излучения гелий-неоновых лазеров на длине волны 1,15 и 3,39 мкм. В настоящее время получена генерация на более чем 450 переходах между уровнями нейтральных атомов. Помимо гелий-неонового лазера весьма популярен ксеноновый лазер, генерирующий на длине волны $\lambda = 3,5$ мкм, и гелий-ксеноновый ($\lambda = 2,02$ мкм). Эти лазеры работают в непрерывном режиме, который обеспечивается газоразрядной накачкой.

Газовые лазеры работают в широком диапазоне длин волн (от 100 нм до 1000 мкм) и мощностей излучения (от 100 мВт до 1 МВт) в непрерывном режиме и до 1 ТВт в импульсном режиме.

3.2.2. Ионные лазеры

Ионные лазеры имеют в качестве рабочих веществ ионизированные инертные газы (Xe, Kr, Ar, Ne), а также ионы фосфора, серы, хлора, кадмия, цинка и др.

В ионных газовых лазерах лазерные переходы происходят между уровнями одно- или двукратно ионизированных атомов. Этому способствуют большие плотности тока при ионизации газов. Инверсия населенностей уровней осуществляется между уровнями E_3 и E_4 . Уровень E_4 с большим временем жизни сильно заселяется возбужденными атомами. Атомы возбуждаются при их соударении с быстрыми электронами в газовом разряде, а также путем перехода с полосы E_5 (рис. 3.10).

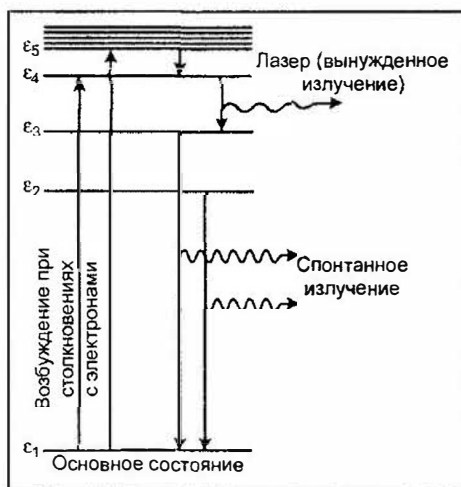


Рис. 3.10. Схема уровней, используемых в ионных лазерах

Уровень E_3 имеет короткое время жизни относительно уровня E_1 , и поэтому он быстро опустошается. Для сильной ионизации газа ток пропускают через длинный капилляр, который охлаждается из-за большой выделяемой энергии. Для увеличения концентрации электронов в центре капилляра создается продольное магнитное поле, которое сжимает разряд и предохраняет стенки от разрушения. На рис. 3.11 приведена одна из схем ионного лазера.

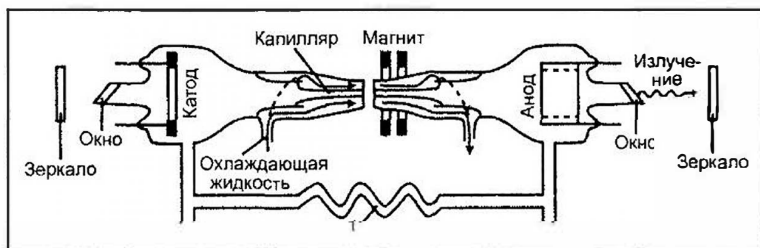


Рис. 3.11. Схема ионного лазера на аргоне: 1 — трубка обратной циркуляции газа

Первичные электроны генерируются катодом и на своем пути в разрядной трубке ионизируют газ. Для компенсации перекачки газов используют длинную трубку, которая не шунтирует газовый разряд.

Разряд в газе возбуждается постоянным, либо переменным высокочастотным полем.

Наибольшее распространение получил аргоновый лазер. Он генерирует в сине-зеленой части спектра на длинах волн 0,488 и 0,514 мкм. Аргоновые лазеры могут излучать и в УФ-части спектра (0,340—0,370 мкм). Мощность излучения достигает киловатта. Аргоновый лазер может работать как в непрерывном, так и импульсном режимах. Криптоновые лазеры генерируют излучение в красной части спектра (0,65—0,86 мкм).

3.2.3. Молекулярные лазеры

Молекулярные лазеры представляют собой оптические квантовые генераторы, в качестве активного вещества которых используются молекулы. В отличие от атомов, молекулы газов имеют кроме электронных энергетических уровней также колебательные и вращательные.

Первый молекулярный лазер был реализован на смеси молекул углекислого газа и азота. Молекула CO_2 имеет три частоты собственных возбуждаемых колебаний, которым соответствуют уровни E_3 , E_4 , E_5 (рис. 3.12). Заселение этих уровней происходит вследствие нескольких одновременно протекающих процессов. Молекулы углекислого газа возбуждаются при соударении с быстрыми электронами в газовом разряде. Одновременно вводится азот и в газовом разряде происходит ионизация молекул азота и углекислого газа. Возбуждается уровень E_2 в молекуле азота и уровень E_5 в молекуле углекислого газа. Молекулы азота возбуждаются весьма интенсивно и энергию возбуждения резонансно передают молекулам CO_2 .

Еще большую инверсную заселенность уровня можно достичь введением в газовую смесь гелия. Благодаря большой теплопроводности гелия понижается температура смеси в охлаждаемом рабочем объеме. Это уменьшает тепловое заселение уровней и возвращает молекулы CO_2 с самых высоких колебательных уровней на уровень E_5 .

В дальнейшем происходит переход с полосы уровней E_5 на нижележащие уровни с одновременным излучением. При этом генерация возникает на большом числе переходов молекулы CO_2 в интервале волн от 9 до 18 мкм.

На рис. 3.13 приведена схема мощного лазера. В электрическом разряде имеют место нежелательные эффекты, такие как разогрев газа и диссоциация его молекул. Эти паразитные эффекты устраняются тем, что газовая смесь непрерывно прокачивается через разрядные трубки лазеров с целью регенерации газовой смеси.

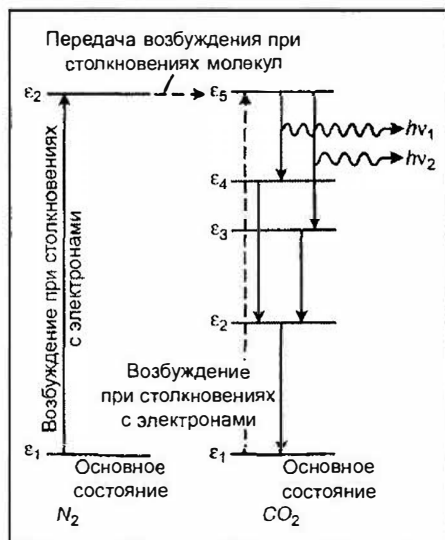


Рис. 3.12. Схема уровней молекул углекислого газа и азота

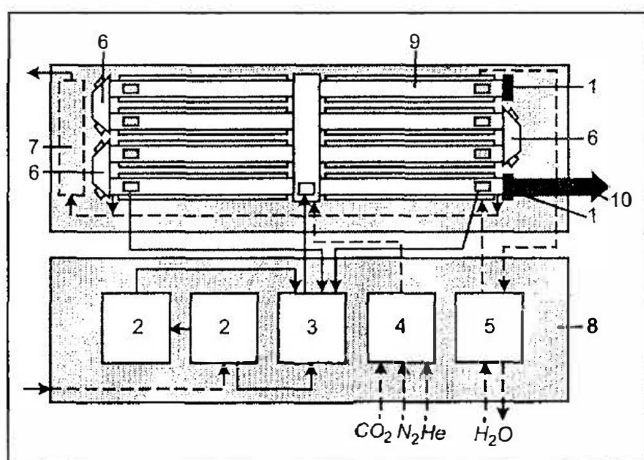


Рис. 3.13. Схема мощного CO_2 лазера: 1 — зеркала радиатора; 2 — блок управления лазером с системой блокировки; 3 — источник питания; 4 — смеситель газов; 5 — блок управления водяным охлаждением; 6 — система поворотных призм; 7 — система прокачки; 8 — анодный блок; 9 — многократно свернутый оптический резонатор; 10 — лазерный луч

Газовая смесь прокачивается через разрядные трубки, составляющие многократно свернутый оптический резонатор. С одного метра активной среды такого лазера можно получить лазерное излучение мощностью более 50 Вт на длине волны 10,6 мкм. Повышение удельной мощности можно получить путем интенсивного охлаждения рабочей смеси.

Лазеры, в которых возбуждение рабочей смеси достигается за счет возбуждения разряда высокого давления и пучком быстрых электронов с энергией до 500 кэВ, носят название электроионизационных CO_2 -лазеров. В таком типе лазеров с одного метра активной среды можно снять до 100 Дж и достигнуть максимальной энергии свыше 10 кДж.

Особую техническую трудность представляет вопрос лазерных зеркал. Их делают из кристаллов KCl, ZnSe, CdTe, n-Ga, NaCl. Большая выходная мощность излучения выводится не как обычно, через полупрозрачное зеркало, а через прозрачное для ИК-излучения окно в зеркале. В этом случае изготавливают охлаждаемые водой зеркала из бронзы, молибдена и других металлов.

Помимо молекулярных лазеров на CO_2 разработаны лазеры на моноокиси углерода CO, на парах воды, работающие на длинах волн $\lambda = 27,9$ мкм, $\lambda = 118,6$ мкм.

В далеком ИК-диапазоне ($\lambda = 337$ мкм) работают лазеры на молекулах HCN.

Коэффициент полезного действия молекулярных лазеров лежит в пределах 20—40%, а максимальная мощность 100 кВт и более.

3.2.4. Эксимерные лазеры

Эксимерный лазер — газовый лазер, активная среда которых представляет собой электронные переходы эксимерных молекул.

Эксимерные молекулы существуют только в электронно-возбужденном состоянии. Основному электронному состоянию таких молекул соответствует разлетный терм.

Зависимость потенциальной энергии взаимодействия E атомов эксимерной молекулы от межатомного расстояния R является монотонно спадающей функцией (рис. 3.14).

Для возбужденного электронного состояния, которое является верхним уровнем лазерного перехода, имеется минимум. Этот минимум и определяет существование эксимерной молекулы. Время жизни возбужденной эксимерной молекулы определяется временем ее радиационного распада. Нижнее состояние лазерного перехода опустошается в процессе разлета атомов эксимерной молекулы.

Газ, содержащий эксимерные молекулы, является активной средой. Эксимерные молекулы, как правило, представляют собой короткоживущие соединения атомов инертных газов друг с другом, с галогенами и кислородом.

Конструкция эксимерных лазеров типична для газовых лазеров. Возбуждение активной среды производится электронными пучками, газовым разрядом, оптической накачкой или комбинацией этих способов.

Длина волны излучения лежит в видимой или ближней УФ-области спектра. В табл. 3.2 приведены некоторые параметры лазеров на наиболее распространенных эксимерных молекулах.

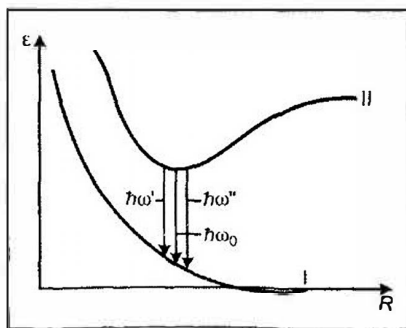


Рис. 3.14. Зависимость энергии E эксимерной молекулы от межатомного расстояния R

Особенностью эксимерных лазеров является высокое значение ширины линии усиления активного перехода. Это обстоятельство позволяет создавать мощные лазеры длин волн УФ-диапазона. Одновременно можно производить плавную перестройку длины волны в достаточно широком диапазоне. Особой мощностью обладают эксимерные лазеры на основе F_2 . Например, KrF-лазер имеет выходную энергию в импульсе до 100 кДж и длительностью импульса порядка 1 нс, что позволяет использовать его в экспериментах по термоядерному синтезу.

Таблица 3.2. Параметры эксимерных лазеров

Эксимерная молекула	Длина волны в центре линии перехода, нм	Ширина линии усиления, нм
Ar ₂	126,1	8
Kr ₂	146,7	14
Xe ₂	172,0	20
ArF (B-X)	193,3	7
KrCl (B-X)	222,0	7
KrF (B-X)	248,4	15
XeBr (B-X)	281,8	1,0
XeCl (B-X)	308,0	12
XeF (B-X)	351,1	25
XeF (C-A)	485	100
XeO	540	25,0
KrO	558	1,5
ArO	558	4,0

3.2.5. Газодинамические лазеры

Газодинамический лазер — газовый лазер, в котором инверсия населенностей создается в системе колебательных уровней энергии молекул газа с помощью адиабатического охлаждения нагретых газовых масс, движущихся со сверхзвуковой скоростью. Другими словами, в качестве активной среды используется быстро охлаждаемая смесь газов, инверсная населенность уровней в которой достигается при следующих условиях:

- ☐ скорость опустошения нижнего уровня лазерного перехода выше скорости опустошения внешнего уровня;
- ☐ время опустошения верхнего уровня больше времени движения газа в резонаторе.

Эти условия обеспечивают так называемую инверсию населенностей верхних уровней. На рис. 3.15 приведена схема газодинамического лазера.

В камере сгорания сжигается углеводородное топливо с воздухом в качестве окислителя. Нагретая газовая смесь аэродинамическими средствами разгоняется до сверхзвуковой скорости (~ 1,8 км/с) и резко расширяется. Молекулы газа CO₂, Ar, Ne опустошают свой инверсный уровень, что создает условия для генерации излучения. Оптический резонатор,

зеркала которого параллельны потоку, имеет значительные размеры и способен усиливать колебания на длинах волн $\lambda_1 = 18,4$ мкм, $\lambda_2 = 16,7$ мкм, $\lambda_3 = 16,2$ мкм. На СО-лазере получено излучение на длине $\lambda = 5$ мкм.

В непрерывном режиме газодинамические лазеры позволяют получать излучение мощностью $\sim 10^2$ кВт при КПД $\sim 1\%$.

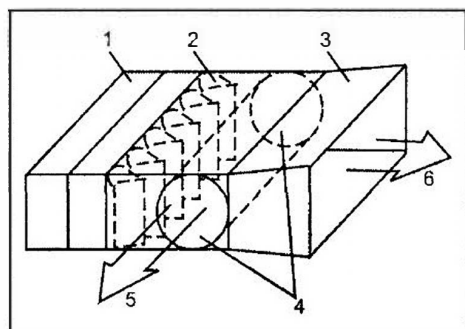


Рис. 3.15. Схема газодинамического лазера: 1 — камера сгорания смеси; 2 — сверхзвуковые сопла Лавалья; 3 — диффузор; 4 — зеркала оптического резонатора; 5 — лазерное излучение; 6 — выход отработанного газа

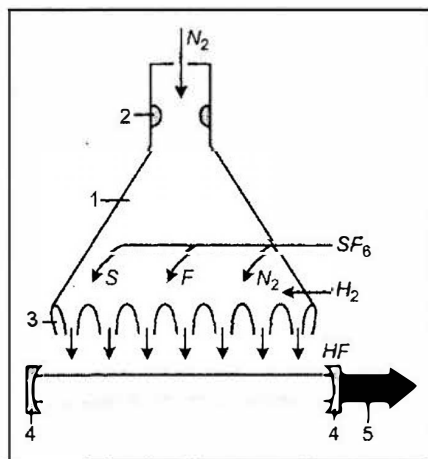


Рис. 3.16. Схема химического лазера с использованием реакции фторирования водорода: 1 — реактивная камера; 2 — нагреватель; 3 — сопла; 4 — зеркала резонатора; 5 — лазерное излучение

3.2.6. Химические лазеры

Химические лазеры — лазеры, в которых инверсия населенностей создается во время экзотермических химических реакций, приводящей к преобразованию химической энергии в энергию электромагнитного излучения.

Различают три вида химических реакций, на основе которых созданы химические лазеры:

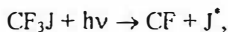
- ❑ фотодиссоциация или распад молекул под действием света;
- ❑ диссоциация молекул при электрическом разряде в газе;
- ❑ взаимодействие молекул, атомов и соединений.

Химический лазер с использованием реакции фторирования водорода представлен на рис. 3.16.

Молекулярный азот N_2 нагревают в камере до $T = 2000$ К и одновременно в реактивную камеру вводят газообразный гексафторид серы (SF_6). В процессе смешения с горячим азотом происходит диссоциация с образованием атомов фтора. Смесь продувается со сверхзвуковой скоростью через сопла Лавалья. Одновременно вводится водород H_2 . В результате взаимодействия фтора и водорода образуется колебательно-возбужденные молекулы фтористого водорода (HF)*, которые проникают через оптический резонатор из параллельных зеркал. В оптическом резонаторе возбуждается когерентное излучение на

длинах волн $2,6 \div 3,6$ мкм. Мощность непрерывной генерации достигает десятки киловатт при КПД $\sim 10\%$.

Фотодиссационные лазеры содержат в качестве активной среды результат распада, например, по схеме



где J^* — возбужденный атом иода.

Йодный фотодиссационный лазер работает на длине волны $\lambda = 1,31$ мкм с энергией в импульсе до 100 Дж.

Существуют и другие типы лазеров с использованием различных типов химических реакций. Разрабатываются лазеры, работающие в видимом и ультрафиолетовом диапазоне длин волн.

3.2.7. Лазеры на парах металла

Лазеры на парах металлов — газовые лазеры, активная среда которых является парами металла.

В настоящее время генерация газовых лазеров осуществляется на переходах атомов и атомарных ионов более 50 элементов, из них половина на переходах атомов или ионов металлов.

Разрядная трубка с металлом помещается в трубчатую печь. Для поддержания разряда в ненагреваемых частях трубки используется буферный инертный газ, например, гелий (рис. 3.17). В некоторых конструкциях буферный газ используется в процессе создания инверсной заселенности уровней, например, в гелий-кадмиевом лазере. Вместо термического нагрева иногда используют явление катафореза или движения дисперсионных частиц под действием электрического поля. Это явление объясняется существованием на границе двойного электрического слоя.

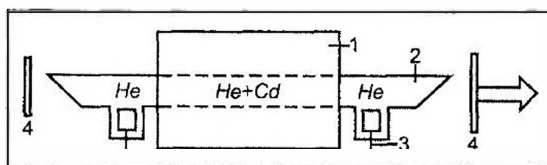


Рис. 3.17. Схема лазера на парах металла:
1 — нагреватель;
2 — резонаторная трубка;
3 — электроды;
4 — зеркала резонатора

Одной из перспективных конструкций является лазер на парах меди. В активной системе лазера на парах меди при переходах с резонансного уровня на метастабильный возникает уникально высокий коэффициент усиления. Эти лазеры могут работать в режиме сверхсвечимости, при котором использование оптического резонатора необязательно. Такая возможность позволяет использовать лазер на парах меди в качестве когерентных усилителей света, способных за один переход активной среды на несколько порядков усилить яркость световых пучков, формирующих изображение или какую-нибудь оптическую информацию. Лазер на парах меди генерирует в зеленой части спектра $\lambda = 0,510$ мкм. Возможна генерация также на желтой линии $\lambda = 0,578$ мкм. Накачка активной среды в лазерах на парах металлов осуществляется газовым разрядом, оптическим путем, процессом перезарядки, в процессе рекомбинации двукратно заряженных ионов.

3.2.8. Лазеры на свободных электронах

Лазеры на свободных электронах представляют собой генератор когерентных электромагнитных колебаний оптического диапазона длин волн, принцип работы которого основан на взаимодействии пучка свободных релятивистских электронов с пространственно-периодическим электрическим или магнитным полем.

Лазер на свободных электронах близок к приборам релятивистской высокочастотной электроники. Это прибор, принцип работы которого основан на эффектах квантовой и вакуумной электроники.

Пучок релятивистских электронов создается ускорителем заряженных частиц и направляется в ондулятор. Ондулятор представляет собой устройство, в котором создаются электромагнитные поля, действующие на движущуюся в нем заряженную частицу с периодической силой. Так, что среднее за период значение силы равно нулю.

Движущаяся заряженная частица, понав в ондулятор, совершает периодические колебания и генерирует ондуляторное излучение. В общем случае цуги волн ондуляторного излучения представляют собой гармоники, кратные основной частоте. Частоты ω_n n -й гармоники в соответствии с эффектом Доплера определяются как

$$\omega_n = \frac{n\Omega}{1 - \beta_{||} \cos \theta},$$

где $\Omega = \frac{2\pi\beta_{||}c}{\lambda_0}$ — частота колебаний частицы в ондуляторе. При $\theta = 0$ частоты ондуляторного излучения максимальны. Длина волны первичного излучения в направлении движения электронов

$$\lambda = \frac{\lambda_0}{2\gamma^2},$$

где λ — период электрического или магнитного поля в ондуляторе, γ — отношение кинетической энергии электрона к их энергии покоя. Если $\gamma \gg 1$, то $\lambda \ll \lambda_0$ (частота первичного излучения во много раз превышает частоту поперечных колебаний электронов). В открытом резонаторе, в который направляются релятивистские электроны и создаваемая ими первичная электромагнитная волна, происходит усиление волны и генерация направленного когерентного излучения. Такое излучение возникает из-за явления самосогласованного процесса, включающего в себя группирование электронов в сгустки под действием резонансной первичной волны. Группирование электронов возможно только при условии расходимости пучка, не превышающей несколько миллирадиан, и при условии моноэнергетичности электронов. Усиление происходит за счет когерентного излучения образовавшихся электронных сгустков.

В лазерах на свободных электронах есть возможность плавной перестройки длины волны генерации путем изменения величины кинетической энергии электронов. Так, получена генерация в инфракрасном диапазоне длин волн (10,8 мкм, 3,4 мкм, 0,65 мкм). Средняя мощность излучения составляет около 5 Вт. КПД составляет ~ 1%, но может быть увеличен до 40% при условии возврата электронов в резонатор.

Задачи и упражнения

3.1. Лазер с модулированной добротностью

Оценить основные параметры импульса излучения оптического квантового генератора при мгновенном включении добротности: пиковую мощность, время τ_1 , в течение которого плотность энергии в кристалле достигает половины максимального значения, а также τ_2 , в течение которого плотность энергии в кристалле возрастет от максимального значения и затем потом снова упадет до половины максимального значения.

Оценить энергию, высвечиваемую в импульсе за время τ_2 .

Для численных оценок взять следующие исходные данные: длина кристалла $L = 5$ см, плотность активных частиц в кристалле $n = 2 \times 10^{19}$ см $^{-3}$, частота излучения $\nu = 4 \times 10^{14}$ Гц, время спонтанного перехода с верхнего рабочего уровня на нижний $\tau_{sp} = 3 \times 10^{-3}$ с.

Рабочие уровни считать невырожденными. Принять, что при полной инверсии, создаваемой в кристалле источником накачки, коэффициент квантового усиления равен $\approx 0,4$ см $^{-1}$. Считать потери в кристалле, включая потери за счет излучения через зеркала (основные потери), распределенными и коэффициент потерь равным $\alpha_0 \approx 0,13$ см $^{-1}$.

Решение

Работу оптического квантового генератора опишем кинетическими уравнениями вида:

$$\left. \begin{aligned} \frac{dN}{dt} &= -\frac{N}{\tau_p} + \alpha_B c \frac{n_V}{n_{V_0}} N + \alpha_B c \frac{n_V^B}{n_{V_0}} \\ \frac{dn_V}{dt} &= -2\alpha_B c \frac{n_V}{n_{V_0}} N \end{aligned} \right\}, \quad (3.1.1)$$

где N — число фотонов в типе колебаний; n_V — разность чисел активных частиц на верхнем и нижнем рабочих уровнях во всем объеме кристалла; n_{V_0} — та же разность при полной инверсии (когда все активные частицы находятся на верхнем рабочем уровне); α_B — коэффициент квантового усиления при $n_V = n_{V_0}$; τ_p — время жизни фотонов в резонаторе. Эта система несколько отличается от ранее полученной системы.

Во-первых, коэффициент квантового усиления записан в виде $\alpha_B \frac{n_V}{n_{V_0}}$, т. е. нормирован на n_{V_0} .

Во-вторых, во втором уравнении имеется множитель 2, появление которого связано с тем, что если нижний рабочий уровень нельзя считать пустым, то переход активной частицы с верхнего рабочего уровня на нижний приводит к появлению одного фотона, но разность количества активных частиц уменьшается на две единицы (число частиц на верхнем уровне уменьшается на единицу, а число частиц на нижнем уровне увеличивается на единицу). В-третьих, из уравнения для изменения разностей чисел на уровнях исключены члены, определяющие действие источника накачки на влияние спонтанного излучения, т. к. это уравнение понадобится для описания поведения лазера после мгновенного включения добротности, когда этими членами можно пренебречь ввиду их малости. Кроме того, по-другому записан член, описывающий изменение числа фотонов в типе колебаний за счет спонтанного излучения [сравни с выражением (1.1.5) задачи 1.1].

Ометим также, что в уравнение (3.1.1) входит скорость света в вакууме. Вообще говоря, в эти уравнения должна была входить скорость света в кристалле, но ввиду грубости производимых в дальнейшем оценок считаем ее равной c .

В системе уравнений (3.1.1) перейдем к новым переменным. Введем обозначение $\alpha_0 = \frac{1}{\tau_p c}$, чтобы записать первый член правой части уравнения для числа фотонов в такой же форме, как член, описывающий индуцированные переходы. Вместо разности числа активных частиц перейдем к инверсной населенности

$$n = \frac{n_V}{V}, \quad n_0 = \frac{n_{V_0}}{V}, \quad n^B = \frac{n_V^B}{V},$$

а вместо числа фотонов — к плотности энергии в кристалле

$$\rho = \frac{Nh\nu}{V}.$$

Тогда система уравнений (3.1.1) примет вид:

$$\left. \begin{aligned} \frac{d\rho}{dt} &= -\alpha_0 c \rho + \alpha_B c \rho \frac{n}{n_0} + \alpha_B c \frac{n_B}{n_0} \frac{h\nu}{V} \\ \frac{dn}{dt} &= -\frac{2\alpha_B c}{h\nu} \rho \frac{n}{n_0} \end{aligned} \right\}. \quad (3.1.2)$$

При выключенной добротности потери в лазере велики (для простоты принимаем $\alpha_0 \gg \alpha_B$), генерации нет, n и n_B определяются сигналом накачки и перед включением добротности $n = n_B = n_0$, т. е. все частицы с нижнего рабочего уровня переброшены на верхний, а плотность энергии ρ_0 в типе колебаний определяется спонтанным излучением. Тогда из первого уравнения системы (3.1.2) имеем

$$\rho_0 = \frac{\alpha_B}{\alpha_0 - \alpha_B} \frac{h\nu}{V} \approx \frac{\alpha_B}{\alpha_0} \frac{h\nu}{V}. \quad (3.1.3)$$

После мгновенного включения достаточно высокой добротности коэффициент потерь α_0 резко уменьшается и $\alpha_B \gg \alpha_0$. Плотность энергии быстро нарастает за счет индуцированных переходов, и вклад спонтанного излучения становится ничтожным. Поэтому в первом уравнении системы (3.1.2) можно пренебречь последним членом, описывающим вклад спонтанного излучения в плотность энергии, и поведение лазера будет определяться следующей системой уравнений:

$$\left. \begin{aligned} \frac{d\rho}{dt} &= -\alpha_0 c \rho + \alpha_B c \rho \frac{n}{n_0} \\ \frac{dn}{dt} &= -\frac{2\alpha_B c}{h\nu} \rho \frac{n}{n_0} \end{aligned} \right\}. \quad (3.1.4)$$

Умножая второе уравнение системы (3.1.4) на $h\nu/2$ и складывая с первым уравнением, получаем

$$\frac{d}{dt} \left(\rho + \frac{h\nu}{2} n \right) = -\alpha_0 \nu \rho.$$

Подставляя в правую часть этого равенства выражение для ρ из второго уравнения системы (3.1.4), имеем

$$\frac{d}{dt} \left(\rho + \frac{h\nu}{2} n \right) = \frac{\alpha_0}{2\alpha_B} h\nu n_0 \frac{d}{dt} \ln n. \quad (3.1.5)$$

Решение уравнения (3.1.5) записывается в виде

$$\left(\rho + \frac{h\nu}{2}n\right) = \frac{\alpha_0}{2\alpha_B} h\nu n_0 \ln n + \text{const},$$

где константа определяется из начальных условий: при $t = 0$ (момент мгновенного включения добротности) все активные частицы находятся на верхнем рабочем уровне $n = n_0$, а плотность энергии определяется спонтанным излучением и равна ρ_0 [см. формулу (3.1.3)].

Используя эти условия, получаем

$$\rho - \rho_0 = \frac{h\nu}{2}(n_0 - n) + \frac{\alpha_0}{2\alpha_B} n_0 h\nu \ln \frac{n}{n_0}. \quad (3.1.6)$$

Оценим сначала пиковую мощность излучения.

Плотность энергии максимальна в точке $d\rho/dt = 0$. Из первого уравнения системы (3.1.4) видно, что $d\rho/dt = 0$ при

$$n = n_0 \frac{\alpha_0}{\alpha_B}. \quad (3.1.7)$$

Подставляя (3.1.7) в (3.1.6), находим

$$\rho_{\max} - \rho_0 = \frac{h\nu}{2} n_0 \left[1 - \frac{\alpha_0}{\alpha_B} + \frac{\alpha_0}{\alpha_B} \ln \frac{\alpha_0}{\alpha_B} \right]. \quad (3.1.8)$$

Для оптического квантового генератора, работающего в режиме модуляции добротности, $\rho_{\max} \gg \rho_0$, поэтому в формуле (3.1.8) величиной ρ_0 можно пренебречь по сравнению с ρ_{\max} . Практически во всех производимых ниже оценках $\rho \gg \rho_0$ и всюду в разности $\rho - \rho_0$ величиной ρ_0 можно пренебречь.

Мощность выходного излучения равна

$$P = \alpha c L S \rho_{\max},$$

т. е. пиковая мощность составляет

$$P_{\max} = \alpha c L S \rho_{\max} = \frac{n_0 c L S h\nu \alpha_B}{2} \frac{\alpha_0}{\alpha_B} \left[1 - \frac{\alpha_0}{\alpha_B} + \frac{\alpha_0}{\alpha_B} \ln \frac{\alpha_0}{\alpha_B} \right]. \quad (3.1.9)$$

Величина P_{\max} , как видно из (3.1.9), зависит от отношения α_0/α_B .

Исследование уравнения (3.1.9) на экстремум показывает, что экстремум P_{\max} достигается при $\alpha_0/\alpha_B \approx 0,3$. В условии задачи значения α_0 и α_B указаны как раз такие, что $\alpha_0/\alpha_B \approx 0,3$, т. е. отношение параметров α_0 и α_B выбрано оптимальным.

Вычислим теперь интервалы τ_1 и τ_2 . Прежде всего, нетрудно определить значение инверсной населенности $n = n_1$, при которой плотность энергии поля достигает половины своего максимального значения:

$$\rho = \rho_{\max}/2.$$

Подставляя в левую часть (3.1.6) $\rho_{\max}/2$ из выражения (3.1.8), получаем уравнение для определения n_1

$$\frac{n_1}{n_0} - \frac{\alpha_0}{\alpha_B} \ln \frac{n_1}{n_0} = \frac{1}{2} + \frac{\alpha_0}{2\alpha_B} - \frac{\alpha_0}{2\alpha_B} \ln \frac{\alpha_0}{\alpha_B}. \quad (3.1.10)$$

Это трансцендентное уравнение имеет два решения. Одно соответствует нарастающей, а другое — спадающей части импульса. При $\alpha_0 / \alpha_B \approx 0,3$ для нарастающей части импульса имеем $n'_1 / n_0 \approx 0,7$, а для спадающей $n''_1 / n_0 \approx 0,1$.

Для оценки τ_1 воспользуемся первым уравнением системы (3.1.4), считая $n = \text{const}$ (n меняется от n_0 до $0,7n_0$). Разделим левую и правую части уравнения на ρ и получим легко интегрируемое выражение

$$\frac{dp}{dt} / p = \frac{d}{dt} \ln p = \left(\alpha_B \frac{n}{n_0} - \alpha_0 \right) c \Big|_{n=\text{const}},$$

откуда

$$t = \frac{\ln p / p_0}{\left(\alpha_B \frac{n}{n_0} - \alpha_0 \right) c}.$$

В момент $t = \tau_1$ $\rho = \rho_{\max} / 2$, следовательно,

$$\tau_1 = \frac{\ln \frac{p_{\max}}{2p_0}}{\left(\alpha_B \frac{n}{n_0} - \alpha_0 \right) c}. \quad (3.1.11)$$

Интервал значений τ_1 получаем, если принять для n два значения: $n = n_0$, $n = n'_1$, т. е.

$$\frac{\ln \left(\frac{p_{\max}}{2p_0} \right)}{(0,7\alpha_B - \alpha_0)c} < \tau_1 < \frac{\ln \left(\frac{p_{\max}}{2p_0} \right)}{(\alpha_B - \alpha_0)c}. \quad (3.1.12)$$

При оценке интервала τ_2 следует иметь в виду, что на этом отрезке времени инверсная населенность изменяется очень сильно, а плотность энергии несколько меньше. Поэтому грубую оценку τ_2 можно произвести из второго уравнения системы (3.1.4), считая $\rho = \text{const}$. Это уравнение легко интегрируется:

$$\tau_2 = \frac{n_0 h \nu}{2\alpha_B c p} \ln \frac{n(t = \tau_1, \tau_2 = 0)}{n(t = \tau_1 + \tau_2)}. \quad (3.1.13)$$

Таким образом, τ_2 определяется отношением $n'_1 = n(t = \tau_1, \tau_2 = 0)$ к $n''_1 = n(t = \tau_1 + \tau_2)$. Для проводимой грубой оценки можно взять два значения постоянной плотности энергии: $\rho = \rho_{\max} / 2$ и $\rho = \rho_{\max}$. Тогда

$$\frac{n_0 h \nu}{2\alpha_B c p_{\max}} \ln \frac{n'_1}{n''_1} < \tau_2 < \frac{n_0 h \nu}{\alpha_B c p_{\max}} \ln \frac{n'_1}{n''_1}. \quad (3.1.14)$$

Пользуясь условиями задачи, проведем численные оценки формул (3.1.9), (3.1.12), (3.1.14):

$$\begin{aligned} P_{\max} &\approx 2 \times 10^{10} \text{ Вт}, \\ 5 \times 10^{-9} \text{ с} &< \tau_1 < 10^{-8} \text{ с}, \\ 4 \times 10^{-10} \text{ с} &< \tau_2 < 8 \times 10^{-10} \text{ с}. \end{aligned}$$

Оценим также энергию W , высвечиваемую в импульсе за время τ_2 . Очевидно, что

$$W = \frac{1}{2}(n_1'' - n_1'')h\nu LS \approx 7 \text{ Дж.}$$

Контрольные вопросы

1. Что такое твердотельный лазер?
2. Нарисуйте типичную функциональную схему лазера.
3. Какие активные среды твердотельных лазеров вы знаете? Охарактеризуйте их свойства.
4. Что такое полупроводниковый лазер?
5. Опишите процесс взаимодействия полупроводников с излучением.
6. Что такое инжекционный лазер и как он устроен?
7. Что такое гетеролазер и как он устроен?
8. Что такое жидкостной лазер и как он устроен?
9. Какие типы жидкостных лазеров вы знаете?
10. Что такое газовый лазер?
11. Какие типы газовых лазеров вы знаете?
12. Опишите конструкцию и дайте характеристику He-Ne лазера.
13. Что такое ионный лазер?
14. Что такое молекулярный лазер?
15. Что такое эксимерный лазер?
16. Что такое химический лазер?
17. Что представляет собой лазер на свободных электронах?
18. Какие лазеры на парах металла вы знаете?

□□

Рекомендуемая литература

1. Зверев Г. М., Голяев Ю. Д. Лазеры на кристаллах и их применение. — М.: Рикел, Радио и связь, 1994.
2. Карлов Н. В. Лекции по квантовой электронике. Учебное пособие. — М.: Наука, 1983.
3. Курбатов Л. Н. Оптоэлектроника видимого и инфракрасного диапазона спектра. — М.: Изд. МФТИ, 1999.
4. Пихтин А. Н. Физические основы квантовой электроники и оптоэлектроники. Учебное пособие. — М.: Высшая школа, 1983.
5. Справочник по лазерам. В двух томах. Под ред. акад. А. М. Прохорова. Т. 1. — М.: Советское радио, 1978.
6. Успенский А. В. Сборник задач по квантовой электронике. Учебное пособие. — М.: Высшая школа, 1976.

4. Оптические волноводы

Информационные системы, в основе которых лежат квантовые структуры, включают в себя следующие основные элементы:

- ☐ волноводы, представляющие собой континуальную среду для передачи световой информации к приемнику излучения;
- ☐ устройство управления световыми потоками, с помощью которых в световой луч может быть внесена необходимая информация;
- ☐ приемники излучения, позволяющие детектировать сигнал и расшифровывать всю записанную в световом пучке информацию.

Рассмотрим все эти элементы квантовых систем обработки и хранения информации.

4.1. Плоские волноводы

Передача информации от излучателей (генераторов) света к приемникам излучения осуществляется с помощью волноводов.

Волноводы представляют собой искусственный или естественный канал, способный поддерживать распространяющиеся вдоль него волны, поля которых сосредоточены внутри канала или в примыкающей к нему области. Основным свойством волноводов является существование в нем дискретного набора нормативных волн или мод, распространяющихся со своими фазовыми и групповыми скоростями.

Под *фазовой скоростью* будем понимать скорость перемещения фазы волны в определенном направлении. В случае плоской монохроматической волны фазовые фронты или плоскости постоянной фазы отличаются от волнового вектора, и могут в некоторых случаях превышать скорость распространения света. Существует зависимость фазовой скорости от частоты, которая определяет дисперсию волн. Дисперсия может привести к искажению формы передаваемого сигнала.

Групповая скорость волн представляет собой движение группы или дуга волн, образующих в каждый момент времени локализованный в пространстве волновой пакет.

В линейных средах соблюдается принцип суперпозиции и поэтому волновой пакет представляет собой набор гармонических волн. Длина пакета Δl и ширина спектра $\Delta\omega$ ограничены снизу соотношением

$$\Delta l \cdot \Delta k \geq 1,$$

где волновое число k связано с частотой ω дисперсионным соотношением $\omega = \omega(k)$.

Групповая скорость определяет скорость переноса энергии волнами. Существующие методы измерения скоростей распространения волн дают значение групповой скорости.

Согласно теории относительности групповая скорость не может превышать скорость распространения света в вакууме.

Механизм *канализации* электромагнитных волн в волноводах основывается на явлении полного внутреннего отражения. Если на границу двух прозрачных сред с показателями

преломления n_1 и n_2 из среды с большим показателем преломления, например, $n_1 > n_2$, под углом $\varphi \geq \varphi_{кр}$ падает свет, то происходит полное внутреннее отражение. При этом величина $\varphi_{кр}$ такова, что выполняется соотношение

$$\sin \varphi_{кр} = n_2 / n_1 = n_{21}. \quad (4.1)$$

Обычно волновод реализуется между подложкой и покровным материалом. Такая структура представлена на рис. 4.1 и характеризуется показателем преломления канала n_k , подложки n_n , покровного материала n_0 . Необходимо, чтобы выполнялось соотношение

$$n_k > n_n > n_0.$$

Существуют два критических угла, благодаря которым свет может каналироваться: угол полного внутреннего отражения φ_n на границе раздела "канал—подложка" и угол полного внутреннего отражения φ_0 на границе "канал—покровный слой". В противном случае имеют место потери света в подложке и покровном материале, который пронизывает канал и выходит через покровный материал (рис. 4.1, а). Такая электромагнитная мода называется излучательной.

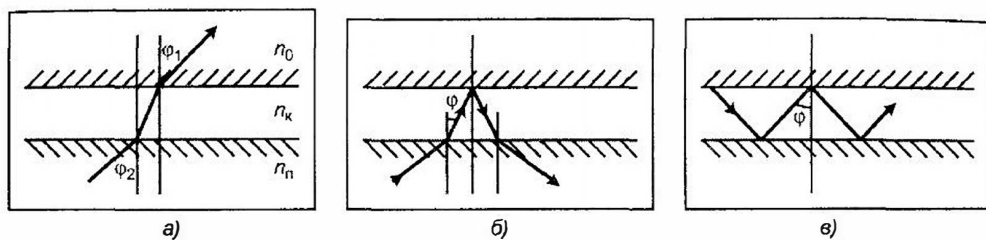


Рис. 4.1. Распространение волн в плоском волноводе: а — излучательные моды; б — излучательные моды подложки; в — волноводные моды

Если увеличить угол падения света так, чтобы выполнялось соотношение

$$\varphi_0 < \varphi < \varphi_n,$$

то распространяющийся свет отражается от границы с покровным материалом и преломляется обратно в подложку, и в этом случае говорят об излучательной моде подложки и речь не идет о волноводном распространении света (рис. 4.1, б).

Если увеличить угол φ так, чтобы выполнялось условие полного внутреннего отражения от подложки, то наблюдается полное внутреннее отражение как от границы раздела "канал—подложка", так и от границы раздела "канал—покровный материал".

Теперь световая волна, попавшая в канал, будет распространяться в нем по зигзагообразному пути. Эта волна называется *волноводной модой* (рис. 4.1, в). Волны света испытывают полное внутреннее отражение на границах канала.

Волны монохроматичны и когерентны, их угловая частота равна ω , длина волны λ , а волновой вектор в направлении нормали к волновой поверхности равен kn_k , так что

$$k = 2\pi / \lambda = \omega / c,$$

где c — скорость света в вакууме. Постоянная распространения β для волновой моды в плоском волноводе определяется выражением:

$$\beta = \omega / V_{ф} = kn_k \sin \varphi, \quad (4.2)$$

где $V_{ф}$ — фазовая скорость.

Значение β лежит в пределах $kn_n < \beta < kn_k$.

Введем понятие эффективного волноводного показателя преломления как величину $N = \beta / k = n_k \sin \phi$, который лежит в пределах

$$n_n < N < n_k. \quad (4.3)$$

На рис. 4.2 представлены различные режимы распространения мод в волноводном канале и соответствующее распределение полей мод $E(x)$.

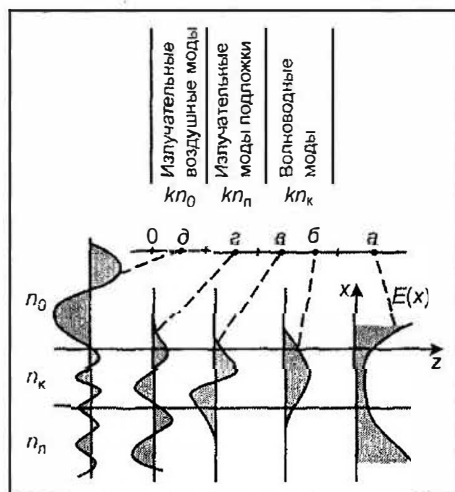


Рис. 4.2. Распространение полей мод при различной постоянной распространения β

При $\beta > kn_k$ распространение моды имеет экспоненциальный характер. Такой тип колебаний не имеет практического применения, и поле неограниченно растет в подложке и в покровном слое (рис. 4.2, а).

При условии $kn_n < \beta < kn_k$ мода имеет поля, представленные на рис. 4.2, б, в. Это гармоническая колебательная функция с максимумом распространения по оси канала или с образованием в поперечном сечении волновода стоячей волны. Это волноводные моды.

Излучательная мода подложки реализуется при условии $kn_0 < \beta < kn_n$. Они могут существовать в волноводной структуре, однако быстро затухают из-за перекачки из канала в подложку.

Если постоянная распространения лежит в интервале $0 < \beta < kn_n$, то мода колебаний характерна для всех областей (рис. 4.2, д). Это соответствует излучательным воздушным модам, которые быстро затухают.

Этот тип мод не используется для передачи сигнала. Проведенный анализ показывает, как важно правильно ввести пучок света в планарный волновод через элемент связи. Различают *поперечные* и *продольные* элементы связи. В поперечных элементах связи пучок фокусируется на открытое поперечное сечение волновода. В эту категорию входят элементы связи прямого действия.

На рис. 4.3 приведена возможная схема ввода излучения в волновод и преобразования светового пучка в поверхностную волну. Преобразование энергии пучка в поверхностную волну происходит путем согласования электромагнитного пучка с полем поверхностной волны. Распределение интенсивности по сечению волновода соответствует моде TE_0 .

Очевидны потери световой энергии на границах планарного волновода. Поэтому поперечная схема элементов связи не нашла широкого применения.

В оптических системах волновод принято называть световодом, который предназначен для направленной передачи света.

В продольных элементах связи пучок света падает наклонно по отношению к световоду. К этим элементам связи относят, прежде всего, призмы и решетчатые элементы.

В случае использования *призменного* элемента для ввода излучения пучок света падает на призму, показатель преломления которой n_p (рис. 4.4). Угол падения выбирается таким, чтобы он удовлетворял условию полного отражения:

$$\theta > \theta_{\text{полн}} = \arcsin(n_0 / n_p).$$

Одновременно угол падения должен удовлетворять условию согласования фаз:

$$k_p \cdot \sin \theta = \beta_m \quad (4.4)$$

где k_p — волновой вектор, β_m — постоянная распространения поверхностной волны, L — длина связи призмы и световода.

Полный обмен энергией может быть осуществлен при реализации условия

$$\zeta L = \pi / 2,$$

где ζ — коэффициент связи между излучательной модой и модой поверхностной волны.

Если высота воздушного промежутка h_0 достаточно высока, то между модами света в призме и волновода существует слабая связь. При незначительных размерах промежутка h_0 энергия пучка призмы передается в канал благодаря эффекту оптического туннелирования света. Этот эффект вызван нарушением полного внутреннего отражения. При этом должно выполняться условие согласования фаз между модами призмы и канала. Перекачка энергии из излучательной моды, совпадающей с полем пучка, может происходить при условии

$$L \cdot \cos \theta = 2w,$$

где w — ширина (апертура) пучка падающего излучения.

Исследованное свойство обратимости световых пучков можно вывести излучение из волновода через призму. Таким образом, призменный элемент связи позволяет осуществлять преобразования пучок света \rightarrow поверхностная волна и поверхностная волна \rightarrow пучок света.

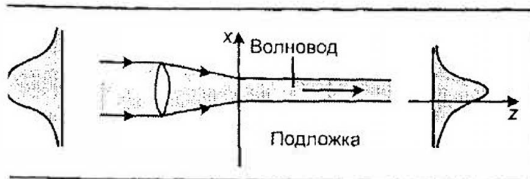


Рис. 4.3. Схема поперечного элемента связи

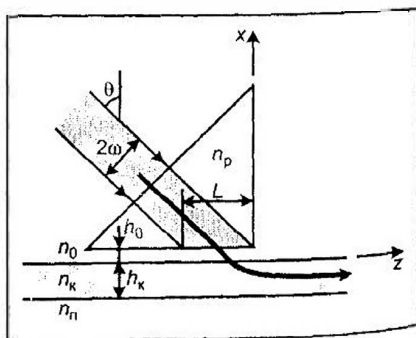


Рис. 4.4. Схема призменного элемента связи

Недостатком призмных элементов связи является требование высокого значения показателя преломления $n_p > n_s$ материала призмы. Например, для волноводов из GaAs, имеющих достаточно высокий показатель преломления, трудно подобрать призму с еще более высоким показателем преломления.

Другая трудность заключается в настройке воздушного зазора h_0 , толщина которого обычно меньше половины длины световой волны. К тому же в зазоре возможно скопление частиц пыли, что резко увеличивает потери световой энергии.

Решеточный элемент связи работает аналогично призмному элементу. Однако в этом случае отсутствует воздушный зазор. Решеточный элемент связи обладает периодической структурой (рис. 4.5). При падении на решетку волны рождаются гармоники, локализованные в области решетки.

Продольные постоянные распространения этих гармоник равны:

$$\beta_v = \beta_0 + 2v\pi / d, \quad (4.5)$$

где v — индекс, $0, \pm 1, \pm 2$; β_0 — постоянная распространения поверхностной волны, $\beta_0 \approx k_0 = 2\pi / \lambda_0$, где λ_0 — длина волны падающего излучения; d — период решетки.

Благодаря отрицательным значениям v возможны значения $\beta_v \leq \beta_0$, при которых будет удовлетворяться уровень согласования фаз $k_0 \sin \theta = \beta_v$, что выполняется при соответствующих значениях v, d, λ, θ .

Поле поверхностных волн состоит из гармоник и, следовательно, энергия, передаваемая пучком одной из гармоник, связана с основной гармоникой. Эта гармоника распространяется в канале и превращается в поверхностную волну.

Решетка может быть изготовлена в виде пленки фоторезиста, предварительно экспонированной интерференционной картиной волн. В зависимости от фоторезиста и способов его обработки решетка может иметь синусоидальную, трапецеидальную или треугольную формы.

Основной недостаток решетчатого элемента связи состоит в том, что значительная часть световой волны гасится в подложке вследствие многократного прохождения через решетку. Это ограничивает эффективность связи, которая меньше, чем у призмных элементов связи.

Существуют и другие элементы связи, например, голографического типа. Однако они применяются редко.

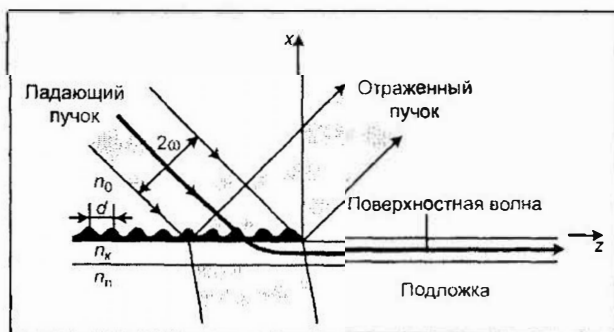


Рис. 4.5. Схема решетчатого элемента связи

Волноводы могут быть выполнены на пассивных и активных подложках.

Для выполнения *пассивных* функций используются волноводы на стеклянных подложках. Это просто и экономично. Одним из методов получения волноводных структур является легирование стекла путем ионного обмена в электрическом поле между ионами стекла и ионами легирующей примеси. Таким способом создаются волноводные структуры различной конфигурации. Возможно использование методов ионной имплантации. Этим методом удастся получить изменение коэффициента преломления на 10—15% и снизить оптические потери до 0,1 дБ/см. Для пассивных волноводных структур используются тонкие пленки, полимерные пленки.

Активные волноводные структуры формируются на основе активных диэлектриков, например, кристаллов ниобата лития, титаната лития.

Оптическими свойствами активных диэлектриков и соответствующих волноводных структур можно управлять с помощью внешних электромагнитных полей.

Волноводные каналы в активных диэлектриках формируются с помощью процессов диффузии ионов, имплантации, эпитаксии. Например, широко распространена структура волновода на основе легирования титаном ниобата лития $Ti : LiNbO_3$.

К активным волноводам относятся также структуры, полученные на полупроводниковой основе, например, на соединениях типа A^3B^5 , A^2B^4 . Необходимый перепад показателя преломления в таких структурах достигается за счет многослойных структур с плавным или ступенчатым изменением состава при переходе от слоя к слою.

Одновременно можно менять концентрацию свободных носителей в полупроводнике и создавать области с различными показателями преломления.

В табл. 4.1 приведены некоторые параметры материалов для волновых структур.

Таблица 4.1. Сравнительные характеристики волноводных структур

Материал волноводного слоя	Материал подложки	Показатель преломления волноводного слоя на длине волны λ (нм)	Способ получения волноводного слоя	Средний уровень потерь на длине волны (мкм), дБ/см
Стекла ТК8-ТК21	Стекло	1,56—1,66 (0,63)	ВЧ-распыление	1 (0,63)
NB_2O_6	"	2,27 (0,63)	"	0,01 (0,63)
ZnO	"	2,00 (0,63)	"	0,01 (0,63)
Ta_2O_5 , $Ta_2O_5N_x$	"	1,93—2,15 (0,63)	"	0,9 (0,63)
Винилтриметилсилан	"	1,52—1,53 (0,63)	Полимеризация в плазме глосю-щего разряда	0,9 (0,63)
Желатин	"	1,54 (0,63)	Осаждение из водного раствора	0,5 (0,63)
$Ti : LiNbO_3$	$LiNbO_3$	2,21 (0,63)	Диффузия	0,5 (0,63)
$LiNbO_3$	$LiTaO_3$	2,20 (0,63)	Эпитаксия	1,0 (0,63)
$Ti : LiNbO_3$	$MgO : LiNbO_3$	2,21 (0,63)	Диффузия	0,8 (0,63)

Таблица 4.1 (окончание)

Материал волноводного слоя	Материал подложки	Показатель преломления волноводного слоя на длине волны λ (нм)	Способ получения волноводного слоя	Средний уровень потерь на длине волны (мкм), дБ/см
H: LiNbO ₃	LiNbO ₃	2,35 (0,63)	Протонный обмен	0,5 (0,63)
LiTaO ₃	LiTaO ₃	2,21 (0,63)	Ионный обмен	2,0 (0,63)
PLZT-керамика	Al ₂ O ₃	2,60 (0,63)	ВЧ-распыление	5,0 (1,32)
GaAs	Al _x Ga _{1-x} As	3,60 (0,82)	Эпитаксия из молекулярного пучка	0,2 (1,30)
CdS _{1-x} Se _x	CdS	2,46—2,50 (0,63)	Эпитаксия	9,0 (0,63)
Cd _{1-x} Zn _x S	ZnS	2,37 (0,63)	Диффузия	3,0 (0,63)
Al _x Ga _{1-x} As	Al _y Ga _{1-y} As	3,10—3,37 (1,15)	Эпитаксия	4,0 (1,15)
Al _x Ga _{1-x} As	GaAs	3,20 (1,15)	» »	1,7 (1,15)
Ga _x In _{1-x} As, P _{1-x}	InP	3,37 (1,32)	» »	5,0 (1,32)
GaAs	GaAs	3,60 (1,15)	» »	1,3 (1,15)

Полупроводниковые волноводы, как правило, изготавливают из тех же материалов, что и интегрируемые с ними инжекционные лазеры и фотоприемники.

4.2. Волоконные световоды

Волоконный световод представляет собой тонкую нить из оптически прозрачного материала с радиусом R и показателем преломления n_2 , внутри которого расположена сердцевина с радиусом r и показателем преломления $n_1 > n_2$.

Если в открытом пространстве передача световой энергии происходит одновременно по всем направлениям в пределах прямой видимости и ограничивается расходимостью, поглощением и рассеянием света, то применение световодов позволяет передавать световую энергию на большие расстояния по криволинейным трассам.

Разработаны различные типы световодов: линзовые, зеркальные, полые трубы и т. п.

Наибольшее распространение получили гибкие диэлектрические световоды с низкими оптическими потерями. Конструкция волоконного световода представлена на рис. 4.6.

В приближении геометрической оптики лучи, входящие на границу "оболочка—сердцевина" волокна по углам $\theta \leq \theta_{кр}$, испытывают полное внутреннее отражение на поверхности раздела двух оптических сред (рис. 4.6, б). Величина критического угла $\theta_{кр}$ определяется из соотношения:

$$\sin \theta_{кр} = \frac{1}{n_1} \sqrt{n_1^2 - n_2^2}. \quad (4.6)$$

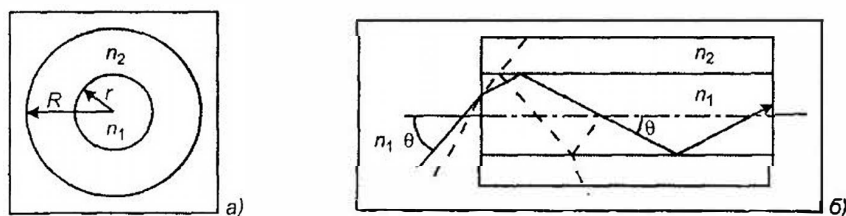


Рис. 4.6. Конструкция волоконного световода (а) и траектория распространения лучей (б)

Свет распространяется по световоду по зигзагообразной траектории.

Величина угла $n_1 \theta_{кр} = \sqrt{n_1^2 - n_2^2}$ является мерой способности захватить световую энергию.

Синус угла $\theta_{кр}$ называется *числовой апертурой световолокна*.

Характер распространения оптического излучения по световолокну зависит от его поперечных размеров и профиля показателя преломления по сечению. Число мод, которые могут распространяться по световолокну, зависит от квадрата диаметра сердцевинки и разности показателей преломления $\Delta n = n_1 - n_2$. Число мод N , которое можно передать по световолокну, в первом приближении определяется из соотношения

$$N \approx \frac{2\pi^2 r^2}{\lambda^2} (n_1^2 - n_2^2). \quad (4.7)$$

На рис. 4.7 приведены типичные поперечные сечения и профиль распределения показателя преломления по сечению. Материалом для оптических световодов обычно служит кварцевое стекло. Различные показатели преломления достигаются путем легирования стекла фтором, германием, фосфором и др.

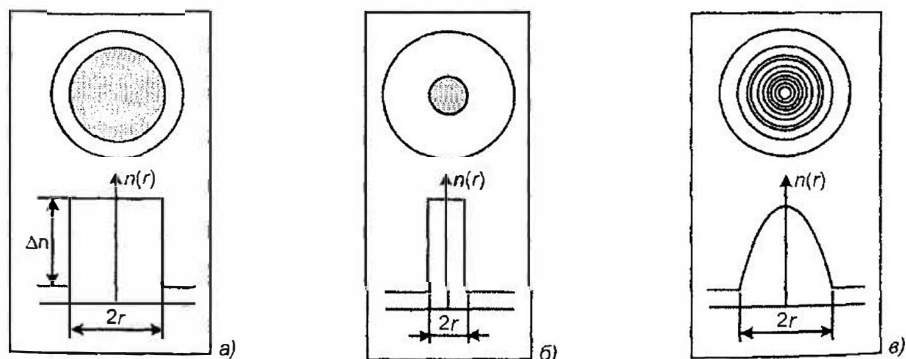


Рис. 4.7. Сечение и профиль показателя преломления по сечению многомодового ступенчатого (а), одномодового (б) и многомодового градиентного (в) световодов

Для кварцевого стекла полоса пропускания световодов имеет максимальное значение на длине волны 1,3 мкм и составляет $\sim 10^{11}$ Гц·км.

Изменением профиля показателя преломления можно сместить нуль дисперсии в область длин волн 1,55 мкм, где расположен минимум оптических потерь.

На рис. 4.8, а приведена схема изготовления оптических волокон. Стекло плавят в кремниевых тиглях с отверстиями в нижней части. В каждом тигле находится стекло с соот-

ветствующим показателем преломления. Волокно вытягивается из тигля и наматывается на барабан. На основе оптических волокон изготавливаются кабели для оптоволоконной связи (рис. 4.8, б, в).

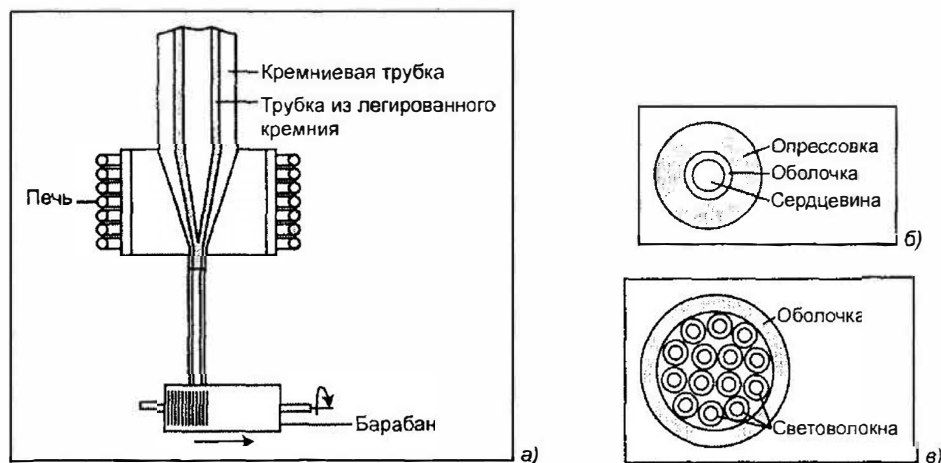


Рис. 4.8. Схема изготовления оптического волокна по методу Корнинга (а) и структура оптического кабеля с одним (б) и многими (в) волокнами

При производстве светового волокна используются также материалы, прозрачные в видимой и ИК-областях спектра: бескислородные стекла и полимеры. Однако они все уступают кварцевому стеклу по прочностным характеристикам. Разработанные эрбиевые волоконные стекла работают в спектральной области вблизи 1,55 мкм, т. е. в области минимальных оптических потерь. Потери при передаче информации по световому волокну незначительны и составляют доли дБ/км в области длин волн 1—1,5 мкм.

В мировой практике используются волоконные линии дальней связи со скоростью передачи сигналов более 10 Гбит/с.

Контрольные вопросы

1. Что такое волновод?
2. Что такое групповая и фазовая скорости распространения сигнала?
3. Как происходит канализация электромагнитной волны в плоском волновом?
4. Какие продольные элементы связи вы знаете?
5. Расскажите о призмочном элементе связи. Какие у него достоинства и недостатки?
6. Расскажите о решетчатом элементе связи. Какие у него достоинства и недостатки?
7. Что такое активные волноводы?
8. Что такое волоконный световод? Опишите его конструкцию.
9. Как получить волоконный световод?
10. Какими свойствами обладает волоконный световод, определившие его широкое применение в системах связи и телекоммуникации.

Рекомендуемая литература

1. Интегральная оптика. Под ред. Т. Тамира. — М.: Мир, 1978.
2. Курбатов Л. Н. Оптоэлектроника видимого и инфракрасного диапазона спектра. — М.: Изд. МФТИ, 1999.
3. Клэр Ж. Ж. Введение в интегральную оптику. — М.: Советское радио, 1980.
4. Пихтин А. Н. Физические основы квантовой электроники и оптоэлектроники. Учебное пособие. — М.: Высшая школа, 1983.
5. Свечников Г. С. Интегральная оптика. — Киев, Наукова думка, 1988.

5. Устройства управления световыми потоками

Световой пучок становится информативным, если им можно управлять во времени и в пространстве, а также если можно ввести информацию в световую волну.

Управление световым пучком путем механического переключения инерционно и ненадежно. Возможен процесс детектирования, электронного переключения и восстановления пучка. Такой процесс связан с потерями и дополнительными искажениями информации, а также уменьшением отношения сигнал/шум.

В настоящее время разработаны методы управления световыми потоками на основе электро- и магнитооптических явлений и эффектов: эффект Керра, эффект Штарка, эффект Поггеля, эффект Фарадея, явление образования фазовых дифракционных решеток в интерференционном поле интенсивных когерентных потоков и др.

На их основе созданы переключатели и сканеры световых потоков, различного типа модуляторы и дефлекторы, детекторы светового потока.

Дефлектор (от латинского *deflecto* --- отклоняю, отвожу) представляет собой устройство, предназначенное для изменения по заданному закону направления распространения луча в пространстве.

Вопрос переключения световых потоков в пространстве успешно решается с помощью электроно-оптических и акустооптических дефлекторов различной конструкции. Наибольшее распространение получили акустооптические дефлекторы, работа которых основана на явлении дифракции света на фазовой ультразвуковой решетке. Эти конструкции будут рассмотрены ниже.

Переключение оптического потока является операцией, с помощью которой целенаправленно изменяется пространственное положение когерентной световой волны.

Сканирование есть расширение понятия переключения, при котором осуществляется целенаправленное непрерывное или дискретное движение луча.

Непрерывная световая волна практически не несет информационного сигнала, равно как и регулярные импульсы света.

Модуляция является операцией, с помощью которой осуществляется ввод информации в световую волну.

Модулятор представляет собой устройство, которое управляет параметрами светового потока и изменяет детектируемые свойства световой волны в соответствии с приложенным электрическим сигналом.

Модулятор интенсивности представляет собой устройство, позволяющее изменять интенсивность когерентной световой волны в соответствии с изменяющимся во времени сигналом.

Например, плоская волна в форме

$$A = A_0 \exp[i(\omega t + \phi)]$$

имеет интенсивность

$$J = A_0(t) \cdot A_0^*(t) = A_0^2.$$

Величина $|A_0|^2$ является функцией приложенного сигнала.

Действие амплитудных модуляторов основывается на физических эффектах, связанных с изменением фазы, а не поглощением света, проходящего через модулятор. Наведенное изменение фазы обуславливает эквивалентное изменение интенсивности.

Фазовую задержку можно получить различными путями, например, поляризовав световую волну и пропустив ее через анизотропный элемент. Анизотропным элементом служат два идентичных кристалла, между которыми расположена полуволновая пластина. Ориентация полуволновой пластины такова, что поляризация проходящего через нее света поворачивается на угол $\pi/2$. Разность фаз за счет естественной анизотропии на выходе равна нулю. Меняя знак управляющего электрического поля при переходе от первого кристалла ко второму, можно навести сдвиг фазы в кристаллах (рис. 5.1).

Разработано более 100 типов амплитудных модуляторов поляризационного типа.

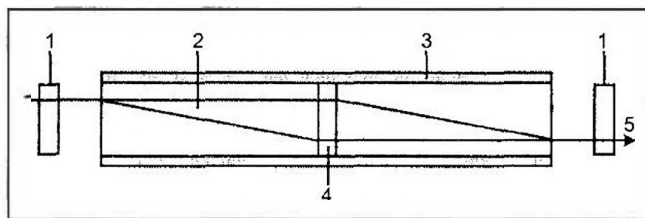


Рис. 5.1. Амплитудный модулятор света поляризационного типа: 1 — поляризатор и анализатор; 2 — электрооптические элементы; 3 — электроды; 4 — полуволновая пластина; 5 — световой пучок

В качестве электрооптических элементов обычно используют в световом диапазоне кристаллы АДР ($\text{NH}_4\text{H}_2\text{PO}_4$), КДР (KH_2PO_4), ДКДП (KD_2PO_4), LiNbO_4 , LiTaO_4 , $\text{Ba}_{12}\text{SiO}_{20}$, а в ИК-диапазоне арсенид галлия (GaAs) и теллурид кадмия (CdTe). Полуволновые напряжения этих материалов лежат в пределах 90—4000 В, полоса модуляции от 1 МГц до 1 ГГц.

Фазовые модуляторы используют явление линейного изменения показателя преломления кристаллов в зависимости от величины электрического поля E , приложенного к кристаллу. Это явление носит название *эффекта Поккельса*. Изменение показателя преломления можно выразить формулой:

$$n = n_0 + n_0^3 r E / 2,$$

где n_0 — показатель преломления кристалла в отсутствии поля E , r — электрооптический коэффициент.

Фазовый сдвиг зависит от длины кристалла l , поля E , длины волны света λ и имеет вид:

$$\varphi = \frac{2\pi n l}{\lambda} = \varphi_0 + \frac{\pi n_0^3 r E l}{\lambda},$$

где φ_0 — начальный сдвиг фаз, приобретенный световым пучком при прохождении кристалла в отсутствии поля. Наличие фазового сдвига, вызванного внешним электрическим полем, и означает фазовую модуляцию света. В качестве материалов для фазовых модуляторов используют кристаллы АДР, КДР, ДКДР, ниобат лития, титанат лития. Напряже-

ные на модуляторах находится в пределах 100—5000 В. Фазовые модуляторы обычно применяются в гетеродинном приеме, а также в качестве переключателя света.

Поляризационная модуляция также основана на использовании электрооптического эффекта. Вектор электрического поля волны можно записать в виде:

$$\vec{E} = \vec{x}E_x \cdot \exp[j(\omega t + \varphi_0)] + \vec{y}E_y \cdot \exp[j(\omega t + \varphi_0)], \quad (5.1)$$

где \vec{x} и \vec{y} — орты. Если $E_x = E_y$, то плоскость поляризации расположена под углом 45° к оси.

При приложении сигнала амплитуда волны примет вид:

$$\vec{E} = \vec{x}E_x \cdot \exp[j(\omega t + \varphi_0 + \varphi_x)] + \vec{y}E_y \cdot \exp[j(\omega t + \varphi_0 + \varphi_y)], \quad (5.2)$$

где φ_x и φ_y — добавочные фазы от приложенного сигнала. Сдвиг фаз $\Delta\varphi$ составит:

$$\Delta\varphi = \varphi_x - \varphi_y.$$

В общем случае речь идет об эллиптически поляризованной волне, у которой углы между главными осями эллипса и осью x зависят от $\Delta\varphi$.

Если $\Delta\varphi = \pi$, то речь идет о линейно поляризованной волне. При прохождении волны, описываемой выражением (5.2), через анализатор, ее интенсивность пропорциональна $\sin^2(\Delta\varphi/2)$.

Ранее рассматривался пример использования поляризационной модуляции в модуляторах на основе электрооптических эффектов.

Магнитооптические модуляторы света также основаны на эффекте вращения плоскости поляризации света. В этом типе модуляторов, как правило, используют эффект Фарадея. Его суть заключается во вращении плоскости поляризации линейного поляризованного света, распространяющегося в веществе вдоль постоянного магнитного поля. Дело в том, что под действием магнитного поля показатель преломления n^+ и n^- для циркулярно право- и левополяризованного света отличаются друг от друга. Следовательно, составляющие линейно поляризованного света распространяются с различными фазовыми скоростями, приобретая разность хода. В результате плоскость поляризации линейного поляризованного света с длиной волны λ , прошедшего путь l , поворачивается на угол

$$\theta = \frac{\pi l(n^+ - n^-)}{\lambda}.$$

Ввиду того, что величина $(n^+ - n^-)$ зависит от магнитного поля H , угол фарадеевского вращения выражается соотношением

$$\theta = VH,$$

где величина V называется постоянной Верде и зависит от свойств вещества.

Эффект Фарадея слабо зависит от температуры. В качестве магнитооптического элемента используются монокристаллы $\text{Y}_3\text{Fe}_5\text{O}_{12}$, C_2Br_4 , RbNiFe_3 и др. Это позволяет получать модуляторы света при магнитных полях ~ 10 А/М на частотах до 200 МГц при управляющей мощности до 0,1 Вт.

Частотные модуляторы используют явление сдвига частоты (длины волны) при приложении электрического поля. Для оптических частотных модуляторов возможен небольшой сдвиг частоты, поэтому детектирование такого сигнала вызывает значительные технические трудности. Обычно в этом случае используют сложный метод гетеродинирования света. Частотные модуляторы разработаны на акустооптических эффектах.

Рассмотрим теперь *акустооптические модуляторы и deflectоры*.

Показатель преломления вещества можно изменить не только с помощью электрического или магнитного полей, но и путем механической деформации. Это явление называется фотоупругостью или инзоптическим эффектом. Разность фаз возникает под действием механических напряжений на эффекте двулучепреломления. Напряжение в среде и, следовательно, изменение показателя преломления периодичны и повторяют период длины акустической волны.

Акустическая волна при распространении в оптически прозрачной среде формирует последовательные изменения показателя преломления. Образуется структура, аналогичная дифракционной решетке. Если период структуры меньше ширины светового пучка, то на ней наблюдается дифракция света.

Различают два режима дифракции света на звуковых волнах.

Физической основой режима Рамана — Ната является условие дифракции на звуковом столбе с образованием ряда дифракционных максимумов, углы которых выражаются как

$$\sin \theta_p = m \lambda / \Lambda,$$

где $m = 0, \pm 1, \pm 2, \pm 3 \dots$; λ — частота света; Λ — частота звуковой волны (рис. 5.2, а). В режиме Рамана — Ната должно выполняться условие

$$l \ll \Lambda^2 / \lambda.$$

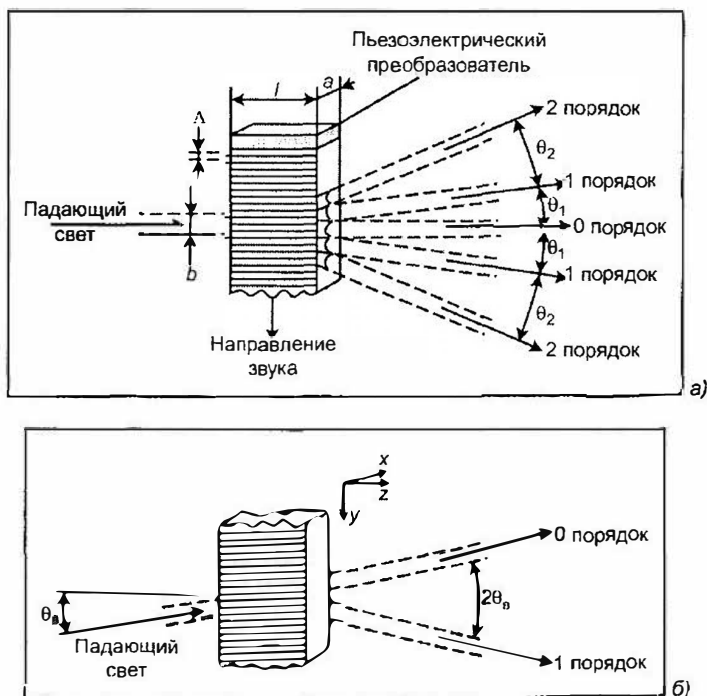


Рис. 5.2. Схема дифракции света на звуковой волне: а — режим Рамана — Ната; б — режим Брэгга

физической основой режима Брэгга является условие того, что дифрагированный свет из падающего пучка вновь сильно дифрагирует перед тем, как покинуть акустическое поле. В этом случае гасятся все порядки, за исключением одного первого. Брэгговский угол определяется зависимостью

$$\sin \theta_B = \lambda_0 / 2\Lambda.$$

В режиме дифракции Брэгга необходимо выполнение условия

$$l \gg \Lambda^2 / \lambda.$$

решетка Рамана-Ната называется тонкой решеткой, а акустическая решетка Брэгга называется толстой. На основе этих эффектов создаются модуляторы и дефлекторы света. В качестве материалов для акустооптических модуляторов на частотах до 250 МГц используются: тяжелый флинт ТФ7, кристаллы PbMoO_4 , TeO_2 , LiNbO_3 , в ИК-диапазоне Ge и GaAs.

Контрольные вопросы

1. Что такое дефлектор? Какие функции он выполняет?
2. Что такое модулятор? Какие функции он выполняет?
3. Какие типы модуляторов вы знаете?
4. Какие физические явления лежат в основе работы амплитудного модулятора?
5. Какие физические явления лежат в основе работы фазового модулятора?
6. Какие физические явления лежат в основе работы магнитооптического модулятора?
7. Расскажите об особенностях дифракции света на акустической волне.

Рекомендуемая литература

1. Интегральная оптика. Под ред. Т. Тамира. — М.: Мир, 1978.
2. Клэр Ж. Ж. Введение в интегральную оптику. — М.: Советское радио, 1980.
3. Свечников Г. С. Интегральная оптика. — Киев, Наукова думка, 1988.

6. Приемники излучения

Приемники излучения предназначены для регистрации сигналов в оптическом и ближнем инфракрасном диапазоне частот. В основе их работы лежат явления и эффекты, связанные с преобразованием оптического сигнала в электронный сигнал. Речь идет, прежде всего, о фотогальваническом (фотовольтаическом) эффекте.

Фотогальванический эффект связан с возникновением электрического тока (фототока) при освещении полупроводника, включенного в замкнутую цепь, или возникновением ЭДС на освещаемом образце при разомкнутой внешней цепи (фото-ЭДС).

Если выходной электрический сигнал возникает при преобразовании энергии фотона в первичную реакцию фотоприемника, то он относится к классу фотонных или квантовых фотоприемников.

Все другие эффекты преобразования световой энергии относятся к тепловым фотоприемникам.

В квантовой электронике и оптоэлектронике в основном используются квантовые фотоприемники, которые подразделяются на фотодиоды с р—п-переходом, фотодиоды со сложной структурой, фоторезисторы, фотоэмиссионные приемники — фотоэлектронные умножители (ФЭУ) и электронно-оптические преобразователи (ЭОП). К фотонным приемникам следует отнести также ПЗС-приемники — матричные или линейные преобразователи на приборах с зарядовой связью.

6.1. Фотодиоды

Фотогальванический эффект возникает при поглощении света полупроводником при одновременной генерации подвижных носителей — электронов и дырок. Эти носители разделяются в пространстве. Причиной разделения носителей может быть электрическое поле приэлектродного барьера Шоттки на контакте "металл—полупроводник", поле р—п-перехода или гетеропереход.

На рис. 6.1 представлена схема р—п-перехода, на который падает излучение с энергией фотонов $h\nu / E_g$, где E_g — ширина запрещенной зоны. При освещении р—п-перехода поглощение фотона может произойти в области объемного его заряда. Под воздействием внутреннего поля перехода носители заряда будут перемещаться в противоположных направлениях: электроны будут стремиться в п-область, а дырки соответственно в р-область. Вероятность этого процесса мала, потому что толщина слоя объемного заряда невелика.

Основная часть носителей генерируется в областях, примыкающих к р—п-переходу (на рис. 6.1 они обозначены как области 2 и 3).

Если генерация происходит на расстояниях меньше диффузионной длины электронов L_n и дырок L_p , то сгенерированные носители дойдут до р—п-перехода, не рекомбинируя с ос-

новными носителями этих областей. В области объемного заряда неосновные носители сортируются полем р—п-перехода: избыточные носители накапливаются в разных областях, что приводит к возникновению фото-ЭДС и фототока при наличии внешней цепи.

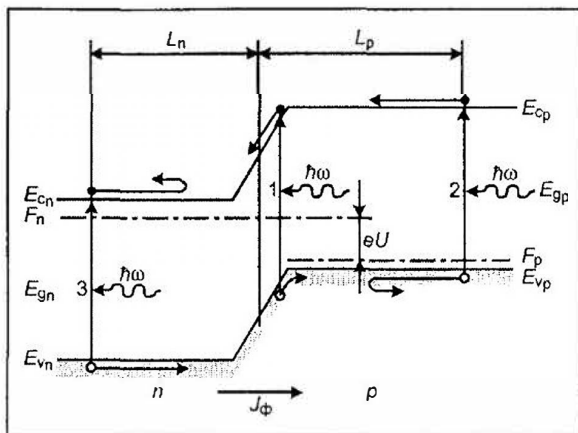


Рис. 6.1. Схема процессов, происходящих в р—п-переходе под воздействием света

В результате разделения носителей р-область будет заряжаться положительно, а п-область — отрицательно.

Фототок через р—п-переход протекает независимо от приложенного напряжения и имеет вид:

$$J = J_s \exp\left(\frac{qU}{kT} - 1\right) - J_\phi, \quad (6.1)$$

где J_s — ток насыщения, создаваемый свободными носителями заряда; J_ϕ — фототок, пропорциональный скорости генерации избыточных электронно-дырочных пар в области диффузионных длин неосновных носителей; $qU = E_{Fn} - E_{Fp}$.

Вольтамперная характеристика фотодиода представлена на рис. 6.2.

Значение фото-ЭДС можно определить, положив $J = 0$. Тогда:

$$E_{\phi} = \frac{kT}{q} \ln \left(1 + \frac{J_\phi}{J_s} \right).$$

Такой процесс разделения зарядов и формирования ЭДС называется всильным или барьерным ЭДС.

Если к р—п-переходу приложить обратное напряжение, по величине равное напряжению лавинного пробоя, то возможно усиление фото-ЭДС за счет лавинного умножения генерированных светом носителей.

Лавинные фотодиоды перспективны при обнаружении слабых оптических сигналов.

Существует и объемная ЭДС, вызываемая разделением пар носителей неоднородностями в объеме образца. Это может быть изменение концентрации легирующих примесей и изменение химического состава сложных полупроводников. Причиной разделения пар яв-

ляется встроенное электрическое поле, возникающее в результате изменения положения уровня Ферми. Такие полупроводники называются варизонными. Фотогальванический эффект применяется в фотодиодах, фототранзисторах, фототиристорах, т. е. во всех фотоприемниках с $p-n$ -переходом.

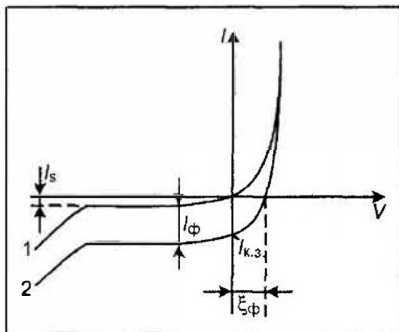


Рис. 6.2. Вольтамперная характеристика фотодиода:
 J_s — ток насыщения; J_ϕ — фототок;
 $J_{kз}$ — ток короткого замыкания;
 E_ϕ — фото-ЭДС;
 1 — освещение отсутствует;
 2 — освещение присутствует

Большое распространение получили $p-i-n$ -диоды, в которых i -область состоит из слабо легированного полупроводника. Электрическое поле в i -области ускоряет транспорт носителей и снижает барьерную емкость фотодиода. Последнее обстоятельство позволяет снизить величину емкости на входе предусилителя и уменьшить время переходных процессов.

Диоды на $p-i-n$ -структуре позволяют обеспечить высокую чувствительность в длинноволновой области спектра при увеличении i -области, в которой поглощается порядка 90% излучения. Малые рабочие напряжения в фотодиодном режиме позволяют обеспечить совместимость $p-i-n$ -диодов с интегральными схемами.

Фотодиоды Шоттки со структурой "металл—полупроводник" позволяют повысить быстродействие приемников излучения до 10^{-10} с.

В таких структурах граница спектральной характеристики сдвигается в сторону более длинных волн. Фотодиоды Шоттки обладают малым сопротивлением базы фотодиода, и инерционность таких приборов определяется временем пролета фотоносителей через область объемного заряда (примерно 10^{-10} — 10^{-11} с).

Простота создания выпрямляющих fotocувствительных структур с барьером Шоттки на различного типа полупроводниках (даже на тех, на которых нельзя сформировать $p-n$ -переход) открывает большие перспективы использования фотодиодов Шоттки.

Фотодиоды на гетероструктурах позволяют создать фотоприемные устройства с КПД, близким к 100%. На рис. 6.3 показано устройство и зонная диаграмма гетерофотодиода.

Излучение вводится через слой GaAlAs. Поглощение происходит в n -области GaAs. Разница в ширине запрещенной зоны по обе стороны от гетероперехода составляет приблизительно 0,4 эВ. Генерируемые в n -области дырки переносятся в p^+ -область. Ширина активной n -области выбирается такой, чтобы происходило полное поглощение излучения. Структура работает при небольших напряжениях. Выбирая соответствующие пары полупроводников можно создать фотодиоды для любой части оптического спектра. Дело в том, что в гетероструктурах длина волны определяется разницей ширины запрещенных зон и не связана со спектральной характеристикой поглощения излучения.

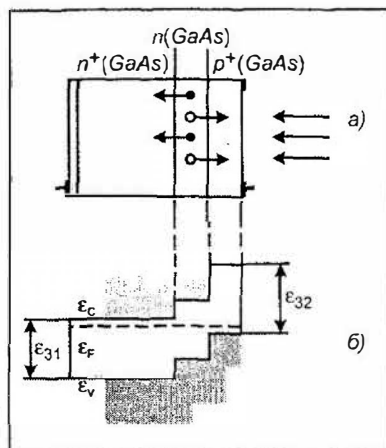


Рис. 6.3. Схема фотодиода с гетероструктурой (а) и его зонная диаграмма (б)

6.2. Фотоприемники с внутренним усилением

Фотоприемники, в которых происходит преобразование оптического излучения с одновременным усилением фототока, называют *фотоприемниками с внутренним усилением*.

Обычно используются *фототранзисторы* биполярного типа, которые включаются в электрическую цепь по схеме с общим эмиттером. База не имеет внешнего вывода (рис. 6.4, а). Управление коллекторным током осуществляется на основе внутреннего фотоэффекта.

Генерация носителей происходит в области базы под действием излучения. Коллекторный переход служит для разделения носителей. Через эмиттерный переход происходит инжекция дырок для восстановления нейтральности базовой области, в которой остаются электроны.

По сравнению с обычным диодом фототранзистор усиливает ток в β раз, где β — коэффициент передачи тока базы.

При использовании фототранзисторов для них выбирают оптимальную конструкцию: уменьшение толщины базы способствует повышению коэффициента передачи и уменьшению времени переключения, но снижает фоточувствительность прибора. Поэтому оптимальное быстродействие транзистора лежит в пределах 10^{-7} — 10^{-8} с.

Составной транзистор представляет собой известную конструкцию из двух транзисторов с общим коллектором. Коэффициент передачи тока базы составляет

$$\beta \approx \beta_1 \cdot \beta_2,$$

где β_1 и β_2 — соответственно коэффициенты передачи тока каждого транзистора. В результате чувствительность составного транзистора повышается и достигает значений $\sim 10^3$, что выше, чем у обычных фотодиодов.

Фототиристор представляет собой четырехслойную структуру типа $n-p-n-p$, которая освещается в плоскости расположения $p-n$ -переходов. К структуре приложено прямое напряжение. Тиристор можно представить как комбинацию двух транзисторов, меж-

ду которыми имеется положительная обратная связь по току. Переход фототиристора из закрытого состояния с низкой проводимостью в открытое состояние с высокой проводимостью происходит под действием света. Тиристоры могут выдерживать токи от 10^{-3} до 10^2 А и напряжения до 10^3 В при мощности светового излучения от 10^{-3} до 1 Вт.

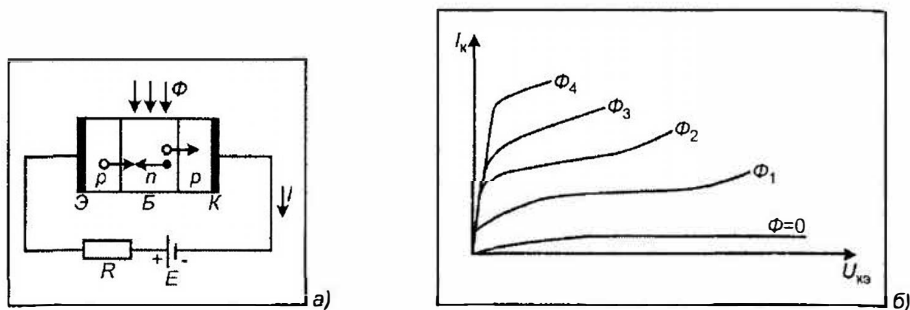


Рис. 6.4. Схема включения фоторезистора (а) и его выходные характеристики (б)

В *фоторезисторах* используется эффект фотопроводимости или фоторезистивный эффект. Он заключается в изменении электропроводности среды, обусловленной действием электромагнитного излучения. В зависимости от механизма поглощения носителей различают примесную, собственную и внутризонную фотопроводимость.

Собственная и примесная фотопроводимость основана на внутреннем фотоэффекте, приводящем к генерации электронно-дырочных пар. Возможен также механизм фотообрыва носителей заряда от заряженного примесного центра. Генерируемые свободные носители получили название *фотоносителей*.

Собственная фотопроводимость сводится к генерации электронов из внешней зоны и генерации дырок из зоны проводимости, если величина кванта энергии не меньше ширины запрещенной зоны.

Примесная фотопроводимость происходит вследствие переходов, обусловленных примесями в полупроводнике, при этом энергия квантов облучения может быть меньше ширины запрещенной зоны.

Внутризонная фотопроводимость связана с изменением подвижности носителей заряда при их перераспределении по энергетическим состояниям, обусловленная поглощением излучения.

На основе такого типа полупроводниковых материалов и изготавливают фоторезисторы. Будучи подключенным к калиброванной электрической цепи, фоторезистор позволяет измерять световые потоки.

В зависимости от назначения фоторезисторы могут быть одно- и многоэлементные, с охлаждением или без охлаждения, открытые и герметизированные. Возможно встраивание фоторезисторов в интегральные схемы для предварительной обработки сигналов.

На рис. 6.5 приведены графики спектральной чувствительности фоторезисторов на основе различных материалов.

В последние годы разрабатываются также различные фотоприемные устройства. Они представляют собой гибридную конструкцию фоточувствительного элемента с трактом обработки сигнала, которые изготавливаются по микронной технологии. Фотоприемные устройства широко распространены в волоконно-оптических системах связи.

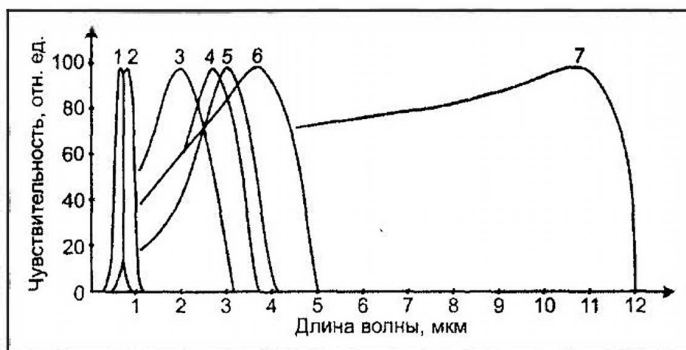


Рис. 6.5. Кривые спектральной чувствительности фоторезисторов на основе CdS (1), CdSe (2), PbS (3), твердого раствора типа PbS-PbSe (4, 5), PbSe (6), PbSn (7)

6.3. Гетеродинный прием оптического излучения

Рассмотренные типы приемников излучения способны реагировать на излучение любой степени когерентности. Это объясняется физическими процессами, лежащими в основе приема излучения: поглощение фотонов и генерация свободных носителей одинаковы как для когерентного, так и для некогерентного излучения. Этот метод регистрации оптического излучения получил название прямого детектирования излучения. Детектор реагирует не на напряженность электрического поля волны, а на поток фотонов или мощность волны. Такие детекторы получили название квадратичных. Для регистрации когерентного излучения был разработан метод *гетеродинирования*, предложенный в 1947 году Г. С. Гореликом. Этот метод был позаимствован из радиотехники.

Идея метода заключается в смешении двух гармонических сигналов, различающихся по частоте или по фазе, и дальнейшем прямом детектировании с помощью обычного квадратичного детектора. При смешении возникает разностная частота, которая и анализируется, в результате чего выделяется мощность колебаний с разностной частотой, а также мощность, эквивалентная шуму (МЭШ).

Мощность шума в полосе с частотой Δf при спектральной плотности шума $S_{ш}$ определяется выражением

$$\overline{I_{ш}^2} = S_{ш} \Delta f, \quad (6.2)$$

где $I_{ш}$ — шумовой ток.

Спектральная плотность мощности шума определяется формулой Шоттки для дробового шума:

$$S_{ш} = 2q\bar{I}.$$

где \bar{I} — среднее значение тока.

Тогда мощность шума в полосе частот Δf определяется как

$$\overline{I_{ш}^2} = 2q\bar{I}\Delta f.$$

В этом выражении среднее значение тока представляет собой сумму тока сигнала I_c мощности P_c и тока фона $I_{фот}$ мощностью

$$P_{фот} \bar{I} = I_c + I_{фот}. \quad (6.3)$$

Пусть сигнал представляет собой когерентное колебание вида $E_c = E_{c0} \cos(\omega_c t)$, а генерируемый генератором колебаний сигнал $E_r = E_{r0} \cos(\omega_r t)$.

При этом мощность генератора сигналов в соответствии с известной теоремой Пойнтинга будет:

$$P_c = \frac{c}{8\pi} E_{c0}^2,$$

а мощность гетеродина соответственно:

$$P_r = \frac{c}{8\pi} E_{r0}^2,$$

где c — скорость света.

Фототоки, генерируемые от сигнала и гетеродина, определим как:

$$I_c = q\eta A \frac{P_c}{h\nu}, \quad (6.4)$$

$$I_r = q\eta A \frac{P_r}{h\nu}, \quad (6.5)$$

где q — заряд электрона, η — потери на отражение и неполное поглощение, A — площадь фоточувствительного элемента, $P/h\nu$ — число падающих фотонов.

В результате суперпозиции двух колебаний E_c и E_r имеем величину полного тока в виде:

$$I(t) = \frac{qnA}{h\nu} \frac{c}{4\pi} (E_c + E_r)^2. \quad (6.6)$$

Проведем несложные арифметические операции и выделим в первую скобку постоянные составляющие фототоков сигнала и гетеродина, а во вторую скобку — токи с суммарными оптическими частотами:

$$I(t) = [I_c + I_r] + 2\sqrt{I_c I_r} \cos(\omega_c - \omega_r)t + [I_c \cos 2\omega_c t + I_r \cos 2\omega_r t + 2\sqrt{I_c I_r} \cos(\omega_c + \omega_r)t]. \quad (6.7)$$

Анализ этого выражения показывает, что на вторую скобку фотоприемники не реагируют, поскольку сумма оптических частот лежит вне предела чувствительности любых фотоприемников. В итоге квадратичный фотоприемник на основе гетеродина будет реагировать на величину:

$$I_{мл} = (I_c + I_r) + 2\sqrt{I_c I_r} \cos(\omega_c - \omega_r)t. \quad (6.8)$$

Значение $[I_c + I_r]$ является константой для данного фотоприемного устройства, а значение фототока с разностной частотой определяется как:

$$I_{\mu\omega}(t) = 2\sqrt{I_c I_r} \cos(\omega_c - \omega_r)t.$$

Средняя мощность этого тока пропорциональна $\overline{I_{\mu\omega}^2}$ или

$$I_{\mu\omega}^2 = 2I_c I_r. \quad (6.9)$$

Анализ этой формулы позволяет заключить, что входная мощность оптического излучения линейно связана с мощностью сигнала разностной частоты.

Мощность сигнала разностной частоты может точно воспроизводить распределение интенсивности спектральной линии сигнала.

Если значение $I_s \gg I_c$, то можно ожидать большого усиления, однако из-за сильной засветки фотоприемника излучением гетеродина возникает сильный радиационный шум, определяемый формулой Шоттки:

$$\overline{I_u^2} = 2qI_s \Delta f.$$

Если мощность фототока с разностной частотой приравнять к мощности радиационного шума гетеродина, то

$$\overline{I_u^2} = 2qI_s \Delta f = \overline{I_{pu}^2} = 2I_c I_s. \quad (6.10)$$

С учетом, что $I_c = q\eta \frac{P_c}{h\nu}$ получаем:

$$P_c = \frac{h\nu}{\eta} \Delta f. \quad (6.11)$$

Из этого следует, что гетеродинный приемник с квантовой эффективностью равной единице способен зарегистрировать один фотон в частной полосе 1 Гц. Гетеродинный прием является весьма эффективным и чувствительным методом при условии согласования волновых фронтов сигнала и гетеродина.

Метод гетеродинирования позволяет выделить слабые оптические сигналы при наличии внутренних тепловых шумов приемника излучения.

Гетеродинный метод широко используется в лазерной интерферометрии, голографии, путем смешения пучков нескольких лазеров или деления пучка лазерного излучения на два — опорный и несущий. В этом случае частоты одинаковы, но при суперпозиции колебаний используется разность фаз.

При использовании пучка одного лазера гетеродинный прием называется *гомодинным*.

Контрольные вопросы

1. Что такое фотодиод?
2. Какие физические явления разделения носителей в фотодиоде вы знаете?
3. Опишите принцип действия фотодиодов на явлении разделения носителей электрическим полем и их характеристики.
4. Опишите принцип действия лавинных фотодиодов и их характеристики.
5. Что такое р—i—n-диод и каковы его характеристики.
6. Опишите принцип действия фотодиодов Шоттки и их характеристики.
7. Опишите принцип действия фотодиодов на гетероструктурах и их характеристики.
8. Опишите принцип действия приемников излучения на фототранзисторах и их характеристики.
9. Опишите принцип действия приемников излучения на фоторезисторах и их характеристики.
10. Опишите принцип гетеродинного приема оптического излучения и его особенности.

Рекомендуемая литература

1. Интегральная оптика. Под ред. Т. Тамира. — М.: Мир, 1978.
2. Курбатов Л. Н. Оптоэлектроника видимого и инфракрасного диапазона спектра. — М.: Изд. МФТИ, 1999.
3. Носов Ю. Р. Оптоэлектроника. — М.: Советское радио, 1977.
4. Лихтин А. Н. Физические основы квантовой электроники и оптоэлектроники. Учебное пособие. — М.: Высшая школа, 1983.
5. Свечников Г. С. Интегральная оптика. — Киев, Наукова думка, 1988.
6. Справочник по лазерам. Под. ред. акад. А. М. Прохорова. В 2-х томах, т. 2. — М.: Советское радио, 1978.

7. Введение в интегральную оптику

Интегральная оптика является разделом оптоэлектроники, в котором изучаются оптические явления в тонких слоях материалов, а также разрабатываются методы создания интегрально-оптических элементов и устройств для целей генерации, преобразования и передачи информации. В рамках интегральной оптики обеспечивается возможность интеграции (объединения) оптических и оптоэлектронных элементов.

7.1. Элементы интегральной оптики

С позиции системного подхода к анализу изделий интегральной оптики можно различить следующие основные элементы.

Первым элементом является источник света, который может излучать когерентную или некогерентную световую волну.

Вторым элементом является миниатюрный волновод — интегрально-оптический световод. Такой световод создается либо на основе световедущего слоя на поверхности или внутри диэлектрической подложки, либо в виде отдельного световолокна. В световедущем слое предусмотрен больший коэффициент преломления, чем в окружающей среде. Это и обеспечивает возможность локализации в нем излучения вследствие явления полного внутреннего отражения.

Третьим элементом интегрально-оптических устройств является устройство управления световым потоком. В его основе лежат либо *пассивные методы*, связанные с возбуждением определенных волноводных мод в зависимости от соотношения показателя преломления световедущего канала, подложки и окружающей среды, величины угла падения света, геометрических размеров волноводов. Сюда включаются планарные линзы и призмы. Возможны и активные способы управления световым потоком за счет использования физических явлений и эффектов: акустооптических, электрооптических, магнитооптических. Они позволяют локально менять показатель преломления материала волновода. Возможно введение в световой канал модуляторов, дефлекторов, частотных фильтров, фазовращателей, направленных ответвителей и других элементов управления световыми потоками. Это так называемые *активные интегрально-оптические элементы*. Они используются для управления параметрами световой моды, а именно модуляцией амплитуды, фазы, поляризации. Возможно пространственное разделение света в дефлекторах. Толщина волновода достаточно мала и поэтому возможна высокая концентрация световой энергии ($\geq 10^6$ Вт/см²). В этом случае возникают нелинейные оптические явления, такие как фазовый синхронизм взаимодействующих мод за счет волноводной дисперсии, эффект удвоения частоты излучения, генерация гармоник.

Четвертым, естественным элементом интегральной оптики, является носитель информационного сигнала — световая волна. Она может иметь форму, необходимую для эффективной обработки информации: цуг монохроматической волны, плоская волна и т. п.

Пятым элементом является приемник оптического излучения. С его помощью детектируется световой сигнал и снимается информация.

7.2. Устройства и элементная база интегральной оптики

Элементной базой интегральной оптики служат интегрально-оптические элементы, представляющие собой миниатюризированные оптические и оптоэлектронные устройства, предназначенные для передачи и обработки световых сигналов.

Различают три группы интегрально-оптических элементов. Пассивные элементы управления излучением, элементы преобразования электрической энергии в световую и обратно. К пассивным интегрально-оптическим элементам относятся, прежде всего, устройства ввода и вывода излучения. Они предназначены для согласования световых потоков входящих и выходящих из волновода. На рис. 7.1 представлены схемы интегрально-оптических элементов связи.

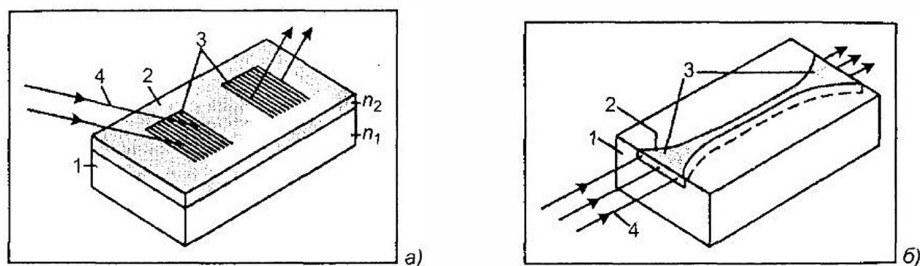


Рис. 7.1. Схема интегрально-оптического элемента связи на основе дифракционных решеток (а) и с использованием рупорных переходов (б): 1 — подложка; 2 — волновод; 3 — элемент связи; 4 — световой поток

Возможны и другие формы интегрально-оптических элементов ввода-вывода, например, волноводы с изменяемой геометрией. К этой же группе относятся интегрально-оптические разветвители излучения с разным числом каналов. Интегрально-оптические линзы фокусируют излучение волновода в заданном месте. Различают геодезические линзы, линзы Льюнеберга и линзы Френеля (рис. 7.2).

Геодезическая линза конструктивно выполняется в виде углубления на поверхности волновода, такого, что происходит фокусировка излучения.

Линза Льюнеберга представляет собой область определенной конфигурации с отличным от волновода показателем преломления:

$$n_2 > n_1,$$

где n_1 — показатель преломления волновода, n_2 — показатель преломления воздуха.

Линза Френеля является фрагментом известной в оптике зонной пластинки Френеля и представляет собой дифракционную решетку с переменным шагом. Линза Френеля может быть голографическим элементом связи.

К этой группе интегрально-оптических элементов можно отнести интегрально-оптические фильтры, представляющие собой дифракционные решетки, а также кольцевые интерферометры и резонаторы (например, типа Фабри — Перо). Эти интегрально-опти-

ски элементы способны изменять пространственные характеристики световых сигналов, оставляя неизменным их энергию.

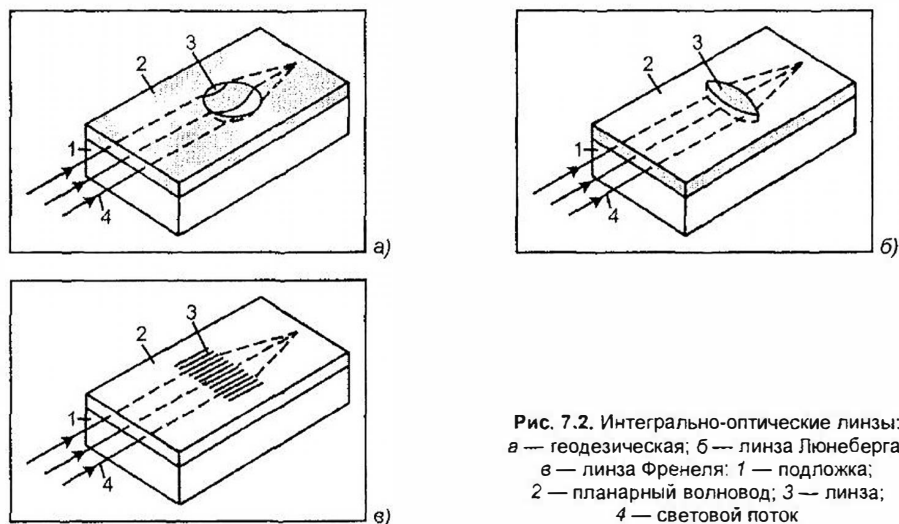


Рис. 7.2. Интегрально-оптические линзы:
а — геодезическая; б — линза Луненберга;
в — линза Френеля; 1 — подложка;
2 — планарный волновод; 3 — линза;
4 — световой поток

Вторая группа интегрально-оптических элементов позволяет управлять излучением путем изменения амплитуды, фазы или поляризации. В этих элементах используются явления, связанные с изменением показателя преломления световода за счет электро-, магнито- или акустооптических эффектов.

Так интегрально-оптические модуляторы амплитуды светового потока строятся по схеме интерферометра Маха — Цандера путем разветвления волновода на основе электрооптических материалов (рис. 7.3, а). В каналах при подаче управляющих сигналов изменяются фазы световых волн, и при их новой интеграции на основе явления интерференции изменяется амплитуда световой волны.

В интегрально-оптических переключателях осуществляется управляемое перераспределение оптической энергии между волноводами (рис. 7.3, б). Это происходит благодаря изменению показателя преломления в области связи между волноводами. Эти изменения происходят под действием управляющего напряжения.

Работа интегральных акустооптических модуляторов основана на изменении направления распространения световых пучков в планарном волновод в результате явления дифракции света на фазовых неоднородностях. Фазовые неоднородности возникают в узлах и пучностях поверхностной акустической волны, возбуждаемой встречно-штыревыми преобразователями (рис. 7.3, в).

Разработаны интегрально-оптические преобразователи частот, действие которых основано на изменении характеристик нелинейно-оптических материалов волноводов с последующим смешением световых частот.

Третья группа интегрально-оптических элементов предназначена для генерации, усиления и детектирования оптических сигналов.

Генерация оптического излучения производится в интегрально-оптическом волновод в результате рекомбинации электронно-дырочных пар в области р—п-перехода полупроводникового излучателя (фотодиода, лазера).

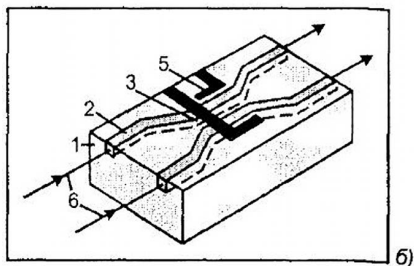
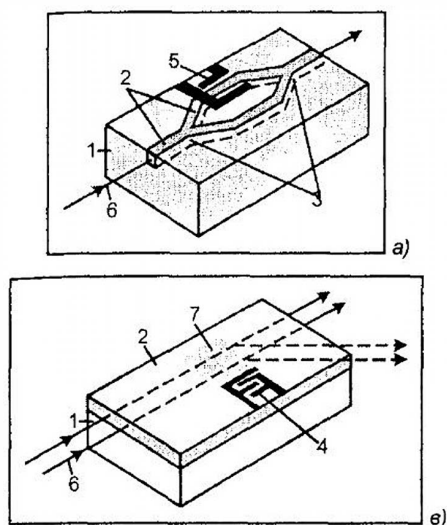


Рис. 7.3. Интегрально-оптические элементы для модуляции по амплитуде (а); переключения световых потоков (б); дефлектор (в):
1 — подложка; 2 — каналные и планарный волноводы; 3 — области связи пучков; 4 — встречно-штыревой преобразователь; 5 — электроды управления; 6 — световой пучок; 7 — поверхностно акустическая волна

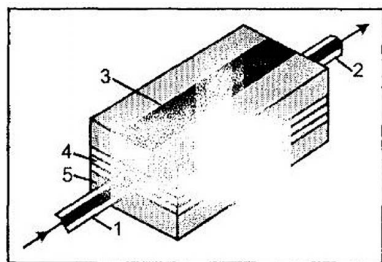


Рис. 7.4. Структура интегрально-оптического усилителя:
1 — входной волновод; 2 — выходной волновод с усиленным излучением; 3 — полосковый электрод; 4 — активная инверсная среда; 5 — полупроводниковая подложка

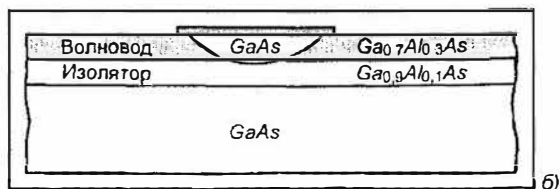
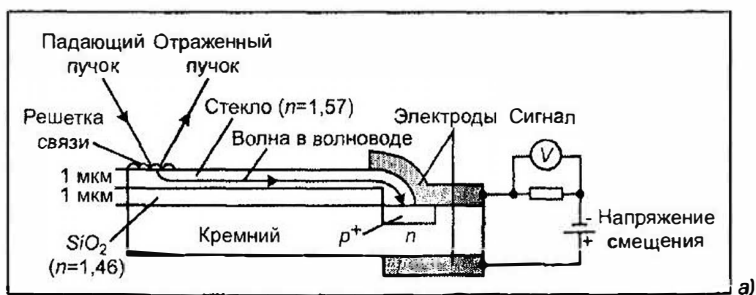


Рис. 7.5. Схема интегрально-оптического детектора на основе кремния (а) и на основе AlGaAs/GaAs (б)

В качестве оптического усилителя могут использоваться слоистые структуры с активной средой. В этой среде создаются условия инверсной населенности энергетических уровней с последующим усилением проходящего излучения (рис. 7.4).

Детектирование излучения осуществляется с помощью фотоприемника, в качестве которого используются фоторезисторы, фотодиоды, фототранзисторы. Фотоприемники обычно сопрягаются с оптическим волноводом.

Широкое распространение получили кремниевые диоды с $p-n$ -переходом (рис. 7.5, а). В качестве подложки используется кремниевая пластина с нанесенным стеклянным оптическим волноводом. Свет из волновода попадает на поверхность $p-n$ -перехода. Детектор на основе фотодиода реализуется с помощью процессов эпитаксии и ионной имплантации. Эти детекторы используются в области видимого оптического излучения. В ИК-области спектра могут быть использованы интегрально-оптические $AlGaAs/GaAs$ детекторы (рис. 7.5, б).

7.3. Интегрально-оптические схемы

Интегрально-оптическая схема представляет собой миниатюризированное функционально законченное устройство обработки информации, связь между элементами которой осуществляется с помощью электрических и световых сигналов.

Различают гибридные и монокристаллические интегрально-оптические схемы.

Гибридные интегрально-оптические схемы создают методом прецизионной сборки из отдельных интегрально-оптических элементов на общей подложке.

В оптических интегральных схемах различают электронную и фотонную части. С помощью многократных ростовых операций и селективного травления на одной пластине размещены электронные устройства и фотонные устройства. Пространственное разделение фотонной и электронной частей может производиться как путем вертикальной, так и горизонтальной интеграции (рис. 7.6).

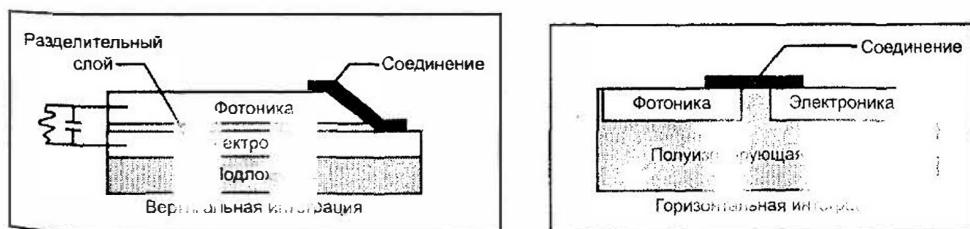


Рис. 7.6. Схема вертикальной и горизонтальной интеграции электроники и фотоники

В вертикально-интегрированной структуре различные слои материалов разделены полупроводящими материалами такой толщины, чтобы избежать паразитных связей между различными участками.

К недостаткам вертикально-интегрированных структур можно отнести непланарное расположение электрических соединений между фотонными и электронными частями схемы. Это затрудняет изготовление сложных схем. Горизонтальная геометрия позволяет разнести в пространстве электронную и фотонную части и тем самым уменьшить паразитные связи.

При проектировании оптических интегральных схем необходимо придерживаться трех принципов:

1. Сохранение объема сигнала. Это означает, что оптоэлектронное преобразование информации должно происходить без потерь информации. Другими словами, объем сигнала определяется соотношением $V = T \cdot F \cdot L = \text{const}$, где T — длительность сигнала, F — ширина спектра частот, $L = \ln(P_c / P_w)$ — динамический диапазон, P_c — мощность сигнала, P_w — мощность шума. Этот принцип определяет возможность изменения параметров сигнала в процессе оптоэлектронного преобразования.
2. Оптимальность структуры. Этот принцип выражает необратимость фотон-электронного и электрон-фотонного преобразования сигнала в одной системе. При оптоэлектронном преобразовании потери информации должны быть минимальными. С этой целью фотоприемные звенья следует формировать в полупроводнике с меньшей шириной запрещенной зоны, чем излучающие звенья.
3. Информативность преобразования. Принцип заключается в том, что при равных информационных объемах сигналов в результате оптоэлектронного преобразования информативность ансамбля фотонов больше информативности ансамбля электронов. Это следует из физических свойств электронов и фотонов. Электрон обладает зарядом и спином и вектор скорости электрона при его движении определяется векторами внешнего электрического и магнитного полей. Фотон движется с постоянной скоростью и независимо от внешних электромагнитных полей.

На рис. 7.7 приведена интегрально-оптическая схема, выполняющая функцию спектроанализатора.

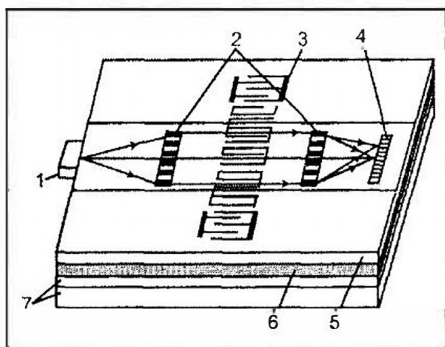


Рис. 7.7. Интегрально-оптическая схема спектроанализатора: 1 — генератор света (лазер); 2 — градиентные линзы Френеля; 3 — встречно-штыревой преобразователь поверхностной акустической волны; 4 — фотодетектор; 5 — волновод; 6 — слой оксида цинка; 7 — подложка

На кремниевой подложке нанесен слой оксида цинка, использующийся в качестве электроакустического преобразователя.

Дифракционные линзы Френеля обладают высокой дифракционной эффективностью (> 90%) при небольшой мощности потоков (> 200 мВт).

Сложный высокочастотный сигнал возбуждает поверхностную акустическую волну (ПАВ) и создает в передней фокальной плоскости Фурье-линзы соответствующую картину распределения пространственных частот. Эта картина движется со скоростью ПАВ. Коллимированный пучок дифрагирует на определенном распределении пространственных частот. Частотному спектру радиосигнала соответствует Фурье-преобразование, которое фокусируется в плоскости линейки фотодетекторов.

Подложка спектроанализатора изготавливается из кремния. Так как кремний не обладает пьезоэлектрическими свойствами, то наносится акустооптическая пленка. Кремниевая подложка позволяет интегрировать элементы. Например, изготавливать пленочные линзы Люнеберга и другие элементы. Такого типа интегрально-оптические спектроанализаторы работают в полосе частот до 500 МГц с лазерным источником в ближнем ИК-диапазоне.

Монолитные интегрально-оптические схемы изготавливают в основном на базе соединенный A^3B^5 . На рис. 7.8 приведен пример интегрально-оптической схемы, состоящей из лазера с распределенной обратной связью модулятора и детектора. Свет переводится из слоя в слой с помощью сужающегося волновода из $AlGaAs$. В этом волноводном слое должна осуществляться модуляция света. Показатель преломления этого слоя должен быть больше, чем у находящегося под ним слоя.

В волноводной области наряду с модуляцией могут быть выполнены функции переклю- чения света, фильтрация, модовая селекция и т. п.

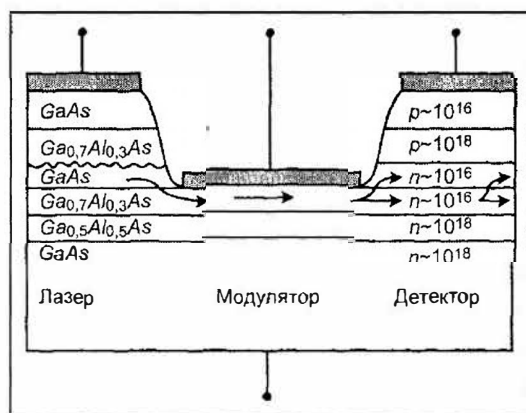


Рис. 7.8. Фрагмент интегрально-оптической схемы (стрелками обозначены световые потоки)

В области детектора промодулированный свет попадает в поглощающую область $p-n$ -перехода на $GaAs$. Полс обратного смещения превращает электронно-дырочные пары в фототок.

Изготовление монолитных интегрально-оптических схем сопряжено с рядом проблем. Среди главных следует выделить технологические: эпитаксиальное наращивание слоев с нужными электрооптическими свойствами, селективное травление, рост через маску и др.

Решив эти и другие проблемы, можно будет надеяться на создание интегрально-оптических схем, превосходящих полупроводниковые ИС по скорости обработки информации, высокой помехоустойчивости. Они найдут широкое применение в системах оптической обработки информации, включая волоконно-оптические линии связи. Особое внимание будет уделено оптическим вычислительным машинам.

Контрольные вопросы

1. Что такое интегральная оптика?
2. Какие элементы интегральных оптических устройств вы знаете?
3. Что является элементной базой интегральной оптики?

4. Что представляют собой интегрально-оптические линзы?
5. Что представляют собой интегрально-оптические модуляторы?
6. Как осуществляется детектирование оптических сигналов в устройствах интегральной оптики?
7. Что такое интегрально-оптическая схема?
8. Назовите области перспективного применения интегрально-оптических схем.

Рекомендуемая литература

1. Голубков В. С., Евтихнев Н. Н., Папуловский В. Ф. Интегральная оптика в информационной технике. — М.: Радио и связь, 1985.
2. Курбатов Л. Н. Оптоэлектроника видимого и инфракрасного диапазона спектра. — М.: Изд. МФТИ, 1999.
3. Свечников Г. С. Интегральная оптика. — Киев, Наукова думка, 1988.
4. Интегральная оптика. Под ред. Т. Тамира. — М.: Мир, 1978.

8. Введение в оптоэлектронику

Оптоэлектроника представляет собой раздел электроники, в котором изучаются эффекты и явления взаимного преобразования электромагнитного излучения оптического диапазона (0,5—20,0 мкм) с электрическими сигналами в твердом теле и методы создания приборов и устройств, использующих эти эффекты для генерации, передачи, хранения и обработки информации.

В оптоэлектронике условно выделяют фотонику и оптронику.

В *фотонике* исследуются методы создания устройств, предназначенных для хранения, передачи, обработки и отображения информации, представленной в виде оптических сигналов.

В рамках *оптроники* развиваются методы создания оптронных схем — электронных устройств с внутренними оптическими связями.

Обработка и хранение информации оптическими методами является центральной задачей оптоэлектроники. Оптическая обработка информации основана на использовании в качестве носителя информации оптического излучения, а в качестве преобразователей информации оптических и оптоэлектронных элементов. Отличительной особенностью оптической обработки информации является естественная пространственная многоканальность, которая обеспечивает одномоментную обработку больших информационных массивов. Оптическая обработка информации позволяет производить аналоговые и цифровые операции над информационными сигналами. В зависимости от характера используемого излучения различают когерентную и некогерентную обработку оптической информации.

8.1. Элементная база

Отличительной особенностью оптоэлектронных приборов является использование электрически нейтральных квантов оптического излучения, с высокой частотой колебаний и соответственно малым значением длины волны оптического излучения, малой расходимостью светового луча и возможностью его острой фокусировки. В этом плане наибольшее предпочтение отдается когерентным излучателям.

8.1.1. Некогерентные излучатели

Некогерентные излучатели представляют собой светодиоды на основе полупроводниковой структуры, излучающей свет при пропускании тока через р—п-переход в прямом направлении.

Различают два механизма возбуждения светодиода: инжекция носителей заряда и ударная ионизация.

В основе спонтанной генерации света лежат процессы инжекции неосновных носителей заряда в активную область р—п-структуры. В результате рекомбинации носителей заряда происходит исчезновение пары свободных противоположно заряженных носителей с од-

новременным выделением избыточной энергии порядка ширины запрещенной зоны в виде кванта света. Может происходить и процесс безызлучательной рекомбинации, при которой происходит рождение фононов.

Процесс инжекции наиболее эффективен в гетероструктурах, где имеются разрывы в валентной зоне и зоне проводимости, и при смещении гетероперехода в прямом направлении наблюдается эффективная инжекция носителей заряда. При этом инжекция происходит из широкозонного материала в узкозонный. Вывод излучения осуществляется перпендикулярно плоскости гетероструктуры через верхний широкозонный слой. Максимальный квантовый выход или отношение числа вышедших фотонов к числу рожденных составляет $\sim 40\%$. Быстродействие достигает $\sim 10^{-8}$ — 10^{-9} с.

Ударная ионизация возникает при обратном смещении р—п-перехода до напряжения электрического пробоя. Этот механизм менее эффективен, чем инжекционный.

Излучательная рекомбинация осуществляется в прямозонных полупроводниках GaAs, InAs, твердых растворах $\text{GaAs}_{1-x}\text{P}_x$, $\text{Ga}_{1-x}\text{Al}_x\text{As}$. Условием инжекции является прямое смещение и соблюдение условия равенства квазиимпульса минимума зоны проводимости и максимума валентной зоны. Этот переход с сохранением квазиимпульса является излучательным (см. рис. 3.3).

Светодиоды на основе гомопереходов в прямозонных полупроводниках имеют сильное поглощение внутри кристалла. В связи с низким квантовым выходом светодиоды этого типа применяются редко.

В непрямозонных полупроводниках излучательная рекомбинация происходит при наличии определенных примесных центров, на которых локализуются электроны. После захвата электрона к центру кулоновскими силами подтягивается и дырка. В результате реализуется излучательный переход.

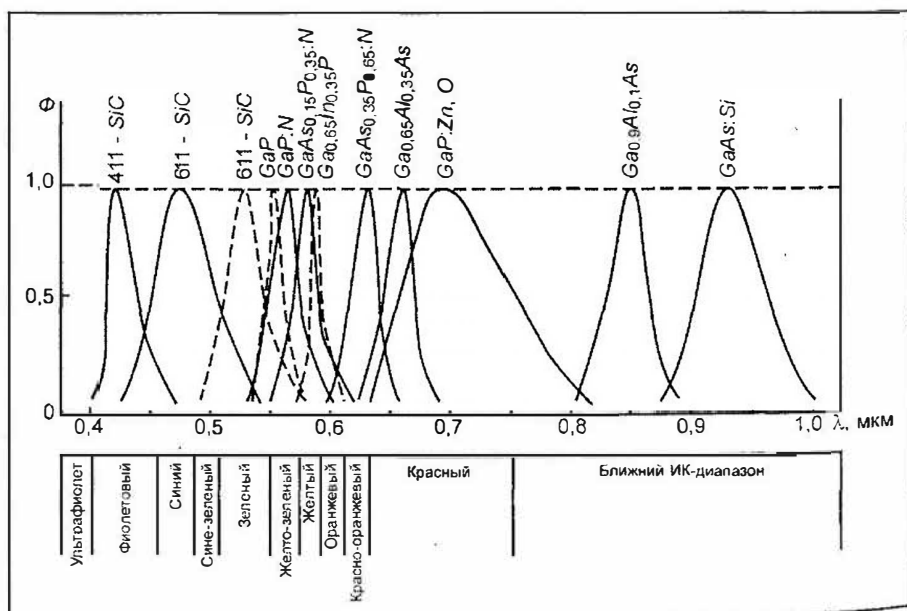


Рис. 8.1. Типичные спектры излучения светодиодов

В целях расширения спектрального диапазона излучения применяют тройные и четверные соединения. Эти соединения характеризуются изовалентным замещением элементов III и V групп периодической системы элементов.

Типичные спектры излучения светодиодов приведены на рис. 8.1.

8.1.2. Когерентные излучатели

Когерентные излучатели получают на основе тех же $p-n$ -переходов и создания условий для вынужденного излучения. Это можно достигнуть, создавая положительную обратную связь путем помещения инверсной активной среды в резонатор.

Рассмотренные ранее полупроводниковые лазеры, которые могут быть изготовлены на основе одного полупроводникового материала (гомолазеры), являются малоэффективными. Гетеролазеры изготавливаются на основе гетеропереходов двух и более полупроводниковых материалов с отличными друг от друга энергиями запрещенных зон.

Гетеролазеры для оптоэлектроники используют, как правило, твердые растворы типа In-GaAsP в различном сочетании. Комбинации четверных и тройных соединений A^3B^5 позволяют получать когерентные излучатели в диапазоне длин волн 0,66—4 мкм.

Лазерные структуры выращиваются по эпитаксиальной технологии. Различают жидкофазную, газофазную и молекулярно-лучевую эпитаксию (см. часть II "Микроэлектроника").

Инжекционные лазеры являются высокоскоростными приборами, у которых максимальная частота модуляции излучения может достигать значения 2×10^{10} Гц.

Для хорошего согласования излучения лазера с каналами передачи информации необходимо введение лазера в одномодовый режим работы. Одномодовые лазеры отличаются низким уровнем шумов и стабильностью.

В большинстве инжекционных лазеров в качестве резонатора Фабри — Перо используются сколотые грани лазерного кристалла, полупрозрачные зеркала. Это упрощает конструкцию, но затрудняет интеграцию источника когерентного излучения с другими элементами.

Для создания положительной обратной связи в лазерах с периодической модуляцией оптических характеристик необходимо менять направление волны, рассеиваемой на неоднородностях, на противоположное. Возникают две связанные волны одной частоты, распространяющихся в противоположных направлениях. Такое возможно, если параметры оптической среды промодулировать по периодическому закону в направлении распространения. Например, по закону, удовлетворяющему условию Брэгга:

$$\Lambda = m\lambda_0 / 2n_{\text{эф}}$$

где Λ — период, m — порядок брэгговского отражения, λ_0 — длина волны излучения, $n_{\text{эф}}$ — эффективный показатель преломления волновой моды.

Такой брэгговский резонатор имеет только одну полосу резонансного усиления. В итоге получается лазер с высокой спектральной избирательностью ($\Delta\lambda_0 \approx 0,1$ нм).

Различают два типа излучателей с периодической модуляцией оптических характеристик.

В лазерах с распределенным брэгговским отражателем (РБО-лазеры) используются модулированные участки оптической среды, вынесенные из активной области (рис. 8.2, а). В лазерах с распределенной обратной связью (РОС-лазеры) модуляция среды осуществляется непосредственно в активной области (рис. 8.2, б).

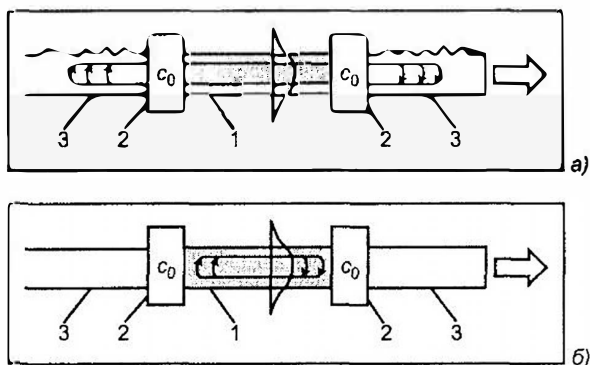


Рис. 8.2. Лазеры с периодической структурой обратной связи: а — РОС-лазеры; б — РБО-лазеры; 1 — активный волновод; 2 — область связи волноводов; 3 — пассивный волновод

Лазерные структуры с периодической модуляцией оптических характеристик различаются порядком дифракции. Этот порядок определяется целым числом полуволн лазерного излучения, которые можно уложить на периоде неоднородности.

Конструктивно неоднородности выполняют в виде диэлектрика с гофрированной поверхностью. Изменение толщины волновода определяет периодичность неоднородности. Коэффициент связи C_0 определяет степень связи активной генераторной и пассивной волноводной компонент.

Высокая эффективность РБО-лазеров достигается за счет малых потерь в областях брэгговского отражения и высокоэффективная связь между внешними и внутренними волноводами.

В РОС-лазерах качество связи в основном определяет эффективность ввода излучения в элементы оптической схемы.

Для увеличения мощности когерентного излучения лазеры объединяют в лазерные решетки.

На рис. 8.3 приведена схема сопряжения в решетку двух гетеролазеров с распределенным брэгговским отражением.

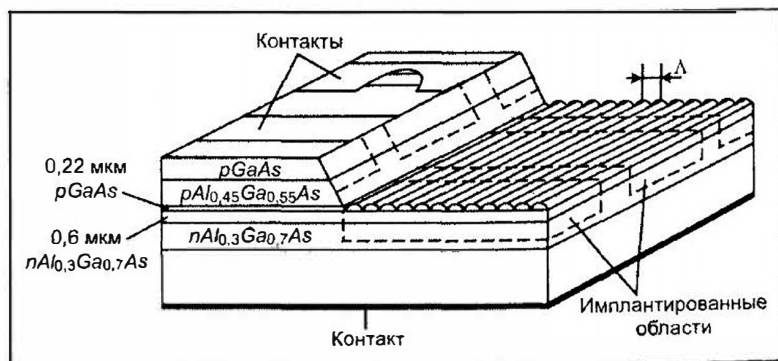


Рис. 8.3. Два полосковых гетеролазера с распределенным брэгговским отражением

Дифракционная решетка в РОС-гетеролазерах используется для вывода излучения. Это позволяет улучшить направленность излучения и повысить мощность.

При использовании в устройствах оптоэлектроники когерентных источников излучения достигается наименьшая расходимость излучения, передается энергия излучения в заданную точку пространства с минимальными потерями, обеспечиваются более широкие функциональные возможности устройств фотоники и оптронных схем.

8.1.3. Приемники излучения

В качестве приемников излучения в устройствах оптоэлектроники используются полупроводниковые фотодиоды, фототранзисторы, фототиристоры, фоторезисторы, которые были рассмотрены ранее.

8.2. Оптоэлектронные устройства обработки информации

В оптоэлектронике разработаны различные устройства и системы для обработки информации. Прежде всего, это различные индикаторные устройства, оптроны, волоконно-оптические системы связи, оптические процессоры. Провести строгую классификацию и отнести то или иное устройство к оптронному или фотонному направлению в оптоэлектронике весьма затруднительно. Однако оптроны традиционно считаются изделиями оптоэлектроники. Эти устройства находят применение в устройствах связи в качестве элементов гальванической развязки.

8.2.1. Оптоны

Оптоэлектронный процессор или **оптрон** представляет собой совокупность оптоэлектронных элементов, выполняющих операции в соответствии с заданной функцией и алгоритмом обработки информации.

С помощью оптронов удается осуществить связь между отдельными частями электронных устройств при условии обеспечения полной гальванической развязки между ними. Помехозащищенность обеспечивается достижением сопротивления изоляции более 10^{14} Ом и емкости связи порядка 10^{-14} Ф.

Оптрон конструктивно состоит из источника излучения, оптического канала передачи информации (световода) и приемника излучения. Управление можно осуществлять путем подачи соответствующего электрического импульса на светодиод, а также изменением геометрии световода.

В зависимости от совокупности характеристик генератора и детектора оптрон может выполнять различные процессы обработки информации: усиление, переключение, согласование, преобразование, индикация и т. п. (рис. 8.4). С помощью оптронов осуществляется гальваническая развязка электрических цепей, оптическая связь с помощью волоконно-оптических линий связи большой информационной емкости. Так, к примеру, на фототранзисторных оптронах легко собрать логические схемы на основе ячеек типа ИЛИ и И (рис. 8.4, б, в). Заметим, однако, что оптроны как элементная база построения вычислительных устройств неконкурентоспособны с транзисторными. Во-первых, двойное преобразование энергии обуславливает низкую энергетическую добротность оптронов. Во-вторых, значительные технологические трудности вызывают необходимость согласования

спектральных характеристик светодиодов и фотоприемников при значительном количестве числа оптронов. В-третьих, оптроны являются элементами оптических схем и поэтому принципиальные ограничения, характерные для схемотехнических методов обработки информационных массивов, накладывают серьезные ограничения и на этот тип схем. На основе оптронов возможно формирование обучаемой системы восприятия и распознавания образов — *персептрона*.

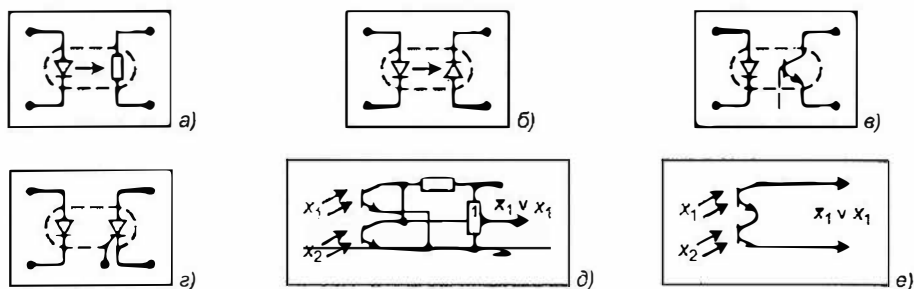


Рис. 8.4. Оптроны различной конструкции: а — резисторный; б — диодный; в — транзисторный; г — тиристорный; д — ячейка типа ИЛИ; е — ячейка типа И

Оптроны обеспечивают высокую электрическую прочность, однонаправленность потока информации, что исключает реакцию приемника на источник, широкую полосу пропускания.

В оптоэлектронике важнейшей проблемой является увеличение функциональной насыщенности элементов, создание приборов и устройств, в которых сочетались бы функции генерирования, детектирования, модуляции, запоминания и избирательности в процессах оптической обработки информации.

Контрольные вопросы

1. Что такое оптоэлектроника?
2. Несогерентные излучатели: конструкции и характеристики.
3. Что такое РОС-лазеры?
4. Что такое РБО-лазеры?
5. Что такое оптрон?
6. Какие типы оптронов вы знаете?
7. Каково функциональное назначение оптронов?

Рекомендуемая литература

1. Интегральная оптоэлектроника. Под ред. Е. Пашенко, М. Кропоткина. — М.: 1990.
2. Курбатов Л. Н. Оптоэлектроника видимого и инфракрасного диапазона спектра. — М.: Изд. МФТИ, 1999.
3. Морозов В. Н. Оптоэлектронные матричные процессоры. Основы оптоэлектроники. Пер. с японского. — М.: Мир, 1988.
4. Носов Ю. Р. Оптоэлектроника. — М.: Радио и связь, 1989.

9. Оптические методы обработки информации

Оптическая обработка информации основывается на использовании оптического излучения как носителя информационного сигнала и оптических элементов для обработки этих сигналов.

9.1. Оптические сигналы

Характерной особенностью оптических сигналов является их двумерность. Это свойство связано с малой длиной световой волны ($\lambda = 0,4 \div 1,5$ мкм).

Оптический световой диапазон неоднороден и подразделяется на поддиапазоны (рис. 9.1).

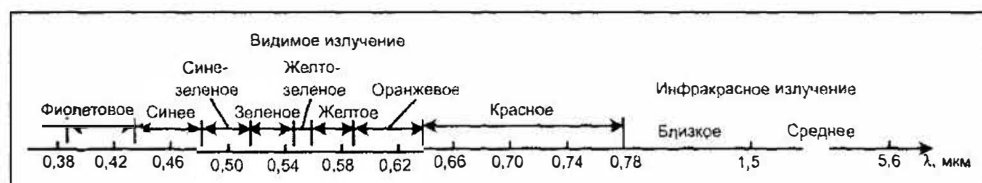


Рис. 9.1. Шкала оптического диапазона. Частота излучения в оптическом диапазоне составляет $\sim 10^{15}$ Гц, энергия излучения лежит в диапазоне 10—1 эВ

Для количественного описания оптического излучения пользуются параметрами, соотношения между которыми приведены в табл. 9.1.

Атомы вещества излучают короткими сериями — *волновыми пакетами*. Длительность волнового пакета и спектр излучения взаимосвязаны: чем больше длительность волнового пакета, тем уже спектр излучения. Это следует из теоремы Фурье, согласно которой любую конечную и интегрируемую функцию можно представить в виде непрерывной суммы бесконечного числа синусоидальных компонент. Колебание $F(t)$, вызванное прохождением пучка волн, можно представить как суперпозицию монохроматических колебаний с различными частотами ν .

Тогда

$$F(t) = \int_{-\infty}^{+\infty} f(\nu) \exp(i2\pi\nu t) d\nu,$$

где $f(\nu)$ определяет распределение по амплитуде монохроматических составляющих $F(t)$. Распределение этих составляющих по энергиям дается выражением $|f(\nu)|^2$. В соответствии со свойствами преобразования Фурье можно произвести обратное преобразование и получить спектр функции $F(t)$. Действительно

$$f(t) = \int_{-\infty}^{+\infty} F(t) \exp(-i2\pi\nu t) dt.$$

Таблица 9.1. Фотометрические параметры излучения

Энергетические параметры			Световые параметры			Формула
Параметр	Обозначение	Единица измерения	Параметр	Обозначение	Единица измерения	
Поток излучения	Φ_e	Вт	Световой поток	Φ_v	лм	$\Phi = dW / dt$, где W — энергия излучения
Сила излучения	I_e	Вт/ср	Сила света	J_v	кд·лм/ср	$I = d\Phi / d\Omega$, где Ω — телесный угол
Энергетическая светимость	M_e	Вт/м ²	Светимость	M_v	лм/м ²	$M = d\Phi / dS$
Энергетическая яркость	L_e	Вт/(ср·м ²)	Яркость	L_v	кд/м ²	$L = \frac{dI}{dS \cos \varphi}$
Энергетическая освещенность (облученность)	E_e	Вт/м ²	Освещенность	E_v	лк·лм/м ²	$E = d\Phi / dS$

На рис. 9.2, а приведены функции $F(t)$, описывающие волновые пакеты и соответствующее распределение энергии (интенсивности I) по частоте (рис. 9.2, б).

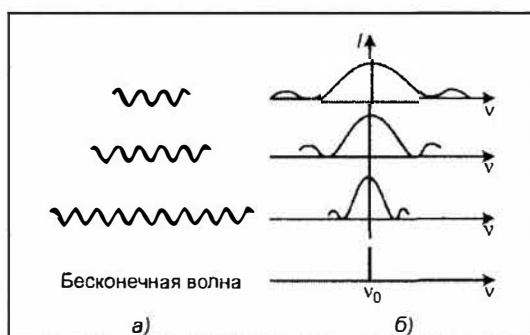


Рис. 9.2. Соотношение между длительностью волнового пакета и его спектром излучения

Особый интерес представляют монохроматические или когерентные сигналы.

Когерентность — явление коррелированного протекания во времени и в пространстве колебательных или волновых процессов, позволяющее получить при их сложении четкую интерференционную картину.

В общем случае световые колебания частично когерентны и количественно когерентность измеряется степенью взаимной когерентности. Эта величина определяется контрастом интерференционной картины. Так интенсивность света в некоторой точке от двух монохроматических источников интенсивностью I_1 и I_2 определяется выражением:

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} |\gamma_{12}(t)| \cos(\omega t + \varphi), \quad (9.1)$$

где $|\gamma_{12}(t)|$ — степень взаимной когерентности, являющаяся функцией расстояния между источниками и времени распространения света, φ — фаза колебаний. Если $I_1 = I_2$, то

$$|\gamma_{12}(t)| = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}},$$

где I_{\max} и I_{\min} — интенсивности светлой и темной полосы интерференционной картины.

Время когерентности τ_0 определяется как минимальная задержка между интерферирующими световыми волнами, снижающая значение $\gamma(t)$ до заданной величины, включая нуль. Величина τ_0 меняется в широких пределах: от $\tau = 10^{-15}$ с для солнечного света до $\tau = 0,1$ с для лазерного излучения высокой степени когерентности.

Длина когерентности $l_0 = c\tau_0$ также меняется в широких пределах.

Введем понятие *оптического сигнала* как электромагнитную волну, в которой в каждой точке пространства электрическое и магнитное поле меняется по гармоническому закону:

$$U(x, y, t) = A(x, y) \cdot \cos[2\pi\nu t + \varphi(x, y)], \quad (9.2)$$

где $U(x, y, t)$ — скалярная функция координат пространства и времени, $A(x, y)$ — амплитуда колебания напряженности электрического поля, ν — частота колебаний, $\varphi(x, y)$ — фаза световой волны.

Рассматривается только электрическая составляющая электромагнитной волны, поскольку именно вектор E ответственен за фиксацию электромагнитной волны квадратичными детекторами (глаз, фотопленка, фотоприемник). Можно записать световой сигнал в комплексной форме, удобной для сложных математических операций.

$$U(x, y, t) = A(x, y) \cdot \exp\{i[2\pi\nu t + \varphi(x, y)]\}. \quad (9.3)$$

Величину $U(x, y, t) = A(x, y) \cdot \exp[i\varphi(x, y)]$ называют комплексной амплитудой, которая основывает пространственное распределение амплитуды $A(x, y)$ и фазы $\varphi(x, y)$ световой волны.

Временной множитель $\exp(2\pi\nu t)$, который является гармонической функцией времени, опускают. Однако он может быть введен на любом этапе математических преобразований.

9.2. Голография

9.2.1. Принципы голографической обработки информации

Голография (от греческого *holos* — весь, полный и *grapho* — пишу, рисую) представляет собой метод записи, воспроизведения и преобразования оптических световых полей.

В основе голографического метода записи лежит регистрация интерференционной картины, образованной волной отраженной от предмета (предметная волна) и когерентной с ней волной от источника света (опорная волна).

При взаимодействии опорной и предметной волн получается новая волна, амплитуда которой является результатом сложения амплитуды опорной волны и амплитуды предметной волны в каждой точке плоскости голограммы; сложение амплитуд происходит в зависимости от фазы каждой волны. Результатом сложения являются точки, в которых наблюдаются интенсивности в пределах от суммы амплитуд до их разности. Пространственное распределение интенсивностей называют *интерференционной картиной*.

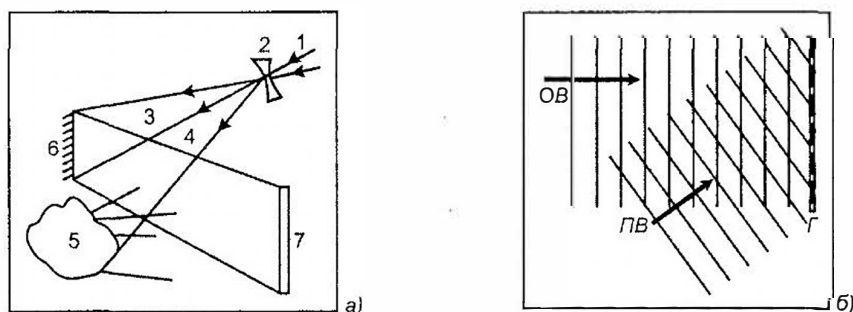


Рис. 9.3. Схема записи голограммы (а) и процесса образования интерференционной картины на голограмме (б): 1 — когерентный пучок света; 2 — линза для расширения пучка; 3 — опорная волна (ОВ); 4 — предметная волна (ПВ); 5 — голографируемый объект; 6 — зеркало; 7 — голограмма

Итак, опорная волна

$$U_{on} = U_{on} \cdot \exp[i\varphi_{on}(x, y)]$$

и предметная волна

$$U_{np} = U_{np} \cdot \exp[i\varphi_{np}(x, y)]$$

взаимодействуя между собой, образуют распределение интенсивности по закону:

$$I = |U_{on} + U_{np}|^2 = (U_{on})^2 + (U_{np})^2 + U_{on}^* U_{np} + U_{np}^* U_{on} \quad (9.4)$$

или

$$I = (U_{on})^2 + (U_{np})^2 + 2U_{on}U_{np}\cos(\varphi_{on} - \varphi_{np}). \quad (9.5)$$

Из этих соотношений следует, что интенсивность излучения в любой точке голограммы от воздействия опорной и предметной волн является суммой интенсивностей отдельных волн и интерференционной составляющей, пропорциональной разности фаз. Под голограммой подразумевается плоскость, в которую помещается фотопластинка для фиксации распределения светового поля. Из выражения (9.5) также следует, что разность фаз модулирует амплитуду интерференционной составляющей. Вот поэтому этот процесс и называется голографией, поскольку фаза опосредованно записывается и является составляющей интенсивности. Другими словами, записывается вся информация о световой волне и фаза, и амплитуда. Записанную интерференционную картину можно интерпретировать как суперпозицию многих интерференционных полос или множество дифракционных решеток.

На стадии восстановления изображения голограмма освещается когерентным светом той же длины волны. При падении лазерного пучка I_0 на дифракционную решетку часть лазерного излучения проходит прямо, одновременно формируются пучки по двум направлениям $+I$ и $-I$, отклоняющихся на угол Θ от основного пучка. Угол Θ зависит от шага (периода) дифракционной решетки (рис. 9.4).

При восстановлении голограммы, представляющей собой суперпозицию множества дифракционных решеток с разным шагом, происходит дифракция света на интерференционных полосах.

Для плоской волны, падающей на дифракционную решетку, справедливо условие Брэгга

$$2d\sin\Theta = \lambda, \quad (9.6)$$

где d — шаг решетки, Θ — направление на главный максимум, λ — длина волны излучения.

Таким образом, при освещении голограммы опорной волной часть дифрагированного на ней света воссоздает волновой фронт, который шел от объекта при записи голограммы. Создается так называемое действительное изображение в том месте, где ранее находился объект. Одновременно восстанавливается сопряженная волна, формирующая мнимое изображение.

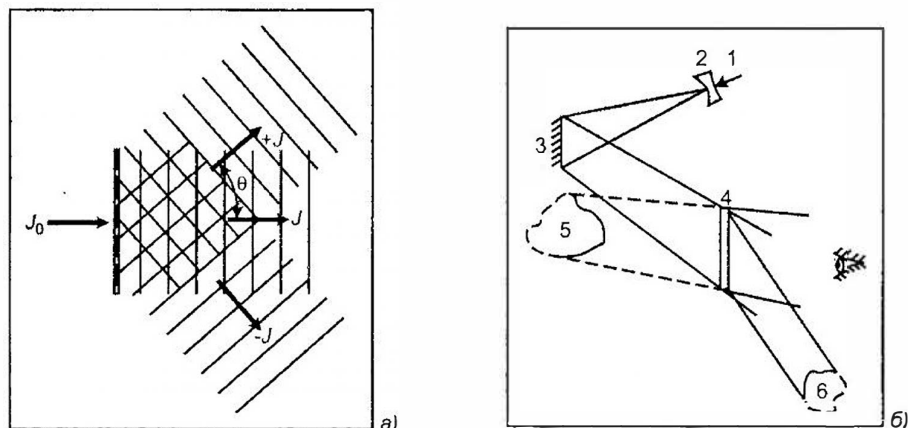


Рис. 9.4. Дифракция плоской волны на дифракционной решетке (а) и схема восстановления изображения с голограммы (б): 1 — лазерное излучение; 2 — линза; 3 — зеркало; 4 — голограмма; 5 — действительное изображение; 6 — мнимое изображение

Пропускание фотопластинки с записью опорной и объектной волн выражается уравнением (9.4) и после проявления фотопластинки будет пропорционально I . Если эту проявленную пластинку-голограмму осветить опорной волной, то пройдет волна, описываемая уравнением:

$$U_{np}I = (|U_{on}|^2 + |U_{np}|^2)U_{np} + U_{np}|U_{on}|^2 + U_{np}^*U_{on}^2. \quad (9.7)$$

В этом уравнении есть три слагаемых:

1. $(|U_{on}|^2 + |U_{np}|^2)U_{np}$ — предметная волна U_{np} , амплитуда которой промодулирована коэффициентом $(|U_{on}|^2 + |U_{np}|^2)$.
2. $U_{np}|U_{on}|^2 = 2U_{on}U_{np} \cdot \exp(i\varphi_{np})$ — объектная или предметная волна U_{np} , промодулированная коэффициентом U_{on} .
3. $U_{np}^*U_{on}^2 = U_{on}^2U_{np} \cdot \exp[i(2\varphi_{on} - \varphi_{np})]$ — волна, комплексно-сопряженная с объектной. Она несет информацию с отличной от объектной волны фазой. Поэтому мнимое изображение псевдоскопично или изображение видно как бы изнутри.

Голограмма обладает рядом уникальных свойств:

- Каждый фрагмент голограммы обладает свойством воспроизводить полную записанную информацию.
- Голограмма способна регистрировать и воспроизводить состояние поляризации предметной волны. Картина интерференции в этом случае характеризуется не изменением интенсивности поля, а модуляцией состояния поляризации. В этом случае слой с ли-

нейной поляризацией соседствуют со слоями, в которых поляризация циркулярна. В свою очередь эти слои граничат со слоями, где поляризация линейна, но уже в ортогональном направлении. Глаз это не различает, но такой способ эквивалентен записи двух голограмм на одной пластинке.

- Голограмма способна формировать обращенную волну, что позволяет компенсировать искажения, вносимые оптически неоднородными средами.
- Двумерная голограмма позволяет масштабировать изображение при изменении длины волны источника излучения.
- Объемная голограмма способна хранить информацию, пропорциональную трехмерной гармонике с характерным размером $\lambda \times \lambda \times \lambda$, где λ — длина волны. Таким образом, в объеме 1 см^3 при длине волны $\lambda = 5 \text{ мкм}$ помещается 10^{13} независимых гармоник.
- Ассоциативные свойства голограммы сводятся к следующему: если зарегистрировать два изображения А и В и осветить голограмму излучением от одного из изображений, например А, то получим другое изображение В. Возможно и наоборот.
- Голограмма осуществляет восстановление одного из изображений, если восстанавливаемое изображение смещается параллельно самому себе во входной плоскости. Это свойство Фурье-голограммы, получаемой от транспарата в фокальной плоскости линзы.

9.2.2. Голографическая элементная база

С помощью методов голографии можно создать *голограммные оптические элементы*, способные осуществлять различные преобразования волновых полей. Их действие основано на явлениях дифракции и интерференции света. Обычно голографические оптические элементы получают на основе фотопластинок с последующей обработкой. Различают *отражательные* и *пропускающие* голографические элементы.

Голограммные линзы образуются при регистрации интерференционной картины от двух сферических волн на плоских или сферических поверхностях. Голографические и классические линзы одного знака обладают хроматической аберрацией противоположных знаков. Поэтому их комбинация может использоваться для ахроматизации оптических систем. Отражательные голограммные линзы могут одновременно выполнять функции светоделиителя, светофильтра и формирователя изображений.

Голограммные фильтры создаются за счет угловой и спектральной селективности трехмерных голограмм.

Существуют также голографические оптические элементы, выполняющие функции *поляризаторов*. С помощью голографического *коррелятора* легко осуществляются операции по распознаванию образов.

В основе метода распознавания образов и идентификации лежит следующее явление. Если голограмму восстанавливать излучением зарегистрированного на ней объекта, то произойдет восстановление точечного опорного источника излучения. Появление точки на экране свидетельствует о том, что на голограмме находится именно тот объект, чьим излучением освещается голограмма.

9.2.3. Голографические запоминающие устройства

В *голографических запоминающих устройствах* (ЗУ) используют методы голографии для записи, хранения и воспроизведения информации. Информация может быть представлена

либо в двоичном коде (побитово), либо в виде изображения. Изображение может быть записано как плоское или объемное, амплитудное или фазовое, цветное или одноцветное. При голографической записи информации достигается высокая плотность записи ($\sim 10^5$ бит/мм²), высокая помехоустойчивость и надежность.

Классификация голографических ЗУ проводится по ряду независимых признаков:

- ☐ по способу ввода информации (одноканальный или многоканальный ввод);
- ☐ по принципу ввода информации (поэлементный ввод или ввод с накоплением информации);
- ☐ по способу записи (локальная запись на отдельном участке носителя или пространственно-распределенная запись, когда каждый элемент записывается по всей поверхности носителя, или объемная запись в объеме носителя);
- ☐ по длительности хранения информации (долговременные, оперативные или промежуточные);
- ☐ по способу обращения к памяти (выборка по адресу или ассоциативная выборка);
- ☐ по способу считывания (последовательное считывание или параллельное, одноэлементное).

Рассмотрим одну из отечественных разработок голографической информационной системы, реализованной в 90-е годы в электронной промышленности.

В основу этой информационно-поисковой системы был положен принцип согласованной фильтрации в сочетании с инвертированной организацией информационного массива, координатной привязкой и параллельной обработкой всего информационного массива (рис. 9.5). Голограмма в виде матрицы 100×100 элементов, представляющих собой массив кодированных поисковых образов документов, размещается в плоскости P_1 . В фокальной плоскости линз L_2 и L_3 в плоскости P_2 размещается пластина с матрицей голограмм-фильтров дескрипторов. Под дескриптором будем понимать лексическую единицу, которая выражает основное смысловое содержание документов. В плоскости P_3 поле сигналов взаимной корреляции дескрипторов и голограммы-фильтра дескриптора.

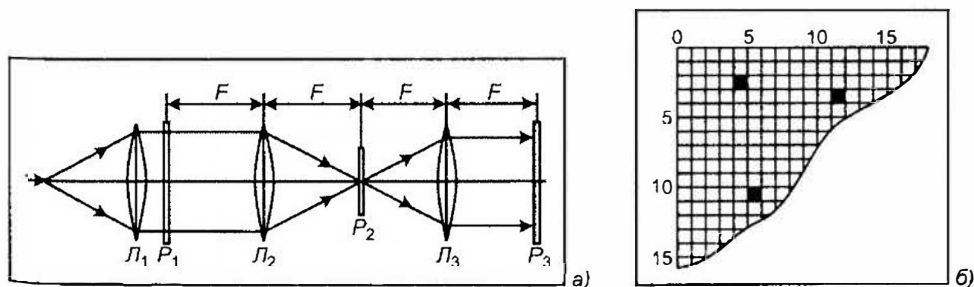


Рис. 9.5. Голографическая информационно-поисковая система:

а — схема согласованной фильтрации; б — выходная плоскость с сигналами

При вводе запроса в систему голограммы-фильтры каждого дескриптора запроса последовательно устанавливаются на оптической оси системы в плоскости P_2 . В выходной плоскости P_3 с координатной сеткой 100×100 элементов высвечиваются ячейки, координаты которых (номер строки и столбца) соответствует номеру документа в массиве (рис. 9.5, б). Результаты поиска регистрируются на фотопленке. Их суммирование по всем

дескрипторам запроса происходит путем последовательной регистрации сигнала корреляции на один кадр. Суммарная экспозиция каждой ячейки координатной сетки зависит от числа совпавших дескрипторов данного документа. Уровень порога регистрации устанавливался изменением общего времени экспозиции. Это позволяет не фиксировать сигналы малой амплитуды.

На следующем этапе каждая поисковая голографическая пластина комплектуется пластиной с матрицей 100×100 голограмм, на каждой из которых записаны поисковые данные одного документа и реферат объемом до 1000 знаков. При считывании рефератов релевантных документов пластина с микроголограммами рефератов передвигается относительно восстанавливающего луча лазера в соответствии с координатами сигналов, полученными в результате поиска. Восстановленные изображения текстовой информации регистрируются на микрофильм.

Голографический процессор реализован на гелий-неоновом лазере (ЛГ-52-1) / на длине волны $0,63 \text{ мкм}$ по стандартной схеме восстановления голограмм (рис. 9.6).

С помощью юстировочного зеркала лазерное излучение направляется на блок голограмм под углом $\sim 45^\circ$. Одновременно луч попадает на голограмму, восстанавливая изображение поисковой матрицы, соответствующее одному из дескрипторов запроса. Световой пучок от голограммы делится зеркалом на две части. Изображение поисковой матрицы просцируется одновременно на фотопленку и экран визуального контроля. С экрана считывается название дескриптора и его числовой код, который записывается в поисковую матрицу на восстановленной голограмме. Точная установка голограммы осуществляется координатным механизмом перемещения блока голограмм и совмещается с изображением поисковой матрицы с контрольными метками на экране.

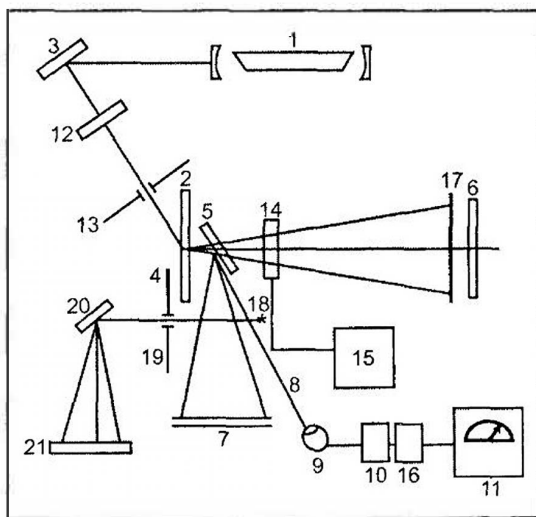


Рис. 9.6. Оптическая схема голографического процессора: 1 — лазер; 2 — блок голограмм; 3 — юстировочное зеркало; 4 — голограмма с координатным механизмом; 5 — делитель потока; 6 — фотокассета; 7 — экран визуального контроля; 8 — световод; 9 — фотоумножитель; 10 — усилитель; 11 — индикатор; 12 — поляризатор; 13 — диафрагма; 14 — фотозатвор с блоком управления (15); 16 — синхронный детектор; 17 — светофильтр; 18 — источник света; 19 — вторая диафрагма; 20 — зеркало; 21 — экран

Яркость изображения на каждой голограмме контролируется световодом и фотоумножителем, сигнал с которого усиливается усилителем и индицируется на индикаторе. По максимальному показателю индикатора выставляется максимум яркости восстановленного изображения при условии совпадения центра голограммы с осью луча лазера. Яркость восстановленного изображения может быть скорректирована с помощью вращающегося поляризационного фильтра. Диафрагма 13 позволяет снизить шумовые помехи.

Экспонирование изображения поисковой матрицы осуществляется с помощью фотоэавтора. Режим экспонирования подбирается так, чтобы суммарная экспозиция в выходной плоскости процессора не зависела от числа дескрипторов в запросе. Корректирующий светофильтр 17 предназначен для выравнивания яркости элементов восстановленного изображения перед плоскостью фотоматериала. Экспонирование производилось на высокочувствительную фотопленку (типа "Микрат-300").

Контроль положения блока голограмм осуществлялся с помощью точечного источника света 18, зеркала 20 и экрана с координатной сеткой 21. По этому оптическому пути высвечивается ячейка с кодом восстанавливаемой голограммы. Световое пятно на экране перемещается одновременно с блоком голограмм.

Элементы оптической системы голографического процессора были разработаны на основе микрометрических винтов оригинальной конструкции на жестких металлических основаниях для их фиксации в любом месте оптического стола (размер 1000×540 мм).

Поиск релевантной информации в информационном голографическом массиве состоял из нескольких этапов.

На этапе подготовки запроса из словаря дескрипторов (тезауруса) выбираются те дескрипторы, совокупность которых отражает содержание искомого документа. С целью повышения смыслового содержания искомых документов отдельным дескрипторам присваиваются весовые коэффициенты в пределах значений 0,7—0,9. Числовые коды дескрипторов представляют собой координаты соответствующих голограмм в голографическом блоке (номер строки ij и столбца mn в матрице голограмм). Матрица голограмм обычно состоит из 100×100 элементов. Формальная сторона запроса представляет собой список числовых кодов типа $ij - mn$, где i, j, m, n лежат в пределах от 1 до 100.

Ввод запроса производится с помощью линейных шкал с блока управления процессором. Экспонированный кадр микрофильма (или отдельного кадра в кассете) с результатами поиска по смысловой (или формальной) части запроса обрабатывается в оригинальном фотопроцессе и расшифровывается. При числовом декодировании с кадра микрофильма считываются координаты сигналов, которые служат адресами искомых документов. Поиск адресов документов производится с помощью таблиц соответствия координат и адресов документов.

Возможен диалоговый режим или многократный поиск с последовательной корректировкой запроса по результатам первичного поиска. Такая корректировка производилась по составу дескрипторов запроса, их весовым коэффициентам и порогу релевантности.

В случае поиска вторичных документов на одном голографическом блоке можно записать до 5×10^5 рефератов объемом свыше 10^3 печатных знаков каждый.

Чтение найденных документов производилось на специально сконструированном читальном голографическом аппарате с жидкокристаллическим экраном. Такой оптический процессор обладал высокой производительностью, оперативностью при поиске релевантной информации, возможностью оперативного получения копии документа. Однако слабым местом голографического процессора являлся встроенный жидкий фотохимический

процесс обработки фотопленок. Этот недостаток был частично ликвидирован за счет встраиваемой высокоразрешающей телевизионной техники, реализованной на ПЗС-структурах.

Отраслевая голографическая информационно-поисковая система (ИПС) состояла из двух подсистем. Главная подсистема находилась в отраслевом центре НТИ и осуществляла функции обработки входящего потока документов, формирования информационного массива на голографических носителях, его выборочное копирование и рассылку абонентам. В третьем контуре отраслевой ИПС хранились оригиналы документов, на основе которых по запросу абонентов высылались копии оригиналов, релевантных запросу (рис. 9.7).

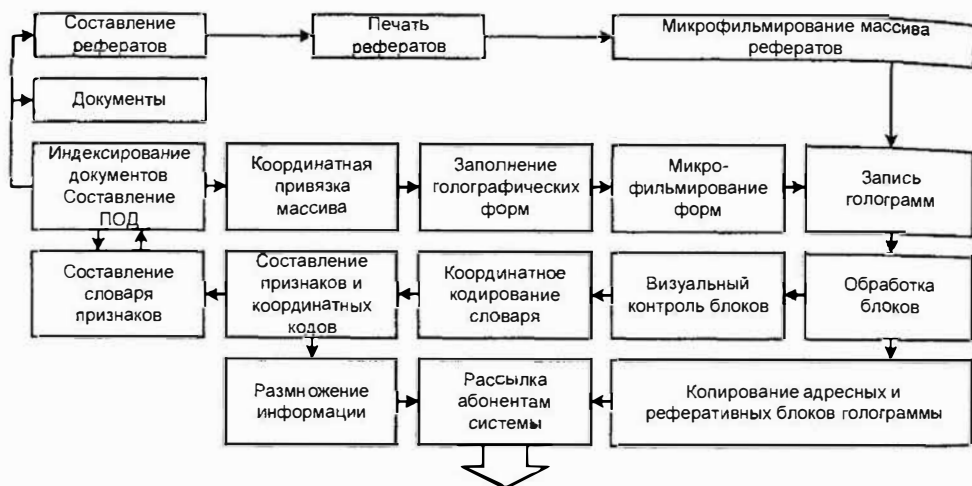


Рис. 9.7. Блок-схема обработки информации в отраслевом центре голографической ИПС

Вторая часть голографической ИПС формируется на основе локальных центров системы, оснащенной аппаратурой для поиска информации и копирования материалов из голографического массива по запросам. Первый контур системы (адреса документов) и второй (поиск рефератов документов) тиражируются и комплектуются вместе со специальным комплектом голографической аппаратуры.

Опытная эксплуатация системы была реализована на основе отраслевого фонда автоматизированного хранилища микрофиш, используемого в качестве третьего контура ИПС. Информационный фонд состоял из 6000 микрофиш по 60 кадров в каждом. Обработка информационного массива осуществлялась по 10^4 параллельным каналам, а общий информационный массив содержал 3×10^8 бит. Заказчик оперативно получал микрофильм с рефератами релевантных документов.

Анализируя голографические информационные системы, можно предположить, что объемом записанной информации может составить 10^8 — 10^9 бит, а скорость обработки информации в режиме поиска составит величину $\sim 10^{12}$ бит/с. Разрабатывается дисковая система ЗУ на одномерных голограммах. Запись осуществлялась на голографическом диске в виде дорожки голограмм. Интенсивность исследований в области голографических ЗУ снизилась из-за отсутствия оптимальных регистрирующих сред, а также из-за сильной конкуренции со стороны полупроводниковых цифровых запоминающих устройств.

Существует целый ряд задач, решение которых затруднительно и даже невозможно в рамках побитовой организации вычислительного процесса. Например, необходимо найти слабоформализованную или ассоциативную информацию в ЗУ.

При обычной организации поиска в ЗУ обращение происходит адресным способом. Однако существует ассоциативная организация ЗУ, позволяющая осуществить обращение сразу ко всем данным, хранящимся в ЗУ. Обращение производится по определенному формальному признаку.

Голографические методы обработки информации позволяют реализовать ассоциативный поиск информации.

Ассоциативные ЗУ голографического типа предполагается применять в "гетеро-ассоциативных" системах памяти, в которых была бы реализована концепция нейронных сетей и имелась "способность" к обучению.

На рис. 9.8 приведена схема ассоциативной памяти, в которой используются зеркала, отражающие волновой фронт. Такие зеркала устанавливаются по обе стороны от голограммы и образуют объемный резонатор. В нем возникает нелинейное взаимодействие вводимого фрагмента изображения и хранимых в банке данных изображений. Обеспечивается обратная связь, задание порога и соответствующего коэффициента усиления. Это устройство позволяет получить точную и полную информацию на выходе без обязательного ввода точной информации — достаточно 10% полного набора данных. Информация в таком ЗУ идентифицируется на основе общих ассоциаций, обработка информации идет с высокой степенью параллелизма: одновременно обрабатывается весь информационный массив.

9.2.4. Интерферометрические методы

Интерферометрические методы основаны на сравнении волновых фронтов реальных объектов с восстановленными голограммами образцовых объектов. При этом используется один оптический тракт. Возможно также сравнение волновых фронтов от одного и того же объекта, полученных в разные моменты времени.

В голографической интерферометрии используется два основных метода.

Метод двойной экспозиции предусматривает регистрацию двух голограмм объекта при одном и том же опорном луче. Первая голограмма регистрируется до приложения возмущающих сил. Вторая голограмма регистрирует состояния объекта после воздействия возмущающих сил через некоторый промежуток времени (рис. 9.9).

При восстановлении двух волновых фронтов, записанных через некоторый промежуток времени, получается изображение объекта с интерференционными полосами в местах деформации. Специалисты называют этот метод "интерференцией Наполеона с Клеопатрой". Преимущество этого метода заключается в возможности измерения интерференционных полос и численного определения деформации. Метод применим для быстропротекающих процессов.

Метод наблюдения в реальном масштабе времени производят, непосредственно сравнивая ранее записанное и восстановленное изображения объекта с текущим состоянием. Этот метод позволяет управлять процессом деформации. Специалисты называют его "живым". Он пригоден для исследования медленных процессов.

Разработаны методы *спекл-интерферометрии*, основанные на явлении диффузного рассеяния когерентного лазерного излучения шероховатой поверхностью. Если изменение



Рис. 9.8. Ассоциативное оптическое запоминающее устройство с обращением волнового фронта



Рис. 9.9. Голографическая интерферометрия мембраны, подвергнутой давлению газа. Проведена расшифровка интерференционных полос

рельефа поверхности составляет порядка длины волны излучения, то имеет место диффузное рассеяние света, напоминающее авантюрин. Заметим, что авантюрин представляет собой зернистый кварц с характерным мерцающим блеском. Такая пятнистая структура — распределения интенсивности когерентного света получила название *спеклы*.

Каждая точка отражающей лазерное излучение поверхности в соответствии с принципом Гюйгенса является источником сферических волн. Амплитуда напряженности поля в наблюдаемой точке формируется путем суммирования векторов со случайными фазами. При изменении положения наблюдателя интенсивность рассеиваемого когерентного света принимает различные значения. Это явление называется *спекл-эффектом*.

Спекл-интерферометрия позволяет измерять смещения с точностью порядка длины волны лазерного излучения. Разработано два метода спекл-интерферометрии.

Корреляционная спекл-интерферометрия основана на анализе интерференционного распределения оптического поля при одновременном наложении волн, рассеянных от поверхности исследуемого объекта в начальном и смещенном положениях. С этой целью используется интерферометр типа интерферометра Майкельсона, в котором вместо зеркал используются шероховатые поверхности. Интерференционную картину корреляции спектров исследуют с помощью кино- или телекамеры, имеющей разрешение порядка 50 линий/мм. Этот метод позволяет определять форму поверхности и анализировать ее деформации.

В основе метода *спекл-фотографии* лежит одновременная регистрация двух рассеянных предмостом волн на одной фотопластинке. Одна волна формируется рассеянием излучения объектом в начальном положении, а другая волна формируется уже смещенным объектом. Измерив относительный сдвиг спекл-структур на проявленной фотопластинке, можно определить смещение объекта. Определение смещения спекл-структуры осуществляется с помощью Фурье-преобразования функции пропускания дважды экспонированной фотопластинки. С этой целью она освещается плоской волной и исследуется световой

поток в фокальной плоскости линзы. Этот метод позволяет исследовать динамику смещения с использованием стробоскопической или двухимпульсной спекл-фотографии.

9.3. Когерентные оптические системы аналоговой обработки информации

Основными элементами оптических систем аналоговой обработки информации являются тонкие сферические линзы, выполняющие преобразование Фурье (рис. 9.10).

Изложение предварим некоторыми сведениями из математики.

Если задана действительная или комплексная функция $f(x)$ действительной переменной x , меняющейся в пределах от $-\infty$ до $+\infty$, такая, что существует интеграл

$$\int_{-\infty}^{+\infty} |f(x)| dx, \quad (9.8)$$

то Фурье-образ $F(u)$ этой функции $f(x)$ определяется выражением

$$F(u) = \int_{-\infty}^{+\infty} f(x) \exp(2\pi i u x) dx. \quad (9.9)$$

По определению свертка $f(t)$ получается из двух функций $f_1(t)$ и $f_2(t)$ с помощью интегральной операции

$$f(t) = \int_{-\infty}^{+\infty} f_1(\tau) f_2(t - \tau) d\tau, \quad (9.10)$$

где τ — независимая (текущая) переменная, а t представляет ряд последовательных значений сдвига функции $f_2(t)$ относительно $f_1(t)$. Интеграл от произведения перекрывающихся частей функции $f_1(t)$ и $f_2(t)$ определяется при одном таком сдвиге t . Другими словами, функция $f(t)$ в аналитической или графической форме получается путем последовательного определения интеграла от произведения двух функций $f_1(t)$ и $f_2(t)$, когда эти две функции последовательно сдвигаются друг относительно друга. Значения интеграла находятся для каждого сдвига t функции $f_2(t)$ относительно $f_1(t)$, включая значение $f(0)$, которое соответствует значению $t = 0$, когда две функции $f_1(t)$ и $f_2(t)$ "совпадают", т. е. имеют общую абсциссу.

Ради компактности интеграл свертки часто удобно записать в виде

$$f(t) = f_1(t) \otimes f_2(t). \quad (9.11)$$

Необходимо подчеркнуть существенное различие между сверткой, с одной стороны, и функцией корреляции — с другой, особенно в связи с тем, что по виду они очень близки.

Если одна из двух функций является сопряженной с некоторой функцией, то в качестве подходящего представления выступает функция корреляции, а не свертка.

По определению функция автокорреляции $\varphi_{11}(\tau)$ равна

$$\varphi_{11}(\tau) = \int_{-\infty}^{+\infty} f_1(t) f_1^*(t + \tau) dt, \quad (9.12)$$

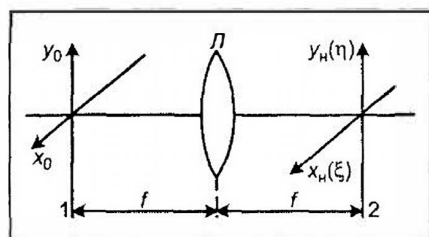


Рис. 9.10. Схема операции преобразования Фурье: Л — линза; 1 — входная фокальная плоскость; 2 — выходная фокальная плоскость

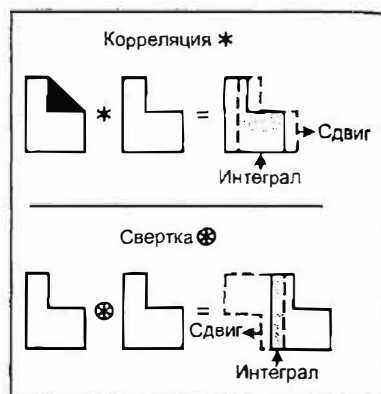


Рис. 9.11. Графическая интерпретация операций корреляции и свертки

а функция кросс-корреляции φ_{12} определяется выражением:

$$\varphi_{12}(\tau) = \int_{-\infty}^{+\infty} f_1(t) f_2^*(t + \tau) dt. \quad (9.13)$$

Аналогично имеем

$$\varphi_{21}(\tau) = \int_{-\infty}^{+\infty} f_1^*(t + \tau) f_2(t) dt. \quad (9.14)$$

Графическая интерпретация операции свертки-корреляции дана на рис. 9.11.

Рассмотрим процесс преобразования Фурье, осуществляемый тонкой линзой (рис. 9.10). Если на вход такой системы подать оптический сигнал $U_0(x_0, y_0)$, то на выходе получится сигнал, связанный с входным сигналом следующим соотношением:

$$\begin{aligned} U_H(x_H, y_H) &= \frac{1}{i\lambda f} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U_0(x_0, y_0) \exp\left[-\frac{2\pi i}{\lambda f}(x_H, x_0 + y_H, y_0)\right] dx_0 dy_0 = \\ &= k_{0H} U_0\left(\frac{x_H}{\lambda f}, \frac{y_H}{\lambda f}\right). \end{aligned}$$

Таким образом, выходной сигнал рассматриваемой простейшей оптической системы с точностью до постоянного множителя совпадает с Фурье-образом входного сигнала. Поэтому выходную плоскость такой системы называют спектральной или Фурье-плоскостью. Пространственно-частотные координаты ξ и η в этой плоскости определяются отношением:

$$\begin{aligned} \xi &= \frac{x_H}{\lambda f}, \\ \eta &= \frac{y_H}{\lambda f}. \end{aligned} \quad (9.15)$$

Фурье-образ входного оптического сигнала существует в виде физически реального пространственного распределения комплексных амплитуд света.

Линза L_2 осуществляет преобразование Фурье, поэтому в спектральной плоскости системы непосредственно перед операционным фильтром распределение комплексных амплитуд света пропорционально Фурье-образу входного сигнала, т. е.

$$U_H(\xi, \eta) = \frac{k_0 A}{i\lambda f} U_0(\xi, \eta).$$

Амплитудный коэффициент пропускания операционного фильтра будет равен

$$t_H = k_H H(\xi, \eta),$$

где k_H — комплексная константа.

Функция $H(\xi, \eta)$ соответствует математической операции, которую необходимо выполнить над входным сигналом, и ее называют передаточной функцией фильтра. В результате пространственной фильтрации получается сигнал, описываемый распределением комплексных амплитуд:

$$U_H^*(\xi, \eta) = \frac{k_0 k_H A}{i\lambda f} U_0(\xi, \eta) H(\xi, \eta). \quad (9.16)$$

Отфильтрованный сигнал подвергается новому преобразованию Фурье с помощью линзы L_3 . В результате в выходной плоскости системы свет будет иметь распределение:

$$U_B(x_B, y_B) = k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U_0(\xi, \eta) H(\xi, \eta) \exp[2\pi i(x_B \xi + y_B \eta)] d\xi d\eta, \quad (9.17)$$

где

$$k = -\frac{k_0 k_H A}{(\lambda f)^2}.$$

Направление координатных осей в выходной плоскости системы выбраны противоположно направлениям осей координат во входной плоскости для того, чтобы учесть инверсию, которая получается в результате двух последовательных преобразований Фурье и выражается соотношением

$$F\{F[U(x, y)]\} = U(-x, -y),$$

где F — символ операции свертки.

Соотношение (9.17) можно записать в виде свертки:

$$U_B(x_B, y_B) = k U_0(x_0, y_0) * h(x, y) = k \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U_0(x_0, y_0) h(x_0 - x, y_0 - y_B) dx_0 dy_0. \quad (9.18)$$

где $h(x, y)$ — обратное преобразование Фурье передаточной функции фильтра. Таким образом, оптическая система, представленная на рис. 9.11, способна выполнять линейные интегральные преобразования типа свертки, описываемые уравнением (9.18).

Метод пространственной фильтрации позволяет с помощью преобразований Фурье производить различные математические операции.

На рис. 9.12 приведена схема процессора для проведения оптической пространственной фильтрации.

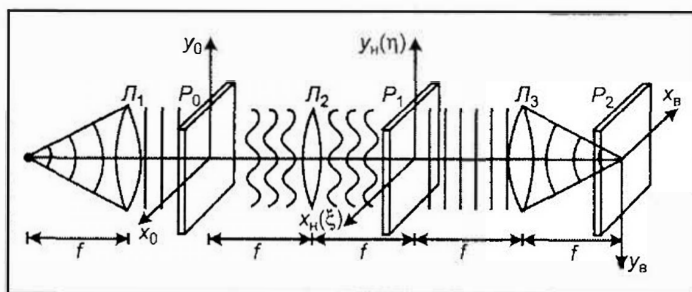


Рис. 9.12. Оптический процессор: L_1, L_2, L_3 — линзы; P_0 — входная фокальная плоскость; P_1 — спектральная фокальная плоскость; P_2 — выходная фокальная плоскость

Входная плоскость освещается плоской световой волной с амплитудой A . Во входной плоскости размещается прозрачный транспарант с заданной информацией. Устройство ввода формирует изображение, соответствующее входной информации $U_0(x_0, y_0)$, так что его можно характеризовать амплитудной функцией пропускания

$$t_0 = k_0 U_0(x_0, y_0).$$

Следовательно, входной оптический сигнал системы имеет комплексную амплитуду:

$$U_0^* = k_0 A U_0(x_0, y_0). \quad (9.19)$$

Описанный процессор позволяет проводить операции по распознаванию образов. Если необходимо из набора произвольных двумерных транспарантов выбрать одну определенную, то эти транспаранты последовательно вводятся в плоскость P_0 процессора. В плоскость P_1 помещается транспарант, являющийся согласованным фильтром для искомого образа. Свойства такого фильтра связаны с компенсацией искажения волнового фронта в том случае, когда падающая на него световая волна является двумерным Фурье-спектром искомого образа. Согласованный фильтр точно компенсирует кривизну падающей на него волны (рис. 9.12). После плоскости P_1 световое поле оказывается плоской волной и фокусируется линзой L_3 в яркое пятно в выходной фокальной плоскости. Процессор позволяет по появлению яркого пятна в центре выходной плоскости P_2 зафиксировать наличие в плоскости P_0 искомого образа.

На рис. 9.13 показан результат согласованной пространственной фильтрации части текста китайских иероглифов. С помощью согласованного фильтра, изготовленного на искомый иероглиф, выявлен искомый образ в тексте. Чем сложнее по структуре и форме сигнал, тем эффективнее решается задача по его идентификации. Поэтому этот метод нашел свое применение в криминалистике и дактилоскопии.

Оптические методы обработки нашли применение при обработке спектрограмм, рентгенограмм, изображений с электронных микроскопов и т. д.

В промышленности использование метода согласованной фильтрации используется при контроле качества фотошаблонов, интегральных схем. Фотошаблон является мультиплицированным изображением маски в виде матрицы. Контроль состоит в обнаружении дефектов отдельных масок, а также дефектов в виде царапин, пылинок. Ручной контроль весьма трудоемкая операция. Метод пространственной фильтрации позволяет визуализировать дефекты с помощью пространственного фильтра. Контроль набора фотошаблонов осуществляют методом сканирования.

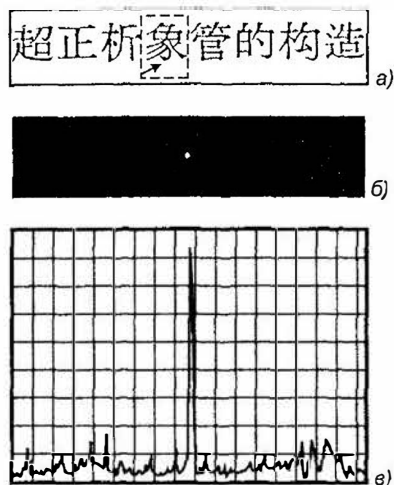


Рис. 9.13. Выявление заданного иероглифа в тексте методом пространственной фильтрации:
а — китайский текст; б — иероглиф "Суперортикон"; в — спектр выходного сигнала

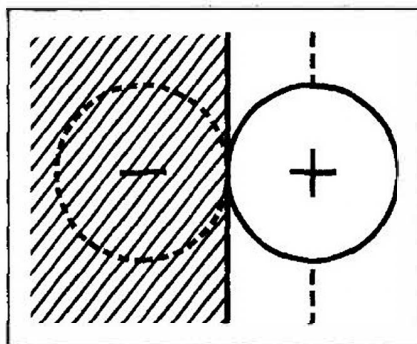
Контрольные вопросы

1. Что такое волновой пакет и как он связан со спектром излучения?
2. Что такое когерентность?
3. Как можно описать оптический сигнал?
4. Что такое голография и как происходит процесс голографирования?
5. Какими уникальными свойствами обладает голография?
6. Какие оптические элементы можно создать методами голографии?
7. Что представляет собой голографическое ЗУ?
8. В чем заключается сущность метода пространственной фильтрации?
9. Какие методы голографической интерферометрии вы знаете?
10. Что представляет собой оптический процессор?

Рекомендуемая литература

1. Акаев А. А., Майоров С. А. Оптические методы обработки информации. — М.: Высшая школа, 1988.
2. Гуревич С. В., Константинов В. Б., Соколов В. К., Черных Д. Ф. Передача и обработка информации. — М.: Советское радио, 1978.
3. Микаэлян А. Л. Оптические методы в информатике. — М.: Наука, 1990.
4. Строук Дж. Введение в когерентную оптику и голографию. — М.: Мир, 1967.





ЧАСТЬ IV

Функциональная электроника

Краткая историческая справка

1. Начала функциональной электроники
2. Функциональная акустоэлектроника
3. Функциональная диэлектрическая электроника
4. Функциональная полупроводниковая электроника
5. Функциональная магнитоэлектроника
6. Функциональная оптоэлектроника
7. Функциональная молекулярная электроника
8. Приборы функциональной электроники второго поколения



Краткая историческая справка

Функциональная электроника оформилась как научное направление в электронике в последней четверти XX столетия. Формированию этого направления способствовала лавина открытий в физике и электронике. Исследователям стало ясно, что помимо традиционных носителей информационного сигнала, таких как свободные электроны в вакууме, твердом теле, ионы в плазме, связанные электроны в средах различного агрегатного состояния, существуют и другие, неявно связанные с электронами. К таким носителям можно отнести различного типа волны в различных средах, домены, квазичастицы, зарядовые пакеты электронов и т. п. Эти носители получили название *динамические неоднородности*. В природе их много, исследователи познакомились с ними давно...

"Я следил за движением баржи, которую быстро тянула по узкому каналу пара лошадей, когда баржа неожиданно остановилась; но масса воды, которую баржа привела в движение, не остановилась. Вместо этого она собралась около носа судна в состоянии бешеного движения, затем неожиданно оставила его позади, катясь вперед с огромной скоростью и принимая форму большого одиночного возвышения, т. е. округлого, гладкого и четко выраженного водяного холма, который продолжал свой путь вдоль канала, несколько не меняя своей формы и не снижая скорости. Я последовал за ним верхом, и когда я нагнал его, он по-прежнему катился вперед со скоростью приблизительно восемь или девять миль в час, сохранив свой первоначальный профиль возвышения длиной около тридцати футов и высотой от фута до фута с половиной. Его высота постепенно уменьшалась, и после одной или двух миль погони я потерял его в изгибах канала. Так в августе 1834 года мне впервые удалось столкнуться с необычайным и красивым явлением, которое я назвал волной трансляции; теперь это название общепринято" — так инженер-изобретатель из Шотландии Джон Скотт Рассел описал уединенную волну в своей научной работе "Докладе о волнах".

Рассел установил следующие основные свойства уединенных волн:

- ☐ постоянство скорости и неизменность формы уединенной волны;
- ☐ зависимость скорости v от глубины канала h и высоты волны y_0

$$v = \sqrt{g(h + y_0)},$$

где g — ускорение свободного падения;

- ☐ процесс распада большой волны на две или более, но это тоже уединенные волны;
- ☐ наблюдаются только волны повышения;
- ☐ большие первичные волны трансляции проходят друг через друга без каких-либо изменений.

Так в 1834 году была открыта, пожалуй, первая динамическая неоднородность — уединенная волна в воде.

Работа Дж. С. Рассела была раскритикована, и о ней долго не вспоминали. Но еще при жизни Рассела лорд Релей приближенно математически описал уединенную волну в мел-

кой воде. Голландские ученые Дидерик Иоханнес Кортевег и его ученик Густав де Фриз в 1895 году нашли уравнение, достаточно точно описывающее эффекты, наблюдаемые Дж. С. Расселом.

Уравнение Кортевега де Фриза получило название *КдФ-уравнение* и оно послужило стартовой площадкой для построения глубокой и стройной математической теории уединенной волны. Однако семь десятилетий эта теория спросом не пользовалась. Может быть потому, что за свойствами уединенной волны исследователи не увидели ее поразительное сходство с частицами.

Волной или частицей является уединенная волна?

Американские исследователи М. Крускал и Н. Забуски в 1965 году изучали явление столкновения уединенных волн. Исследования проводили путем моделирования на вычислительных машинах, и пришли к выводу о поразительном сходстве свойств уединенных волн с твердыми частицами. Именно они из понятия "уединенная волна" убрали слово волна, а определению уединенной "solitari" придали звучание элементарной частицы или soliton. Правда, первоначально *солитон* назывался солитрон. Узнав о существовании фирмы "Солитрон" исследователи не стали вступать с ней в тяжбу за право приоритета на имя и ограничились ныне общепринятым названием.

Солитоны формируются в консервативных системах без диссипации энергии, но с дисперсией. Дисперсионное распыливание компенсируется нелинейными свойствами среды. Поэтому поведение солитонов характеризуется законами сохранения, дающее им свойства частиц.

Принципиально иными свойствами обладают уединенные состояния в диссипативных неравновесных системах, в которых процесс установления равновесного состояния носит не осциллирующий, а релаксационный характер. В таких диссипативных системах при условии подкачки энергии от внешнего распределенного источника возникают также уединенные состояния, получившие название *автосолитонов*. Автосолитоны отличаются от солитонов так же, как автоколебания от колебаний, автоволны от волн.

Итак, автосолитон представляет собой уединенное собственное состояние неравновесной системы и в этом смысле его можно рассматривать как одно из явлений самоорганизации.

В последние годы выяснилось, что природа солитонов и автосолитонов весьма разнообразна.

В высокочастотном газовом разряде возбуждаются автосолитоны в виде движущейся уединенной страты. В полупроводниках и полупроводниковых структурах солитоны могут возбуждаться в твердотельной плазме. В композитных сверхпроводниках автосолитон представляет собой резистивную область, окруженную сверхпроводящей фазой. В магнитных материалах с диссипацией при наличии внешнего источника энергии образуется автосолитон в виде бегущих магнитных доменов, или магнитных вихрей. В средах, в которых протекают автокаталитические реакции типа Белоусова — Жаботинского, образуются автоволны и автосолитоны.

А между тем в электронике происходили открытия так или иначе связанные с динамическими неоднородностями.

В 1964 году молодые сотрудники Института радиоэлектроники АН СССР Юрий Васильевич Гуляев и Владислав Иванович Пустовойт впервые предсказали явление взаимодействия акустических поверхностных волн и электронов в твердом теле и тем самым открыли новое направление в электронике — *акустоэлектронику* (рис. 1).



Рис. 1. Академики Ю. В. Гуляев и Ж. И. Алфиров

В 1968 году в ИРЭ АН СССР и независимо в США был открыт эффект чисто сдвиговых поверхностных акустических волн, которые не обладают дисперсией. Их назвали волнами Гуляева — Блюкштейна по имени авторов открытия. На основе открытых ими волн было создано перспективное поколение акустоэлектронных приборов СВЧ-диапазона длин волн.

Затем последовало открытие Ю. В. Гуляевым акустомагнитоэлектрического эффекта, открытие поперечного акустоэлектрического эффекта. Все эти, а также другие работы в России и за рубежом позволили сформировать не только научное направление — акустоэлектронику, но и промышленное изготовление акустоэлектронных линий задержек, акустоэлектронных фильтров, других акустоэлектронных изделий для сотовых телефонов телевидения, бытовой техники. Объем акустоэлектронного рынка оценивается в 10 миллиардов долларов.

В 1963 году американский физик Дж. Б. Ганн обнаружил токовые неустойчивости в виде электрических доменов в образцах GaAs и InSb с электронной проводимостью. В 1966 году уже появились первые коммерческие *диоды Ганна*, выпущенные фирмой International Semiconductor Inc. Это открытие позволило начать освоение СВЧ-диапазона длин волн полупроводниковыми приборами.

В 1969 году исследователи фирмы Bell Laboratories обнаружили в магнитной пленке так называемые магнитные пузырьки в виде цилиндрических магнитных доменов. Пузырьки (bable) оставались стабильными при приложении внешнего статического магнитного поля. Стало очевидным, что наличие домена можно рассматривать как логическую единицу, а его отсутствие как логический ноль. Родилась новая электронная память, появилась возможность создавать логические схемы на новом принципе.

Сотрудники Bell Laboratories задумались над возможностью создания полупроводниковых электронных устройств на динамических неоднородностях, которые могли бы перемещаться в полупроводнике и переносить заряд подобно тому, как переносят логические единицы магнитные цилиндрические домены. В 1970 году В. С. Бойл и Дж. Е. Смит разработали *приборы с зарядовой связью* (ПЗС). Принцип зарядовой связи позволял хранить носители в областях или потенциальных ямах под электродами, в которых тип проводимости носителей отличался от типа проводимости подложки. Путем подачи соответствующих импульсов на электроды носители или зарядовые пакеты можно было переме-

шать в любом направлении, где расположены электроды. Электроды отделялись друг от друга узкими зазорами, что позволяло сохранить взаимовлияние электрических полей соседних элементов. Первоначально аналогичные структуры были реализованы на биполярных, а затем на МДП-транзисторах. Они получили название "пожарные цепочки". Это изобретение открыло возможность создать ЗУ с неслыханной по тем временам плотностью упаковки элементов. Были разработаны уникальные по своим свойствам линии задержки сигналов, рекурсивные и трансверсальные фильтры, процессоры аналоговых сигналов. Однако основное применение ПЗС-приборы нашли в устройствах получения и обработки изображений.

К концу прошлого века благодаря открытиям в разных областях электроники стало известно о множестве видов динамических неоднородностей различной физической природы.

Ну а что солитон Рассела? Он распался, стал многоликим. Каждый имеет свое лицо, характерное для физической природы, каждый имеет родственников по математической линии.

К одномерным солитонам относят классические уединенные волны в жидкостях, доменные стенки в ферромагнетиках, оптические солитоны, моды коллективной проводимости в органических полупроводниках, кванты магнитного потока в сверхпроводниках, ионно-звуковые и ленгмюровские солитоны в плазме и т. п.

К двумерным и многомерным солитонам следует отнести дислокации в жидких кристаллах, вихревые структуры в тонком слое сверхтекучей жидкости, вихри Абрикосова в сверхпроводниках 2-го рода, каналы самофокусировки в нелинейной оптике и т. п.

Решая различные уравнения для солитонов, математики породили кинки, бризеры, скирмионы, лампы... Теперь осталось лишь найти их реальное воплощение, что и является задачей физиков-исследователей.

В конце семидесятых годов группа отечественных исследователей во главе с Я. А. Федотовым и А. А. Васенковым пришли к заключению, что использование динамических неоднородностей в качестве носителей информационного сигнала в электронных устройствах позволит существенно улучшить их характеристики. Это направление в микроэлектронике и получило название *функциональная электроника*.

Среди перспективных задач функциональной электроники является исследование возможности интеграции различных физических эффектов в одном устройстве. В результате изучения взаимодействия динамических неоднородностей различной физической природы в различных средах открывается перспектива физической интеграции. Функциональная электроника не противопоставляется традиционной схемотехнической микроэлектронике. Напротив, эти два направления в интегральной электронике будут развиваться, взаимно дополняя друг друга, создавая сложные функционально интегрированные системы и устройства. Но в арсенале функциональной электроники еще много неиспользованных динамических неоднородностей, на основе которых можно создать приборы и устройства с заданными функциями, с заданными параметрами.

1. Начала функциональной электроники

Микроэлектроника представляет собой область электроники, связанную с исследованиями поведения заряженных частиц в твердом теле под воздействием электрических, магнитных, электромагнитных, тепловых полей, а также с созданием приборов и устройств в микроминиатюрном исполнении с использованием групповой технологии изготовления.

В микроэлектронике предполагается интеграция элементарных электронных приборов (резисторов, конденсаторов, диодов, транзисторов). Поэтому синонимом микроэлектроники является понятие "интегральная электроника".

Говоря о микроэлектронике, имеют в виду микроэлектронные размеры элементов, говоря об интегральной электронике — интеграцию этих элементов на кристалле микросхемы. Дальнейшее развитие микроэлектронных приборов связано с уменьшением размеров элементарных приборов до субмикронных размеров и переход в нанометровый масштаб измерений. Таким образом, микроэлектронные приборы превращаются в нанoeлектронные. При этом утрачивается групповая технология их изготовления.

С другой стороны, при интеграции на одном кристалле не только элементов, но и физических явлений и эффектов увеличиваются функциональные возможности приборов и устройств интегральной электроники. При этом используется уже не только схемотехнические решения для обработки и хранения информации, но и физические носители информационного сигнала — динамические неоднородности различной физической природы. Это направление в развитии электроники называют функциональной электроникой.

Следует при этом еще раз подчеркнуть, что если с переходом в субмикронный диапазон размеров в нанoeлектронике утрачивается принцип групповой технологии производства элементарных электронных приборов, то в функциональной электронике по-прежнему сохраняется принцип групповой технологии.

Динамическая неоднородность представляет собой локальный объем на поверхности или внутри среды с отличными от ее окружения свойствами, которая не имеет внутри себя статических неоднородностей и генерируется в результате определенных физико-химических процессов. Динамическая неоднородность может быть локализована или перемещаться по рабочему объему континуальной среды в результате взаимодействия с различными физическими полями или динамическими неоднородностями такой же или другой физической природы.

В процессе перемещения динамической неоднородности может происходить, например, перенос информации. Деградация динамической неоднородности не приводит, как правило, к потерям и сбоям в процессах обработки информации.

Известно большое количество динамических неоднородностей различной физической природы (рис. 1.1). Это ансамбли заряженных частиц и квазичастиц (зарядовые пакеты, флюксоны и т. п.), домены (сегнетоэлектрические домены, домены Ганна, цилиндрические магнитные домены и т. п.), динамические неоднородности волновой природы (поверхностные акустические волны (ПАВ), магнитные статические волны (МСВ), волны пространственного заряда (ВПЗ), волны зарядовой плотности (ВЗП) и т. д.).

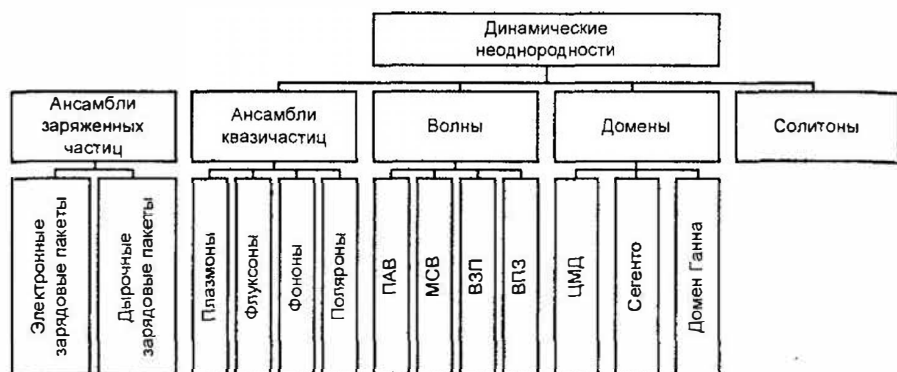


Рис. 1.1. Некоторые типы динамических неоднородностей

Идея использования динамических неоднородностей в приборах по обработке и хранению информации является основополагающей, и ее развитие привело к становлению функциональной электроники. Это не схемотехническое направление в микроэлектронике, использующее в качестве носителей информации динамические неоднородности, основной тенденцией развития которой является интеграция функциональных возможностей приборов и устройств.

Функциональная электроника представляет собой область интегральной электроники, в которой изучается возникновение и взаимодействие динамических неоднородностей в континуальных средах в совокупности с физическими полями, а также создаются приборы и устройства на основе динамических неоднородностей для целей обработки, генерации и хранения информации.

В зависимости от типа используемой динамической неоднородности, континуальной среды, той или иной комбинации физических полей или явлений различают направления в функциональной электронике, например, функциональная акустоэлектроника, функциональная магнетозлектроника, функциональная оптоэлектроника, функциональная диэлектрическая электроника, молекулярная электроника и т. п. Объединяющим их признаком является динамическая неоднородность как носитель или транслятор или хранитель информации. Например, традиционная полупроводниковая схемотехническая электроника отличается от полупроводниковой функциональной электроники носителем информационного сигнала. В приборах схемотехнической микроэлектроники — аналоговых или цифровых ИС — информация хранится или обрабатывается в ячейках в виде заряда, потенциала или тока определенного уровня на определенной статической неоднородности.

Напротив, в ПЗС-матрицах, относящихся по своей физической природе к изделиям функциональной полупроводниковой электроники, информация хранится (либо обрабатывается) в виде динамической неоднородности — зарядового пакета, состоящего из электронов или дырок. Статические неоднородности, различные схемы обрания в этих изделиях играют вспомогательную роль.

В функциональной электронике пока не существует принципиальных ограничений, связанных с размерами статических или динамических неоднородностей. Сравним изделия традиционной схемотехнической и функциональной электроники по быстродействию.

В изделиях схемотехнической электроники перенос информационного сигнала происходит побитово по линиям межсоединений, что снижает помехоустойчивость и надежность

изделий. В изделиях функциональной электроники массив информационных сигналов может быть обработан одномоментно весь, целиком, не обязательно в виде отдельных битов информации. Возможно создание устройства, позволяющего производить обработку информации в аналоговом и цифровом видах одновременно, поэтому в устройствах обработки информации на принципах функциональной электроники достигается производительность более 10^{15} операций в секунду.

В изделиях функциональной электроники перенос информации, как правило, осуществляется в континуальной среде, а не в линиях межсоединений. В схемотехнической электронике с ростом степени интеграции и уменьшением топологической нормы возникает проблема "тирании межсоединений". Она связана с резким увеличением интегральной площади, занимаемой на кристалле межсоединениями (более 60%), деградацией электрических параметров линий межсоединений, ростом энергии на перезарядку линий межсоединений, влиянием погонной емкости линий межсоединений и волнового сопротивления на частотные характеристики схемы, необходимостью многоуровневой разводки при большом числе линий межсоединений. В приборах функциональной электроники межсоединения выполняют функции вспомогательных связей и проблема "тирании межсоединений", характерная для интегральных схем, в этом случае не является ключевой проблемой.

Замечательным свойством приборов функциональной электроники является использование в процессах обработки информации элементарных функций высшего порядка. В качестве примеров элементарных функций высшего порядка назовем: Фурье-преобразование; интегральное преобразование Лапласа; операцию свертки; операцию корреляции; автокорреляцию; управляемую задержку информационного сигнала; хранение информации, в том числе в виде многобитовых носителей; фильтрацию информационного сигнала; когерентное сложение сигналов; ответвление информационных сигналов; комбинированную обработку информационных сигналов и т. д. В то же время в изделиях функциональной электроники элементарные функции высшего порядка могут быть реализованы наряду с традиционными элементарными функциями типа И, ИЛИ, НЕ и их различными комбинациями.

В активной среде прибора функциональной электроники может храниться и одновременно обрабатываться большой объем информации. Поэтому изделие функциональной электроники может рассматриваться как процессор, выполняющий элементарные функции высшего порядка. На рис. 1.2 приведены некоторые примеры процессоров и их общие схемы.

Важно отметить, что обработка информации в таком типе процессорах происходит в аналоговом виде, без перевода аналогового сигнала в цифровой код и обратно. При такой обработке передача информации может осуществляться без проводников или линий межсоединений.

Устройства памяти реализуются путем упорядоченного сохранения в континуальной среде динамических неоднородностей, каждая из которых служит битом информации. Возможно создание запоминающих устройств и на основе многобитовых динамических неоднородностей.

Как уже было сказано, в приборах и устройствах функциональной электроники информационный массив может быть обработан весь и сразу в одномоментном процессе. При этом не обязательно использовать последовательную побитовую обработку двоичной информации. Это эквивалентно случаю предельного распараллеливания процесса обработки массива информации. Например, акустооптический процессор обеспечивает произ-

водительность 10^{10} — 10^{12} операций/с, в то время как специальная микросхема быстрого преобразования Фурье позволяет обеспечить производительность порядка $2,5 \times 10^8$ операций/с. Выигрыш на несколько порядков в производительности вполне существенен.

Рассмотренные процессоры относятся к изделиям функциональной электроники первого поколения. В них используется один вид динамических неоднородностей в одной континуальной среде. Если используются два или более вида динамических неоднородностей в равных средах, то такие изделия относятся ко второму поколению.

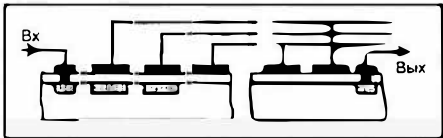
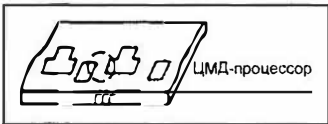
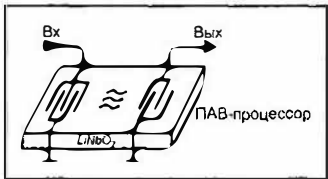
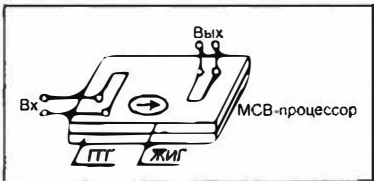
Динамическая неоднородность	Физическая природа	Общая схема
Дырки	Зарядовые пакеты	
ЦМД	Домены	
ПАВ	Волны	
МСВ		

Рис. 1.2. Примеры процессоров, реализованных на динамических неоднородностях различной физической природы

Анализируя устройства функциональной электроники, можно выделить некоторые элементы, характерные для всех конструкций.

Во всех приборах и изделиях функциональной электроники используются динамические неоднородности различного вида и различной физической природы. Это *первый* элемент в предложенной модели прибора функциональной электроники. Так, в акустоэлектронных устройствах используются динамические неоднородности в виде поверхностной акустической волны (ПАВ); в полупроводниковых приборах с зарядовой связью — зарядовые пакеты электронов или дырок; в приборах магнитоэлектроники — магнитоэлектронные волны (МСВ) и т. д.

Все виды динамических неоднородностей генерируют, обрабатывают или хранят информацию в континуальных средах, как правило, в твердом теле. Континуальная среда является *вторым* элементом модели. Вообще говоря, она может иметь любое агрегатное состояние. Наши интересы в области микроэлектроники сосредоточены в использовании твердого тела. По своим физико-химическим свойствам среда должна быть достаточно однородной на всем тракте распространения информационного сигнала. Статические неоднородности, имеющиеся на поверхности или внутри континуальной среды, служат только для управления динамическими неоднородностями и не используются для обработки и хранения информации. Динамические неоднородности, обрабатывая информационный сигнал в континуальных средах, не меняют их физико-химических свойств в условиях термодинамического равновесия. В противном случае динамическая неоднородность может образовать статическую неоднородность, представляющую собой "замороженный" бит информации.

Третьим элементом модели является генератор динамических неоднородностей, предназначенный для их ввода в канал распространения, расположенный в континуальной среде. Динамическая неоднородность может быть введена в информационный канал в континуальной среде и из-за ее пределов или сгенерирована в этом канале.

Устройство управления динамическими неоднородностями в тракте переноса информационного сигнала или в области его хранения является *четвертым* элементом в модели прибора.

Вывод или считывание информации осуществляется с помощью детектора. Это устройство позволяет преобразовать информационный массив, созданный динамическими неоднородностями, в массив двоичной информации. В этом случае можно использовать хорошо развитые устройства и методы цифровой обработки информации. Детектор является *пятым* элементом типовой модели.

Предложенная модель прибора функциональной электроники позволяет развить системный подход к анализу известных конструктивных решений — прототипов приборов, раскрыть физическую сущность явлений, лежащих в основе работы приборов, оптимизировать известные конструкции по технико-экономическим параметрам, а также разработать *прогнозы* — новое, ранее неизвестное конструктивное решение с заданными технико-экономическими показателями. Такого рода таксонометрические исследования имеют вполне самостоятельное значение как интеллектуальные исследования высокого уровня.

Скажем несколько слов о методическом подходе к изложению этой главы. В нем также исповедуется системный подход, который развит на предложенной физической модели. Материал по всем направлениям функциональной электроники изложен по одинаковой методике.

Контрольные вопросы

1. Что такое функциональная электроника?
2. Что такое динамическая неоднородность?
3. Какие типы динамических неоднородностей вы знаете?
4. Назовите основные элементы модели прибора функциональной электроники.
5. Что представляет собой статическая неоднородность?

Рекомендуемая литература

1. Васенков А. А., Федотов Я. А. Интегральная электроника статических и динамических неодиородностей. — М.: Электронная промышленность, 1, 1983.
2. Васенков А. А., Федотов Я. А. Функциональная электроника. Основные направления работ. — М.: Электронная промышленность, 8, 1983.
3. Федотов Я. А. Электроника твердого тела. В кн. Будущее науки. Международный ежегодник. — М.: Знание, 1975.
4. Щука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

2. Функциональная акустоэлектроника

2.1. Физические основы

Функциональная акустоэлектроника является направлением функциональной электроники, в котором исследуются акустоэлектронные эффекты и явления в различных континуальных средах, а также возможность создания приборов и устройств электронной техники для обработки, передачи и хранения информации с использованием динамических неоднородностей акустической и (или) акустоэлектронной, акустооптической природы.

К акустоэлектронным явлениям и эффектам относятся:

- ☐ генерация, распространение, преобразование и детектирование объемных (ОАВ) и поверхностных акустических волн (ПАВ);
- ☐ преобразование электрического сигнала в акустический и обратно;
- ☐ электронное поглощение и усиление акустических волн;
- ☐ акустоэлектронные и акустомагнитные эффекты;
- ☐ нелинейные акустоэлектронные явления: генерация гармоник, акустоэлектронные домены, параметрическое и супергетеродинное усиление звука;
- ☐ взаимодействие света и звука в твердых телах, дифракция, модуляция и сканирование света звуком.

Исследования этих явлений и эффектов показывают, что с их помощью возможна генерация динамических неоднородностей волновой и доменной природы, которые широко используются в процессорах сигналов и устройствах памяти для обработки и хранения информации.

2.1.1. Динамические неоднородности

В функциональной электронике используются динамические неоднородности акустической, акустоэлектронной или акустооптической природы. В твердом теле могут возбуждаться акустические волны, представляющие собой упругое возмущение вследствие деформации материала. Такие деформации имеют место при движении отдельных атомов и сопряжены с изменением расстояний между ними. При этом возникают внутренние упругие силы, стремящиеся вернуть материал в исходное состояние. Колебания атомов происходят вблизи положения равновесия, и при этом генерируется волна механического напряжения и растяжения. Скорость распространения акустической волны лежит в пределах $(1,5-4,0) \times 10^3$ м/с. Диапазон частот акустических волн находится в интервале от нескольких герц до 10^{13} Гц, а их распространение сопровождается переносом энергии. Различают продольные и сдвиговые волны в зависимости от характера движения частиц (рис. 2.1).

Если размеры звукопровода намного больше длины акустической волны, то в нем могут распространяться объемные акустические волны.

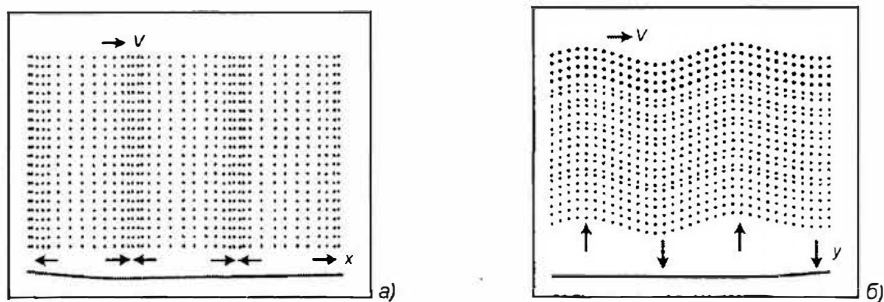


Рис. 2.1. Акустические волны продольного (а) и сдвигового (б) типов

Если же среда, в которой распространяются волны, является ограниченной, то на распространение волн существенное влияние оказывают граничные условия. В однородной среде со свободной плоской поверхностью существуют поверхностные акустические волны (ПАВ). По вектору поляризации волн ПАВ бывают двух типов: для вертикальной поляризации характерно расположение вектора колебательного смещения частицы среды в перпендикулярной граничной плоскости; для горизонтальной поляризации вектор смещения частицы среды параллелен границе и перпендикулярен направлению распространения волн.

ПАВ являются направленными волнами, другими словами, их амплитуда экспоненциально убывает с глубиной. Поэтому $\sim 90\%$ переносимой энергии сосредоточено в слое глубиной не более одной длины волны. Доступность волнового фронта позволяет эффективно управлять распространением ПАВ на всем протяжении звукопровода. В устройствах на ПАВ используются УЗ-волны в диапазоне от 19 МГц до 10 ГГц.

Простейшим типом ПАВ являются волны с вертикальной поляризацией, распространяющиеся вдоль границы твердого тела с вакуумом. Это так называемые *волны Рэлея* (рис. 2.2, а). Энергия рэлеевских волн локализована в приповерхностном слое звукопровода

$$d_p \approx \lambda_v,$$

где λ_v — длина акустической волны. Волны Рэлея не обладают дисперсией, т. е. скорость их распространения не зависит от частоты.

Вдоль границы двух твердых тел могут распространяться *волны Стоуни*, состоящие как бы из двух рэлеевских волн.

К волнам с горизонтальной поляризацией относятся *волны Лява* (рис. 2.2, б). Это сдвиговые волны, существующие в тонком слое на поверхности твердого тела. Волны Лява обладают дисперсией и локализуются в слое, толщина которого $d_p \geq \lambda_v$.

Исследованы чисто сдвиговые волны, не обладающие дисперсией (рис. 2.2, в). Эти волны получили название *волн Гуляева — Блюштейна*. Они могут существовать на свободной поверхности пьезоэлектрических кристаллов. Глубина их проникновения в десятки раз превышает длину акустической волны ($d_p \gg \lambda_v$). Применение таких волн позволяет избежать тщательной обработки поверхности материала. Важной особенностью распространения динамических неоднородностей акустической природы (ПАВ) в континуальных средах с различными физическими свойствами является существующий эффект генерации динамических неоднородностей другой природы. Вследствие явления акустоэлектронного взаимодействия происходит воздействие акустической волны на электроны

проводимости в твердых телах. Результатом такого воздействия является обмен энергией и импульсом между акустической волной и электронами проводимости. Например, передача энергии акустической волны электронам приводит к электронному поглощению звука, а передача импульса акустической волны стимулирует возникновение электрического тока. Возможно явление усиления звука за счет стимулированного дрейфа электронов в твердом теле и частичной передачи энергии акустической волне.

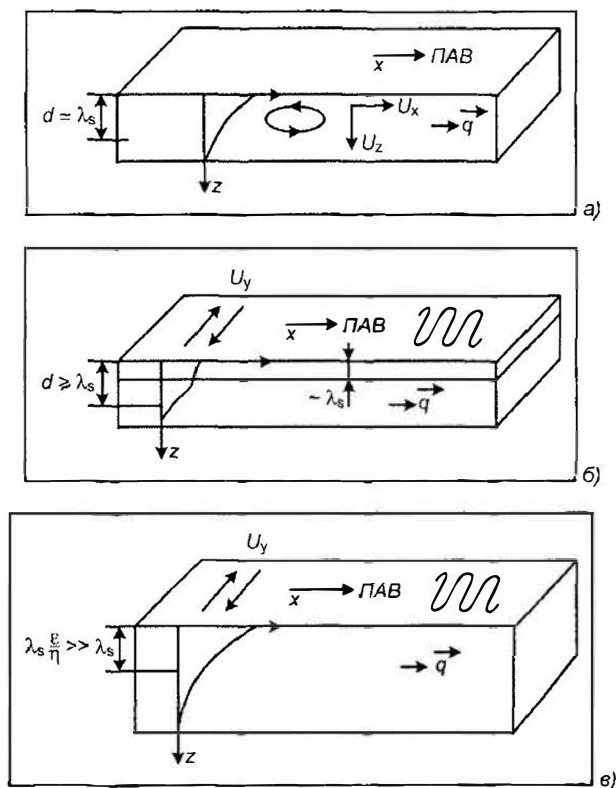


Рис. 2.2. Поверхностные акустические волны различных типов: а — волны Рэлея, б — волны Лява, в — волны Гуляева — Блюкштейна

Возникающая при распространении акустической волны деформация вызывает в пьезоматериалах переменное электрическое поле, амплитуда и фаза которого находятся в прямой зависимости от объемного заряда электронов проводимости (прямой пьезоэффект). В свою очередь это поле вызывает деформацию кристалла и соответственно изменение характера распространения волны (обратный пьезоэффект). Акустическая волна генерирует волны электрических полей. В местах, где кристалл сжимается волной, наведенное электрическое поле замедляет движение электронов, а в местах растягивания кристалла волной наблюдается ускорение электронов за счет внутреннего электрического поля. Под действием этих полей носители стремятся сгруппироваться в областях с минимумом потенциальной энергии. Возникают затухающие волны объемного заряда, несколько запаздывающие по отношению к акустической волне (рис. 2.3, а). Волны объемного заряда

представляют собой динамические неоднородности электрической природы, стимулированные ПАВ.

Если звукопровод поместить в постоянное электрическое поле E_0 , то возникает дрейф электронов со скоростью

$$V_{др} = \mu E_0,$$

где μ — подвижность электронов. В случае, когда $V_{др} > V_{фаз}$ ($V_{фаз}$ — фазовая скорость распространения акустической волны), электроны отдают свою энергию ПАВ и амплитуда ПАВ возрастает (рис. 2.3, б). В случае затухания волны график функции достаточно точно описывается функцией типа

$$U = \exp(-ax) \cdot \sin(bx),$$

а в случае ее усиления

$$U = \exp(ax) \cdot \sin(bx),$$

где a и b — константы. Обмен энергией между динамическими неоднородностями акустической и электронной природы в пределах одной континуальной среды является фундаментальным эффектом, лежащим в основе приборов акустоэлектроники.

В акустооптических средах акустические и электромагнитные волны параметрически связаны упругооптическим эффектом. Упругая акустическая волна индуцирует изменение показателя преломления. Это позволяет сформировать динамические неоднородности в виде оптических неоднородностей по показателю преломления (фазовые решетки), распространяющихся вслед за акустической волной. Свет дифрагирует на таких регулярных динамических неоднородностях оптической природы.

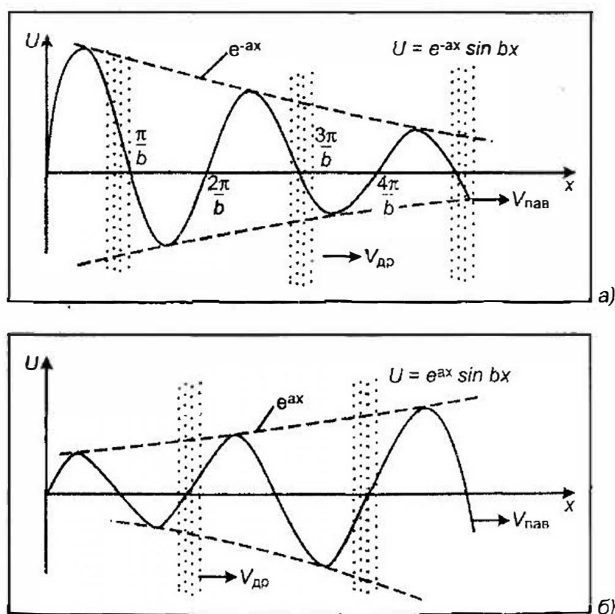


Рис. 2.3. Схема взаимодействия динамических неоднородностей акустической и электрической природы: а — при отсутствии потенциала; б — при наличии потенциала на звукопроводе

Акустическая волна в пьезомагнитной среде (антиферромагнетиках) порождает магнитоупругие волны, которые являются еще одним типом динамических неоднородностей в акустоэлектронике.

2.1.2. Континуальные среды

Континуальными средами функциональной акустоэлектроники являются твердотельные материалы: пьезоэлектрики, пьезополупроводники, сложные слоистые среды. Выбор континуальных сред определяется природой используемых динамических неоднородностей. Основные требования к этим материалам сводятся к минимизации уровня потерь для распространения динамических неоднородностей акустической или иной природы, а также к максимальной температурной стабильности. Другие требования к среде диктуются функциональным назначением прибора, технологией его производства.

Распространение волны в твердом теле сопряжено с локальной деформацией среды. Для небольших деформаций справедлив закон, связывающий смещение частиц y_i и упругие напряжения:

$$\rho d^2 y_i / dt^2 = dT_{ij} / dx_j, \quad (2.1)$$

где T_{ij} — симметричный тензор механических напряжений второго ранга $T_{ij} = T_{ji}$.

В пьезоэлектрических кристаллах тензор механических напряжений и вектор индукции можно описать уравнениями состояния

$$T_{ij} = c_{ijkl} \sigma_{kl} + 1 / 2 c_{ijklmn} \sigma_{kl} \sigma_{mn} - p_{kij} E_k - 1 / 2 Q_{klij} E_k E_l - p_{kijlm} E_k \sigma_{lm} \quad (2.2)$$

$$D_i = \epsilon_{ij} E_j + 1 / 2 \epsilon_{ijkl} E_j E_k + p_{ijk} \sigma_{jk} + Q_{ijkl} E_j \sigma_{kl} + 1 / 2 p_{ijklm} \sigma_{jk} \sigma_{lm}, \quad (2.3)$$

где $\sigma_{ij} = 1 / 2 [dy_i / dx_j + dy_j / dx_i + (dy_k / dx_i dy_l / dx_j)]$ — тензор деформаций, в котором учтен нелинейный характер связи компонент тензора механических деформаций с производными от смещения; p_{ijk}, p_{ijklm} — компоненты тензоров линейного и нелинейного пьезоэффектов; $\epsilon_{ij}, \epsilon_{ijkl}$ — компоненты тензоров линейной и нелинейной диэлектрической проницаемости; Q_{ijkl} — компонента тензора электрострикции.

Тензор T_{ij} имеет четвертый ранг и содержит $3^4 = 81$ элемент. Симметрия тензоров требует, чтобы компоненты были инвариантными относительно индексов i и j или k и l .

Все эти сложные математические соотношения призваны пояснить простую идею: динамические неоднородности акустической природы определенными соотношениями связаны с динамическими неоднородностями электрической природы, и наоборот.

В уравнениях (2.2) и (2.3) первые слагаемые соответствуют линейным составляющим, а остальные члены соответствуют нелинейным составляющим механического напряжения и индукции.

Пьезоэлектрические материалы характеризуются рядом и других коэффициентов и параметров, без учета которых невозможно выбрать континуальную среду для приборов и устройств функциональной акустоэлектроники.

Коэффициент электромеханической связи k_m^2 определяется величиной

$$k_m^2 = \frac{W_e}{W_\Sigma},$$

где W_Σ — генерируемая пьезоэлектриком электрическая энергия, $W_\Sigma = W_u + W_e$ — полная энергия, равная сумме механической энергии деформации W_u и электрической W_e .

Существует точка Кюри T_k , в которой отсутствует спонтанная поляризованность P_s , пьезоэлектрика, а величина пьезомодуля определяется соотношением

$$d \approx P_s \epsilon_0 (\epsilon - 1) = 0.$$

Добротность среды Q обеспечивает частотную избирательность изделий.

В настоящее время известно более 1500 веществ, обладающих пьезоэлектрическими свойствами. Охарактеризуем только некоторые из них.

Классическими материалами являются пьезоэлектрические монокристаллы. Наиболее широко известный кристалл — кварц, представляющий собой кристаллическую модификацию безводной двуокиси кремния SiO_2 . Наряду с природными кристаллами используются и синтетические, превосходящие природные по однородности, размерам, а также по стоимости изготовленных из них изделий. Определенные срезы кварца характеризуются своими параметрами и используются в различных устройствах. Кристаллы ниобата лития (LiNbO_3) и танталата лития (LiTaO_3) обладают более высоким, чем кварц пьезомодулями и коэффициентами электромеханической связи. Во многих областях приборостроения эти кристаллы вытеснили кварц, особенно в СВЧ-технике.

Пьезополупроводники (CdS , ZnS , ZnO) используются для пленочных преобразователей электромагнитных колебаний вплоть до СВЧ-диапазона. Технология их получения довольно хорошо отработана, и можно изготовить пленки с заданной симметрией.

Определенный интерес представляют такие пьезополупроводники, как селениды металлов (CdSe), арсенид галлия (GaAs), антимонид индия (InSb), а также кристаллы иодата лития (LiIO_3) и калия (KJO_3), германата висмута ($\text{Bi}_{12}\text{GeO}_{20}$).

Широкое распространение получили промышленные пьезокерамические материалы, как правило, представляющие собой твердые растворы. Свойства таких материалов задаются путем подбора соотношений компонентов, введением модифицированных добавок, а также технологией их изготовления.

К пьезокерамическим материалам относятся титанат бария (BaTiO_3), титанат свинца (PbTiO_3), ЦТС [$\text{Pb}(\text{Zr}_{0,53}\text{Ti}_{0,47})\text{O}_3$] и др. Добавки, вводимые, например, в ЦТС, расширяют двухфазную область составов и усиливают нужные свойства пьезокерамики. Модификацию пьезокерамики можно осуществить в широком диапазоне свойств за счет введения добавок со скомпенсированной валентностью, другими словами, набора оксидов соответствующего перовскитному соединению.

Следует особо подчеркнуть, что поверхность используемых в акустоэлектронике континуальных сред должна быть тщательно обработана. Это диктуется тем, что звукопровод расположен в приповерхностном слое и тщательная полировка поверхности позволит снизить рассеяние звука и избежать помех.

Анизотропность кристаллов приводит к резкой зависимости скорости распространения ПАВ от направления среза. Поэтому срезы необходимо выбирать так, чтобы добиться выполнения условия коллинеарности векторов фазовой и групповой скоростей.

Континуальные среды могут проявлять нелинейные свойства. Такая ситуация реализуется при достаточно больших смещениях частиц в акустических волнах. В этом случае закон Гука имеет дополнительные компоненты тензора нелинейной упругости. Главным проявлением нелинейности среды является нарушение принципа суперпозиции волн и, как следствие, появление комбинационных частот колебаний, например $(\omega_1 \pm \omega_2)$.

Нелинейность пьезоэлектрических сред порождает еще один эффект. Он связан с тем, что акустическая волна сопровождается волнами зарядовой плотности (электрического поля)

и при достаточно больших амплитудах электрических полей тоже возникают нелинейные явления. Этот эффект получил название *нелинейный пьезоэффект*. В этом случае в уравнении для T_{ij} компонент электрического поля не линеен. Этот эффект сопровождается диэлектрической нелинейностью, связанной с нелинейностью электрической индукцией D_i относительно компонентов электрического поля E_j .

Для получения континуальных сред с заметной нелинейностью подбирают соответствующие параметры, определяемые тензорами высших порядков.

2.1.3. Генераторы динамических неоднородностей

Возбуждение динамических неоднородностей в виде акустической волны в пьезоэлектрике осуществляется с помощью вложенных друг в друга групп электродов, которые получили название *встречно-штыревых преобразователей* (ВШП).

Другое название такой системы электродов — двухфазный преобразователь ПАВ. Две группы чередующихся электродов соединены шинами, которые в свою очередь подключены к источнику напряжения (рис. 2.4). При подаче на ВШП напряжения каждая пара электродов возбуждает ПАВ. Если период преобразователя $(a + b)$ равен длине ПАВ, возникает явление акустического синхронизма. В этом случае волны, возбуждаемые каждой парой электродов, имеют одинаковую фазу, и происходит когерентное сложение волн.

Суммирование происходит за счет локальных деформаций, образующихся под промежутками между электродами. Образовавшаяся деформация начинает перемещаться в оба направления и достигает следующих промежутков как раз в тот момент, когда полуволна внешнего напряжения достигнет максимума и вызовет деформацию под своими электродами.

Сложение деформаций происходит при выполнении условия

$$\lambda = 2(a + b) = V_a / f,$$

где λ , V_a , f — соответственно длина, скорость и частота звуковой волны. Такая картина происходит под каждой парой ВШП. Чем больше штырей содержит преобразователь, тем эффективней происходит преобразование электрической энергии в механическую, и наоборот.

Целесообразно ввести топологические и технологические параметры структуры.

Коэффициент металлизации определим как

$$H = \frac{a}{a + b}.$$

При $a \approx b$ $H \approx 0,5$.

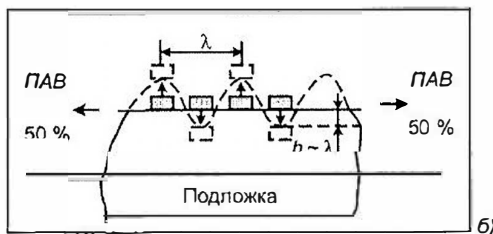
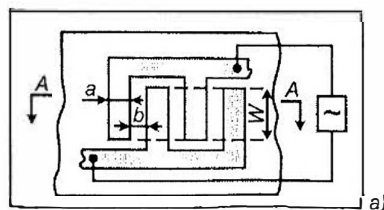


Рис. 2.4. Возбуждение ПАВ с помощью ВШП: а — топология ВШП; б — схема возникновения акустических волн (сечение АА)

Частота синхронизации или центральная (рабочая) частота определяется выражением:

$$f_a = \frac{V_a}{\lambda} = \frac{V_a}{2(a+b)} = \frac{V_a}{d}, \quad (2.4)$$

где $d = 2(a+b)$ — период структуры электродов.

Большое значение имеет топологический параметр W или апертура, определяющая степень перекрытия ВШП. Обычно выбирают значение апертуры $W \sim 100\lambda$. Апертура определяет форму импульсного отклика ВШП. Конструктивно ВШП выполняются в виде тонкопленочных металлических электродов (Au, Al). Толщина электродов h определяется технологией и обычно $h \leq a$.

Коэффициент электромеханической связи k_m^2 является технологическим параметром.

Генерацией динамических неоднородностей можно эффективно управлять путем изменения характеристик преобразователей ПАВ. Исследовано несколько способов управления генерацией ПАВ.

Первый способ предусматривает управление генерацией ПАВ путем изменения топологии ВШП, включая изменение числа штырей, их расстановку, варьирование функции аподизации или изменения длины отдельных электродов (штырей) или их групп, изменение полярности отдельных штырей или их групп. Устройство с управляемой топологией электродов ВШП предусматривает изменение топологии путем перекоммутации отдельных электродов или их групп относительно общих шин преобразователя. Устройства такого типа обладают достаточно широким диапазоном перестройки характеристик в частотной и временной областях, их параметры приближаются к предельно достижимым. Технологические приемы изменения топологии сводятся к проекционным методам, либо навешиванию дискретных элементов и проволочных межсоединений.

Другой способ управления генерацией ПАВ связан с операцией "взвешивания" электродов преобразователей, которая реализуется за счет изменения количества активных пар ВШП. Это позволяет менять ширину полосы генерации ПАВ и, следовательно, изменять характеристики динамических неоднородностей. На рис. 2.5, а представлена конструкция дискретно-управляемой ВШП, позволяющая получить набор амплитудно-частотных характеристик генерируемых ПАВ с различной шириной пропускания, но одинаковой для всех центральной частотой (рис. 2.5, б). Преобразователь разбит на секции a, b, c с числом штырей 1, 2 и 4, соответственно. Количество каналов K определяется числом секций n и равно

$$K = 2^n - 1.$$

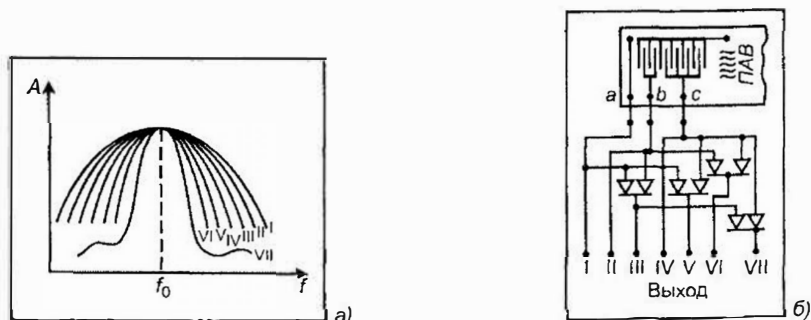


Рис. 2.5. Генерация ПАВ с различными АЧХ (а) и схема коммутации пар ВШП (б)

В схеме коммутации предусмотрены усилители, и обеспечивается контроль выходных параметров за счет применения электронных переключающих схем.

Устройство с управляемым "взвешиванием" электродов представляет собой целенаправленное изменение эффективности преобразования энергии, осуществляемого каждой парой электродов или их группой путем подключения к электродам управляемых импедансных элементов. Это могут быть резистивные, емкостные или индуктивные элементы, а также их комбинации.

Третий способ опирается на управление электрофизическими свойствами подложки под электродами. Мгновенная фаза ПАВ, генерируемая каждой парой ВШП, зависит от знака пьезоэлектрического коэффициента. В свою очередь знак пьезоэлектрического коэффициента может быть изменен переполаризацией в полном соответствии с гистерезисными свойствами материала. Такой прием позволяет управлять процессами генерации ПАВ, создать адаптивные устройства с "памятью". На рис. 2.6, а приведена одна из таких конструкций. Рядом с электродами ВШП 1 размещены электроды переключения поляризации 2, подключенные к источнику поляризующего напряжения $E^- < U_n < E^+$. Это позволяет менять знак поляризации материала звукопровода между металлическим электродом 3 и электродами 2 одновременно или путем перекоммутации локально под каждой парой ВШП. Варьируя значениями U_n и U_{ax} , можно эффективно управлять генерацией ПАВ.

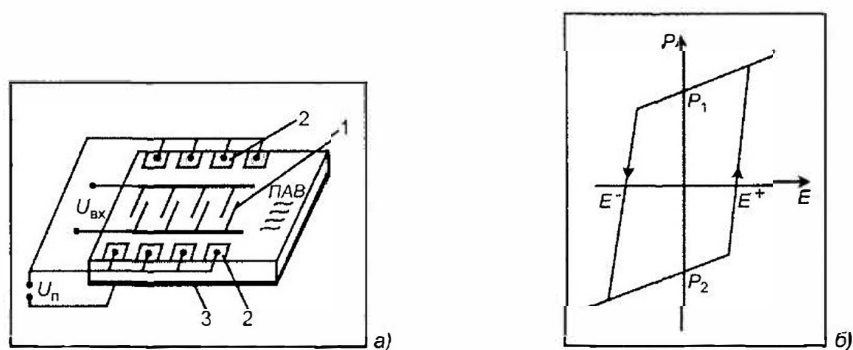


Рис. 2.6. Конструкция адаптивного генератора ПАВ (а) и гистерезис пьезоэлектрического звукопровода (б)

Двухфазные преобразователи позволяют генерировать колебания в обе стороны звукопровода "симметрично", другими словами, 50% энергии направо и 50% энергии налево. Однако можно управлять и направлением генерации ПАВ. С этой целью целесообразно использовать многофазные преобразователи, позволяющие синфазно подпитывать волну в заданном направлении (рис. 2.7, а). Фазовое распределение подбирается так, чтобы в обратном направлении условие синфазности не выполнялось. Однако полоса пропускания трехфазного преобразователя определяется протяженностью ВШП и ограничена. Для расширения полосы пропускания и достижения однонаправленности генерации предложен преобразователь бегущей волны (рис. 2.7, б). Каждый из электродов ВШП запитывается от одного из отводов электрической линии задержки. Линия задержки рассчитывается так, что время задержки между соседними отводами равно времени распространения ПАВ между соседними электродами. Это резко несимметричный генератор ПАВ, обеспечивающий строго однонаправленную генерацию.

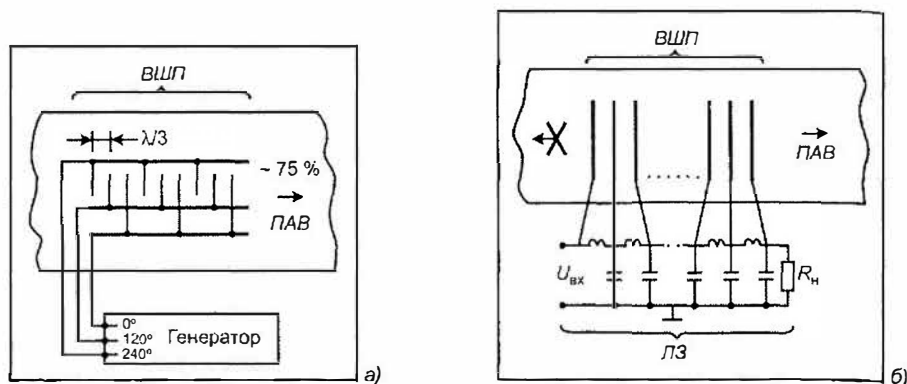


Рис. 2.7. Трехфазный (а) и многофазный (б) генераторы ПАВ

Эффективное управление генерацией осуществляется путем подбора необходимых топологических и технологических параметров.

2.1.4. Устройство управления динамическими неоднородностями

Исследовано несколько способов управления прохождением динамических неоднородностей по тракту передачи информации. Эти процессы определяются типом динамических неоднородностей, функциональным назначением прибора или устройства. Так, если используется только ПАВ, распространение которой происходит в тонком приповерхностном слое, то для эффективного управления можно использовать интерференцию, дифракцию, отражение, преломление, переизлучение, фокусирование волн. Можно варьировать также параметры звукопровода, изменяя такие характеристики процесса распределения, как скорость ПАВ, дисперсию, удельное затухание и т. п.

Наиболее распространенным методом является управление акустическим трактом путем изменения топологии его элементов. Эти элементы позволяют сформировать каналы распространения ПАВ. К таким элементам относится *многополосковый ответвитель* (МПО), который конструктивно выполняется в виде системы пленочных металлических электродов, нанесенных на подложку. В зависимости от назначения устройства управления топология МПО имеет различную форму и способна переизлучать энергию в любой из каналов, формировать эти каналы.

При выполнении условия синхронизма $\lambda = d$ МПО ведет себя как отражательная структура, и рабочий диапазон выбирается в пределах $0,4 \div 0,9$ центральной частоты f_0 . Число электродов, необходимых для эффективной перекачки энергии из канала, определяется соотношением:

$$N = \frac{\lambda}{k_m^2 d},$$

а общая длина МПО равна

$$L = \frac{\lambda}{k_m^2}.$$

На рис. 2.8 приведены некоторые примеры использования МПО в звуковом канале. С помощью МПО можно переизлучать энергию из одного канала в другой (рис. 2.8, а), раздвигать акустический канал (рис. 2.8, б), разворачивать волновой фронт в обратном направлении (рис. 2.8, в).

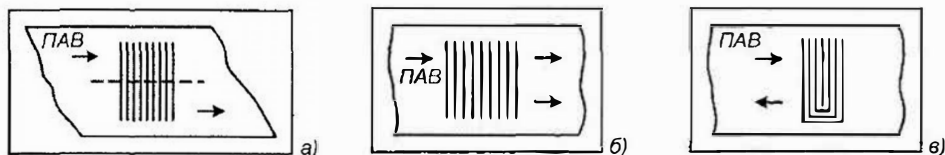


Рис. 2.8. Некоторые операции в тракте, осуществляемые МПО:

а — акустическая связь между двумя независимыми звукопроводами; б — раздвоение канала;
в — обращение акустической волны

В качестве элементов акустического тракта можно использовать периодические неоднородности на поверхности звукопровода в виде пазов, выступов металлических или диэлектрических полосок, а также комбинации этих структур.

Разработаны методы локализации (канализирования) волны в топографическом волноводе (рис. 2.9, а, б). Часто это связано с необходимостью увеличения протяженности тракта, чтобы обеспечить максимальное время задержки. Помимо топографических волноводов для локализации акустической волны можно применять и плоские слоистые волноводы (рис. 2.9, в).

Слоистые волноводы изготавливают путем нанесения вещества, скорость ПАВ в которых отлична от скорости в звукопроводе. Степень локализации волны зависит от соотношения скорости ПАВ в звукопроводе и в слое нанесенного вещества.



Рис. 2.9. Акустические волноводы: топографические волноводы треугольной (а) и прямоугольной (б) формы, слоистый волновод с нанесенным звукопроводом (в) и щелевой волновод (з)

Разработаны конструкции МПО с динамически управляемой конфигурацией электродов, в которых под воздействием локального излучения создаются области повышенной проводимости.

Эффективное управление прохождением динамических неоднородностей является изменением граничных условий распространения ПАВ. К граничным условиям будем относить большое число физических параметров, характеризующих среду распространения волны вдоль границы твердого тела, и параметры, отражающие структуру звукопровода. Этот метод управления базируется на локальном изменении свойств среды, что весьма эффективно с энергетической точки зрения. Например, если покрыть поверхность звукопровода тонким слоем селенида кадмия, удельное сопротивление которого зависит от уровня ос-

вещности, то можно менять мнимую часть акустического импеданса. В этом случае появляется возможность изменять амплитуду ПАВ в достаточно широком диапазоне значений.

Однако заметим, что методу управления удельной проводимостью поверхностного слоя присущи недостатки, связанные с большим энергопотреблением, громоздкостью проекционных систем.

Если же использовать магниточувствительные пленки, нанесенные на поверхность звукопровода, то изменять упругие свойства пленки, а также акустический импеданс звукопровода становится проще. Возникающие магнитоупругие поверхностные волны обладают рядом специфических свойств. В частности, их разовая скорость зависит от ориентации вектора управляющего магнитного поля, что позволяет эффективно управлять скоростью распространения в пределах 20%.

Третий метод управления свойствами звукопровода основывается на целенаправленном изменении электрофизических свойств материала звукопровода. Возможность управления свойствами материала может быть реализована за счет термодинамической взаимосвязи тепловых, электрических, магнитных и упругих параметров среды. Управление в этом случае осуществляется с помощью полей различной физической природы.

В некоторых конструкциях приборов используется линейная зависимость между деформацией звукопровода и скоростью распространения ПАВ. Изменение скорости распространения ПАВ также линейно зависит от температуры и аналитически записывается в виде:

$$\Delta l/l = \frac{\Delta V_a}{V_a} - \beta \Delta t,$$

где β — температурный коэффициент расширения, Δt — температурный интервал. Этот метод отличается простотой, отсутствием дополнительных потерь, позволяет эффективно использовать материалы с большим коэффициентом электромеханической связи.

Управление скоростью распространения ПАВ можно осуществлять, используя эффект электроупругого взаимодействия. С этой целью звукопровод помещается в электрическое поле с напряженностью $\sim 10^3$ В/см. Однако использование высоковольтного источника напряжения сопряжено с известными трудностями.

Скоростью распространения ПАВ можно также управлять, используя термоупругий эффект. Его применение позволяет изменять упругие константы материала в тепловых полях. Однако большая инерционность тепловых процессов ограничивает использование этого метода управления.

2.1.5. Детектирование динамических неоднородностей

Детектирование динамических неоднородностей является, как правило, физическим процессом, обратным их генерации. Если при генерации ПАВ используется прямой пьезоэффект, позволяющий преобразовывать энергию электрического поля в энергию акустической волны, то в процессе детектирования используется обратный пьезоэффект. Устройство, позволяющее детектировать ПАВ, аналогично генератору ПАВ и представляет собой ВШП.

Одним из способов детектирования является управление топологией электродов. С этой целью можно менять их геометрическую конфигурацию, перескоммутировать отдельные электроды или их группы, формировать латентные электроды и управлять локальным облучением световым или электронным потоком.

Форма импульсного отклика ВШП зависит от закона изменения перекрытия электродов, другими словами, от их частоты и апертюры. На рис. 2.10 представлены формы выходных сигналов, зависящие от топологии детектора, при подаче на вход единичного импульса.

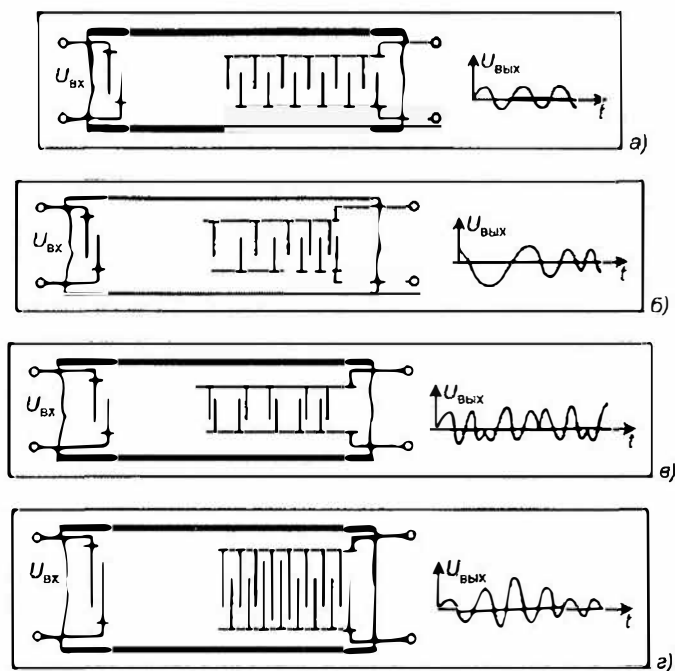


Рис. 2.10. Управление генерацией импульсов топологией ВШП: а — эквидистантный, неаподизированный; б, в — неэквидистантный; неаподизированный; г — эквидистантный, аподизированный по закону $\sin x / x$

Если произвести перекоммутацию штырей детектора, то можно управлять выходным сигналом, аналогично уже рассмотренному случаю (рис. 2.5). Детектирование ПАВ можно осуществлять также методом управляемого взвешивания, осуществляемое путем подключения к электродам преобразователя управляемых импедансных элементов любого типа. Такой прием обеспечивает регулируемую амплитудную модуляцию импульсного отклика, а также управление формой амплитудно-частотной и фазово-частотной характеристик.

И наконец, детектирование можно осуществить, управляя электрофизическими свойствами подложки. В этом случае конструкции выходного ВШП аналогичны входному, а физические процессы детектирования дополнены (обратны) процессам генерации.

2.2. Приборы функциональной акустоэлектроники

Приборы функциональной акустоэлектроники предназначены для преобразования, аналоговой обработки и хранения информации.

С помощью акустоэлектронных приборов и устройств можно производить различные операции с сигналами: задерживать по времени, изменять длительность и форму сигналов

по амплитуде, частоте и фазе, преобразовывать частоты и спектр сигналов, сдвигать их по фазе, производить модулирование сигналов, производить кодирование и декодирование сигналов, интегрирование сигналов, получать функции свертки и корреляции сигналов. Устройства, производящие эти операции, отнесены к процессорам сигналов.

Акустоэлектронные устройства позволяют также осуществлять хранение информации. Такие устройства отнесем к запоминающим устройствам.

Схема, отображающая классификацию приборов функциональной акустоэлектроники, приведена на рис. 2.11. В этих устройствах используются акустические волны в диапазоне 5×10^6 — 2×10^9 Гц. Использование приборов и устройств функциональной акустоэлектроники позволяет осуществлять процесс обработки аналоговой и цифровой информации достаточно просто и надежно. При этом удастся получить выигрыш в габаритах, массе, энергии и стоимости, что характерно для приборов микроэлектроники.

2.2.1. Линии задержки

Простейшим процессором сигналов является *линия задержки* (ЛЗ), которая предназначена для временной задержки сигналов без заметных их искажений.



Рис. 2.11. Классификация устройств функциональной акустоэлектроники

Основным преимуществом ЛЗ на ПАВ являются их небольшие габариты, достаточно широкий диапазон частот (до 10^9 Гц), хорошая температурная стабильность ($\sim 10^{-6} \text{ } ^\circ\text{C}^{-1}$). Эти свойства обусловлены, прежде всего, особенностями ПАВ, а именно невысокой ($\sim 10^5 \text{ см/с}$) скоростью распространения, бездисперсионностью, эффективным преобразованием электрической энергии в акустическую и наоборот. Классификация линий задержки приведена на схеме, представленной на рис. 2.11.

Линии с однократной задержкой сигнала предназначены для однократного или единичного съема информационного сигнала и должны обеспечивать заданную задержку и форму АЧХ-сигнала с максимальной точностью в пределах широкого диапазона температур. Конструктивное решение линии задержки определяется требуемым временем задержки сигнала, а полоса пропускания выбранной топологией ВЧП (рис. 2.12, а). Исходя из

обобщенной модели изделия функциональной электроники, заметим, что в качестве континуальной среды в ЛЗ чаше используются кварц Y -среза (ST), ниобат лития, германат висмута. Выбор конкретного материала осуществляется на основе энергетического критерия, позволяющего обеспечить минимальные потери, или на основе критерия температурной стабильности параметров.

Генератором динамических неоднородностей в виде ПАВ служит ВШП, как правило, неаполитизированный и эквидистантный. Расчет ВШП ПАВ производится исходя из принципа оптимального энергетического согласования в полосе частот. Проще было бы согласование произвести на центральной частоте — частоте акустического синхронизма. Поэтому необходимо обеспечить в акустическом тракте звукопровода равенство электрической и акустической добротностей.

Электрическая добротность Q_e определяется коэффициентом электромеханической связи и числом пар штырей ВШП N :

$$Q_e \approx \frac{\pi}{8k_m^2 N}.$$

Акустическая добротность Q_a тракта определяется как

$$Q_a = \frac{f_0}{\Delta f}.$$

Число электродов ВШП N выбирается исходя из соотношения

$$N = \frac{2f_0}{\Delta f} \alpha,$$

где α — коэффициент, учитывающий уменьшение полосы пропускания, соответствующей произведению амплитудно-частотных характеристик (АЧХ) входного и выходного ВШП. Тогда при $\alpha \approx 0,6 + 0,8$ величина

$$Q_a \approx \frac{N}{1,6}.$$

Зависимости $Q_e = f(N)$ и $Q_a = f(N)$ приведены на рис. 2.12, б.

Можно выделить три области на графиках. При $N < N_{opt}$ на частоте акустического синхронизма при $R_s = R_n$ входная и выходная цепи будут согласованы. При этом 50% энергии источника преобразуется в ПАВ. Однако в этой области $Q_e > Q_a$ и полоса частот будет урезана. Для увеличения полосы частот акустической цепи необходимо зашунтировать электрический контур, например, увеличив сопротивление генератора R_s . Оптимальным режимом является $Q_e = Q_a$, выполняемое при оптимальном числе пар штырей ВШП

$$N_{opt} = \sqrt{\frac{0,6}{k_m^2}}.$$

При $N > N_{opt}$ ширина полосы пропускания электрической цепи будет больше, чем у акустической. С точки зрения минимизации вносимых потерь такой преобразователь предпочтительней малоэлектродного ($N < N_{opt}$). Вводится величина P , характеризующая степень рассогласования:

$$P = \frac{Q_e(N)}{Q_a(N)} = \left(\frac{N_{opt}}{N} \right)^2. \quad (2.5)$$

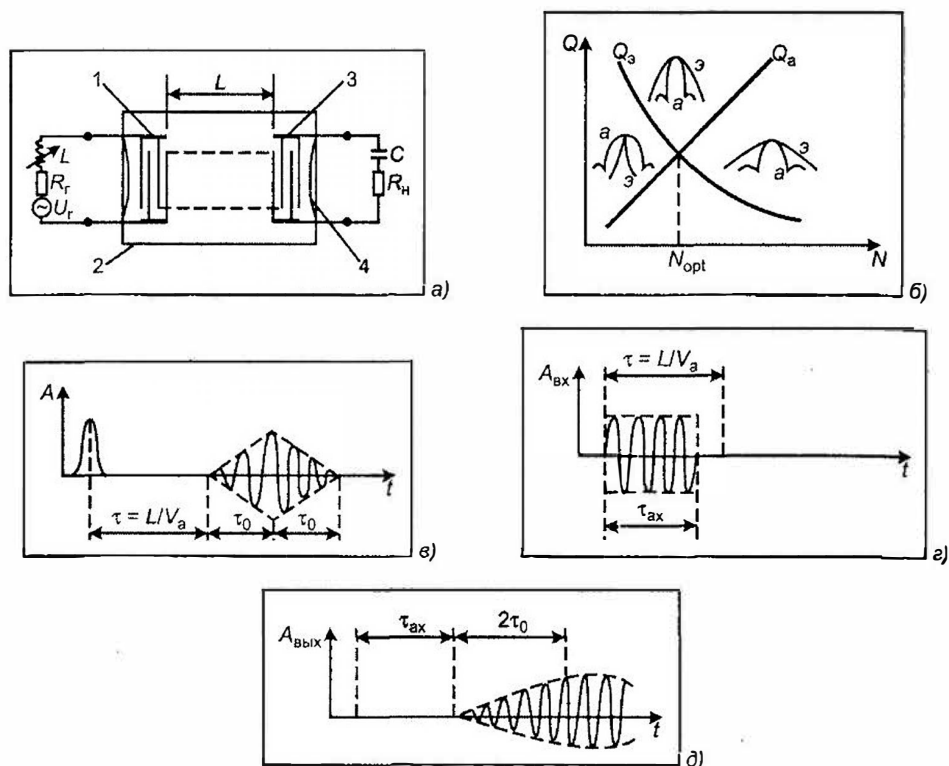


Рис. 2.12. Линия с однократной задержкой сигнала: а — общий вид; 1 — входной ВШП; 2 — непрерывная среда; 3 — выходной ВШП; 4 — поглотители ПАВ; L , R_r , r — индуктивность, сопротивление и напряжение генератора входной цепи, соответственно; C , R_n — емкость и сопротивление нагрузки выходной цепи; б — зависимость акустической и электрической добротности от числа электродов ВШП; в — импульсный отклик на сигнал $\delta(\tau)$; г — форма радиоимпульса; д — отклик ЛЗ на радиоимпульс

Шаг электродов d выбирается исходя из заданной центральной частоты

$$d = V_a / 2f_0,$$

а ширина электродов a , как правило, равна расстоянию между ними b ($a = b$) и тогда

$$a = V_a / 4f_0.$$

Минимальное значение апертуры ВШП определяется дифракционными потерями и может быть оценено из соотношения:

$$W_{\min} \geq \sqrt{V_a \lambda \tau}.$$

Время задержки сигнала

$$\tau = L / V_a,$$

где L — длина звукопровода. Максимальное значение апертуры ограничивается конечной шириной звукопровода.

Переотражение акустической волны между преобразователями вызывает ложный сигнал тройного прохождения. Его задержка по отношению ко времени входного сигнала со-

ставляет величину 3τ . Этот сигнал искажает АЧХ. При подаче на вход ЛЗ одиночного импульса $\sigma(t)$, обладающего единичной амплитудой и бесконечно малой длительностью, на выходе возникнет акустический сигнал, задержанный на время

$$\tau = L / V_a$$

Если входной и выходной ВШП одинаковы, то отклик имеет огибающую в форме ромба с общей длительностью, равной сумме длительностей переднего и заднего фронтов. Огибающая заполнена колебаниями частоты f_0 . Если же на вход ЛЗ подать радиоимпульс, спектр входного сигнала которого уже полосы пропускания ЛЗ (рис. 2.12, з), то на выходе появится сигнал с фронтами $2\tau_0$ (рис. 2.12, д). При этом следует иметь в виду, что входной и выходной преобразователи идентичны, а частота заполнения радиоимпульса равна частоте акустического синхронизма ВШП. Задержка в этом случае также определяется соотношением $\tau = L / V_a$. На точность воспроизведения задержки влияет, прежде всего, точность изготовления фотошаблонов, ошибки при совмещении и экспонировании, точность ориентации звукопровода. Сказывается и ограниченная термостабильность скорости распространения ПАВ. Линии задержки на ПАВ отличает широкий динамический диапазон D [дБ], определяемый выражением:

$$D = 20 \cdot B \cdot \lg U_{\text{вх, max}} / U_{\text{вых, min}}$$

где $U_{\text{вх, max}}$ — максимально возможное входное напряжение; $U_{\text{вых, min}}$ — минимальное напряжение на выходе ЛЗ, обеспечивающее выделение его на фоне шумов, B — вносимые ЛЗ потери, измеряемые в децибелах. Величина динамического диапазона составляет $80 \div 120$ дБ.

Многоотводные линии задержки (МЛЗ) предназначены для увеличения максимального времени задержки, увеличения числа дискретных диапазонов задержки, регулировки времени задержки. Максимальную задержку сигналов можно получить, эффективно управляя распространением ПАВ. Увеличивая траекторию распространения ПАВ в пределах одного звукопровода, можно получить заданные параметры. На рис. 2.13 приведены некоторые конструктивно-технологические решения, связанные с увеличением звукового тракта.

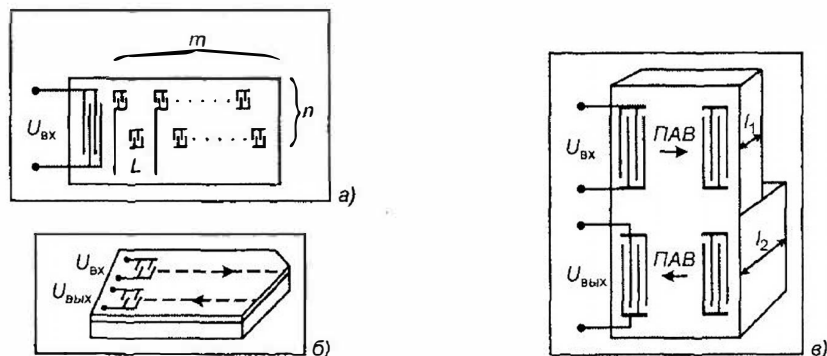


Рис. 2.13. Многоотводная линия задержки: а — матричная конструкция; б — пленочный звукопровод; в — каскадное включение парциальных ЛЗ

Матричная конструкция МЛЗ имеет входной ВШП и матрицу выходных, располагающихся по m в n независимых каналах (рис. 2.13, а). Шаг в одном составляет L , дискрет

задержки $t = L / n$. Другие варианты, например, могут быть связаны с созданием звукового тракта в виде ломаной линии с использованием пьезоэлектрических пленок в качестве континуальной среды (рис. 2.13, б). Рассмотренные конструкции не обеспечивают подавление трехзачодного сигнала и объемных волн, возникающих между ВШП и тыльной стороной звукопровода. Конструкция каскадного включения парциальных ЛЗ позволяет избежать этих помех (рис. 2.13, в). Парциальные ЛЗ выполнены на подложках различной толщины l_1 и l_2 так, что АЧХ объемно-волновых трактов не совпадают и отсутствует результирующая частота их взаимодействия.

Процесс производства МЛЗ заключается в поиске конструктивных и электрических характеристик звеньев линии задержки.

Дисперсионные линии задержки (ДЛЗ) предназначены для формирования зависимости задержки от частоты сигнала.

В ДЛЗ используется генерирующий ВШП с неэквидистантным расположением штырей. Шаг электродов меняется от

$$d_{\max} = V_a / 2f_{\min}$$

до

$$d_{\min} = V_a / 2f_{\max}$$

где f_{\min} и f_{\max} — нижняя и верхняя частоты девиации ДЛЗ. В отличие от физической дисперсии используется "топологическая" дисперсия. На рис. 2.14, а представлена одна из конструкций ДЛЗ. Генерирующий ВШП несимметричен и неаподизирован, а детектирующий имеет малое число штырей и рассчитан на широкую полосу АЧХ. Дисперсионная характеристика имеет линейный характер (рис. 2.14, б). Основное достоинство ДЛЗ несимметричной конструкции заключается в достаточно простом способе изменения наклона дисперсионной характеристики.

С помощью ДЛЗ можно формировать линейно-частотно-модулированные сигналы (ЛЧМ). Линейный закон изменения задержки сигнала от частоты может быть обеспечен уменьшением шага электродов ВШП.

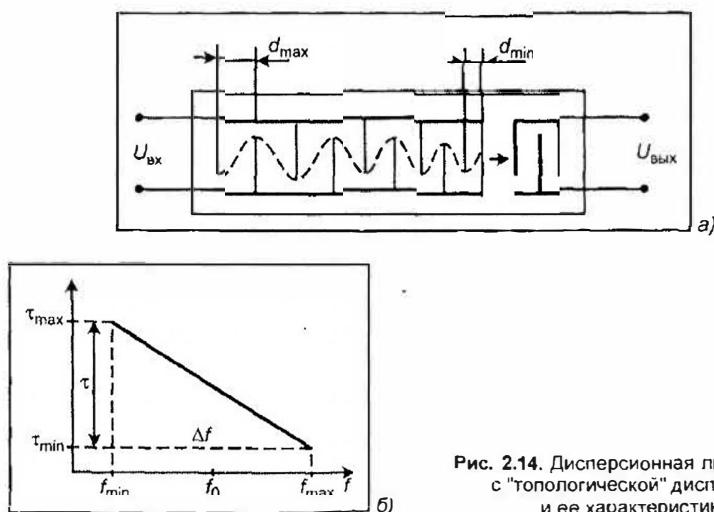


Рис. 2.14. Дисперсионная линия задержки с "топологической" дисперсией (а) и ее характеристика (б)

Можно также сформировать линейную дисперсионную характеристику, используя топологию с увеличивающимся шагом.

2.2.2. Устройства частотной селекции

Устройства частотной селекции представляют собой процессоры, предназначенные для выделения необходимых сигналов на фоне шумов и помех.

К устройствам частотной селекции относятся фильтры, которые можно классифицировать по следующим независимым признакам:

- по виду частотной характеристики: полосовые (пропускающие определенную полосу частот Δf), режекторные (подавляющие определенную полосу частот Δf), нижних частот (пропускающие частоты от 0 до f_0), верхних частот (пропускающие частоты выше f_0);
- по физическому принципу: резонансные (по акустическому или электрическому резонансу), трансверсальные (по фазочастотной характеристике), нетрансверсальные (по амплитудно-частотной характеристике);
- по виду обрабатываемых сигналов: аналоговые (обработка сигнала в виде непрерывной функции), цифровые (обработка сигнала в виде дискретной функции).

К устройствам частотной селекции относятся также резонаторы. Резонаторы на объемных акустических волнах были хорошо известны. ПАВ-резонаторы, как правило, самостоятельного значения не имеют, однако широко используются в конструкциях фильтров.

Полосовые фильтры получили наиболее широкое распространение среди элементов частотной селекции на ПАВ. Различают сверхузкополосные фильтры ($0,01\% < \Delta f / f_0 \leq 0,1\%$), узкополосные ($0,1\% < \Delta f / f_0 \leq 1\%$), среднеспосные ($1\% < \Delta f / f_0 \leq 10\%$), широкополосные ($10\% < \Delta f / f_0 \leq 50\%$) и сверхширокополосные ($50\% < \Delta f / f_0 \leq 100\%$). Полосовые фильтры используются в диапазоне частот от 10 МГц до 2 ГГц. Фильтрация частот в полосовых фильтрах осуществляется за счет селективных свойств ВШП.

Самым простым по конструкции из полосовых фильтров является однопроходной фильтр.

Рассмотрим элементы его конструкции. В качестве континуальной среды используются различные пьезоэлектрические материалы. Генератором динамических неоднородностей в виде ПАВ служит входной ВШП, имеющий многоэлектродную структуру. Амплитудно-частотная характеристика эквидистантного ВШП (рис. 2.15, а) может быть определена с помощью преобразования Фурье.

Спектр входного сигнала $U_{вх}(t)$ может быть представлен в виде суммы бесконечного числа синусоидальных колебаний с непрерывной последовательностью частоты и с бесконечно малыми амплитудами

$$dU_{вх} = \frac{1}{2\pi} \Phi_1(\omega) \exp(j\omega t) d\omega,$$

где $d\omega$ — бесконечно малый частотный интервал. Функция

$$\Phi_1 = \int_{-\infty}^{+\infty} U_{вх}(t) e^{-j\omega t} dt$$

представляет собой спектральную плотность амплитуды.

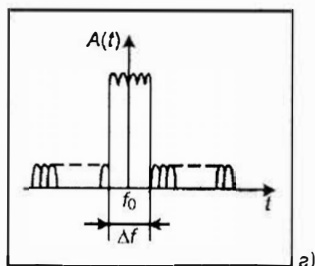
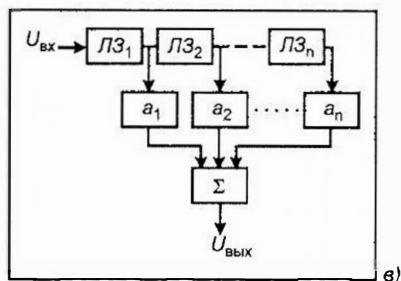
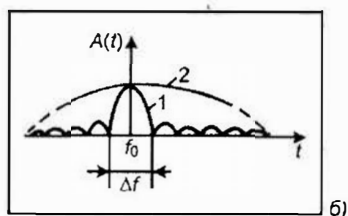
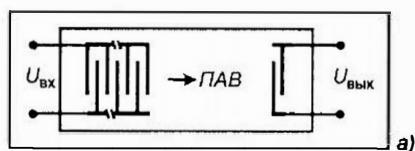


Рис. 2.15. Полосовой фильтр: а — конструкция; б — АЧХ фильтра; в — схема формирования трансверсального фильтра; г — суммарная АЧХ трансверсального фильтра

Если для любой частоты известен коэффициент передачи четырехполюсника

$$K(j\omega) = \frac{U_{\text{вых}}}{U_{\text{вх}}},$$

то по амплитуде элементарного воздействия можно найти амплитуду элементарного отклика:

$$d\tilde{U}(t) = dU_{\text{вх}} e^{j\omega t} = K(j\omega) e^{j\omega t} d\tilde{U}_{\text{вх}}.$$

Таким образом, элементарный отклик можно представить в виде:

$$dU_{\text{вых}}(t) = \frac{1}{2\pi} K(j\omega) \Phi_1(\omega) e^{j\omega t} d\omega.$$

Полный отклик равен сумме элементарных откликов:

$$U_{\text{вых}}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} K(j\omega) \Phi_1(\omega) e^{j\omega t} d\omega,$$

$$U_{\text{вых}}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi_2(\omega) e^{j\omega t} d\omega,$$

где $\Phi_2(\omega)$ — спектр отклика. Другими словами, спектр отклика равен спектру воздействия, умноженному на коэффициент передачи, или $\Phi_2(\omega) = K(j\omega) \Phi_1(\omega)$. Величина $K(j\omega)$ является *передаточной функцией*.

В исследованной конструкции фильтра входной сигнал представлен в виде импульса с прямоугольной огибающей, заполненной колебаниями с частотой f_0 . Входной ВШП

преобразует входной сигнал. Фурье-спектр этого сигнала может быть аппроксимирован функцией вида

$$A(\omega) = \text{const} \frac{\sin X}{X},$$

где $X = \frac{(N-1)}{2} \pi \frac{\Delta\omega}{\omega_0}$ (рис. 2.15, б, кривая 1).

Излучаемый ВШП сигнал распространяется в акустическом канале без искажения и поступает на выходной ВШП, который состоит из малого количества штырей и имеет достаточно широкую полосу пропускания (рис. 2.15, б, кривая 2). Полоса пропускания Δf такого полосового фильтра определяется входным ВШП и регулируется количеством штырей N . Конструкция фильтра мало чем отличается от ранее рассмотренной линии задержки. Такие конструкции, у которых свойства определяются фазочастотной характеристикой известны как *трансверсальные фильтры*. Селекция частот в них осуществляется с помощью топологии ВШП. В более сложных конструкциях используются несколько фильтров (линий задержек), сигналы от которых складываются синфазно. Фильтр имеет n линий задержки, отводы от которых характеризуются весовыми коэффициентами a_i . Сигналы от каждой линии задержки синфазно складываются в сумматоре и формируют сигнал на выходе фильтра (рис. 2.15, в). У нетрансверсальных фильтров свойства определяются амплитудно-частотной характеристикой.

Если входной сигнал представлен в виде $U_{вх} = U_0 e^{j\omega t}$, то выходной сигнал с сумматора с помощью преобразования Фурье описывается зависимостью:

$$U_{вых} = U_0 \sum_{i=1}^n a_i e^{j\omega t \tau_i},$$

где τ_i — время задержки сигнала в линии. Частотная характеристика фильтра с множеством ЛЗ представлена на рис. 2.15, г и представляет собой сумму функций типа $(\sin X)/X$ для большого числа N . Частотная характеристика трансверсального фильтра имеет почти прямоугольную форму.

Следует заметить, что каждую пару электродов ВШП можно интерпретировать как отводы фильтра, шины в качестве сумматора, а весовые коэффициенты a_i определяются топологией электродов. Очевидно, что, изменяя топологию электродов, их взаимное перекрытие можно менять значения a_i и, соответственно, АЧХ полосового фильтра. Такой метод весовой обработки компонентов сигнала получил название *аподизации преобразователя*. В принятой модели изделия функциональной электроники аподизированный преобразователь представляет собой устройство управления.

Степень перекрытия электродов при постоянной амплитуде ПАВ изменяет величину фронта волны, что и сказывается на значении весового коэффициента. Различают два основных метода аподизации преобразователей.

Внешнее взвешивание реализуется с помощью внешних шунтов емкостного или резистивного характера, которые позволяют задавать на электродах (рис. 2.16, а) разность потенциалов (рис. 2.16, б), необходимую для формирования ПАВ. Интенсивность электрического поля между соседними электродами ВШП определяется только разностью потенциалов, поскольку апертура постоянна. Интенсивность формируемой ПАВ пропорциональна интенсивности электрического поля. Энергетические характеристики волны отражены на эпюрах. Амплитуда волны пропорциональна напряжению при постоянной апертуре W .

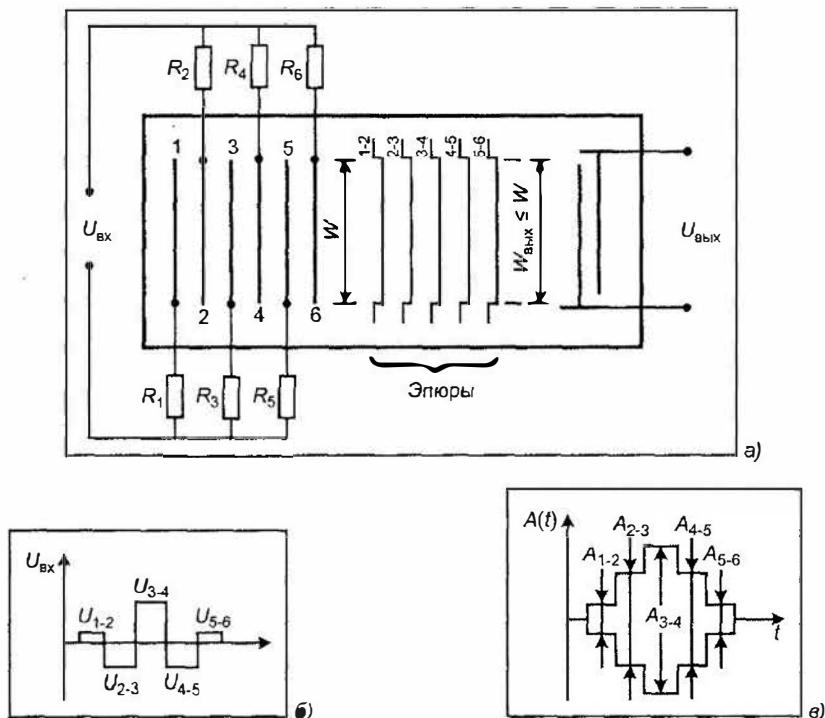


Рис. 2.16. Фильтр с внешним взвешиванием: а — общая схема; б — входной сигнал; в — амплитуда откликов элементов ВШП

Детектирование сигнала происходит на выходном ВШП, имеющем любую апертуру, но достаточно широкополосном, позволяющем получить суммарный сигнал от каждого элемента импульсного отклика в виде $A = f(t)$ (рис. 2.16, в).

Другой метод аподизации — топологический или метод непосредственного взвешивания. Сущность этого метода заключается в том, что амплитудная модуляция импульсного отклика осуществляется заданным изменением апертуры перекрытия электродов. Энергия ПАВ, излучаемая каждой парой электродов ВШП, зависит от степени их перекрытия и излучается в канале соответствующей ширины (рис. 2.17, а). Модуляция в этом случае осуществляется апертурой акустического потока и моментальный входной импульсный отклик равен площади апертуры W_{ijl} (рис. 2.17, б), а выходной (детектирующий) преобразователь должен иметь апертуру

$$W_{вых} \geq W_{max}$$

где W_{max} — максимальная апертура перекрытия электродов и его импульсный отклик представлен на рис. 2.17, в.

Фильтр с топологическим взвешиванием электродов конструктивно проще фильтра с внешним взвешиванием электродов, прежде всего потому, что внешние шунты нарушают микроэлектронную технологию. Недостатком такого фильтра является невозможность разделения фронтов двух аподизированных преобразователей, работающих в одном акустическом канале. В этом случае применяют резонаторы ПАВ, позволяющие разделить акустические каналы.

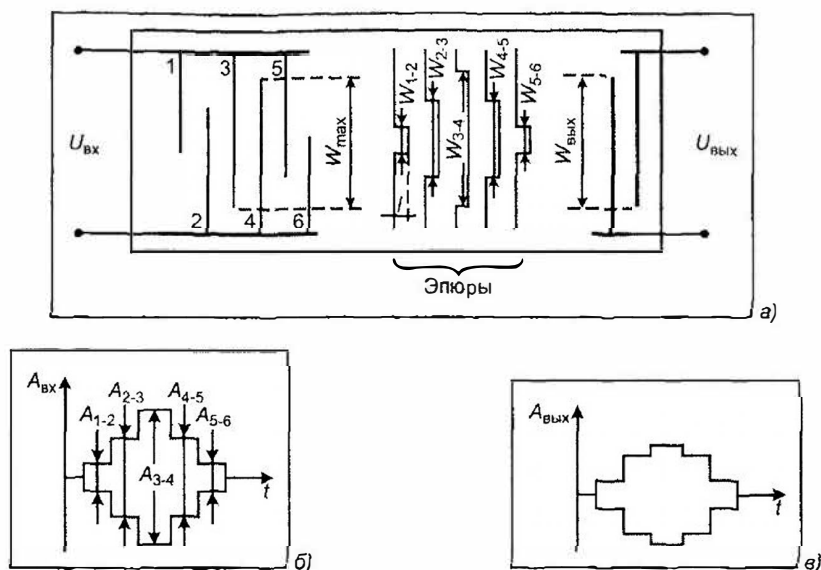


Рис. 2.17. Фильтр с топологическим (непосредственным) взвешиванием: а — общая схема; амплитуда импульсного отклика входного (б) и выходного (в) преобразователей

ПАВ-резонаторы предназначены для стабилизации частоты генерируемых колебаний, а также используются в качестве узкополосных фильтров. Амплитудно-частотная характеристика представляет собой резонансную кривую, и подавление вне полосы пропускания оказывается недостаточно сильным.

Резонанс колебаний формируется в резонаторе, реализованном на ВШП либо зеркалах из отражательных решеток. Условие резонанса определяется соотношением

$$L = n\lambda / 2,$$

где n — целое число, L — размер резонирующей полости. Коэффициент отражения зеркала R определяется выражением

$$R = N|r|,$$

где N — число элементов решетки, $|r|$ — модуль коэффициента отражения каждого элемента. В типичных случаях $|r| \approx 0,01$ и тогда $N \geq 100$ элементов. Резонатор с шагом d проектируется, исходя из условия подавления всех колебаний, кроме одного собственного колебания резонатора $f_0 = V_0 / 2d$. Достигается это в том числе и за счет частотной зависимости коэффициента и фазы отражения решетки. На рис. 2.18, а представлена конструкция узкополосного фильтра, реализованного на основе резонаторов. Они размещены на одной подложке и акустически связаны с помощью многополоскового отражателя.

ПАВ-резонаторы очень часто используются в устройствах для фильтрации частот. Отражательные структуры могут быть выполнены в виде металлических электродов (рис. 2.18, б), топографических (рис. 2.18, в) или диффузионных статистических неоднородностей (рис. 2.18, г). В любом случае основой является технология микроэлектроники. Фильтры на основе ПАВ-резонаторов предназначены для перестройки частоты с помощью управляемого изменения параметров ПАВ и разделения акустических каналов двух аподизированных преобразователей.

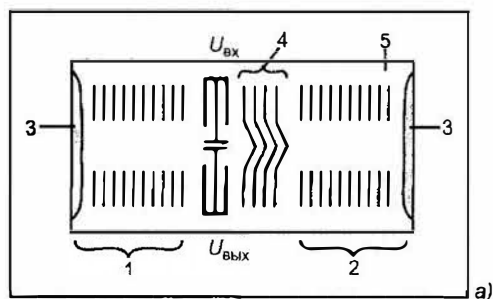
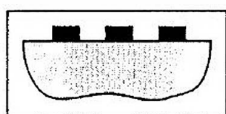
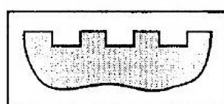


Рис. 2.18. ПАВ-резонатор, выполняющий функцию фильтра:

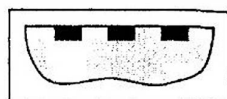
- а — резонаторные полости (1, 2), акустические поглотители (3), многополосковый ответвитель (4), кварцевая подложка (5);
б, в, г — отражательные структуры



б)



в)



г)

В соответствии с принятой моделью устройства функциональной электроники между генератором и детектором динамических неоднородностей вводится дополнительное устройство управления в виде многополоскового отражателя (МПО). Например, МПО расположены непосредственно в акустическом канале на рабочей поверхности звукопровода и конструктивно могут быть выполнены в виде напыленных электродов или распределенных отражателей — *канавок* (акустических неоднородностей). Как правило, МПО представляют собой эквидистантную решетку отражателей с шагом d , так что центральная частота $f_0 = V_a / 2d$. Добротность ПАВ-резонаторов определяется в основном потерями на распространение волны в резонирующей полости, дифракционными эффектами, генерацией паразитных объемных волн, потерями энергии волны в самих отражателях и составляет $\sim 10^4$.

Введением входного ВШП между ПАВ-отражателями можно получить ПАВ-резонатор с одним входом. Такая структура поддерживает несколько стоячих волн и позволяет выделить нужную моду путем выбора соответствующей конструкции электродов преобразователя и определенной формы отражателя.

Повышения селективности полосовых фильтров можно достичь путем использования аподизированных входного и выходного преобразователей, расположенных в параллельных акустических каналах. Связь между каналами и управление распространением ПАВ осуществляется с помощью МПО (рис. 2.19, а).

В процессе переизлучения волны из верхнего в нижний канал происходит преобразование апертурной модуляции ПАВ в модуляцию интенсивности звуковой волны. Эпюры (Э1) топологического взвешивания с помощью МПО преобразуются в эпюры внешнего взвешивания (Э2). Это позволяет принимать сигнал выходным ВШП, аподизированным непосредственным взвешиванием электродов. В таком фильтре происходит достаточно сильное подавление сигнала вне полосы пропускания входного (рис. 2.19, б) и выходного (рис. 2.19, в) преобразователей. Амплитудно-частотная характеристика такого фильтра получается путем перемножения парциальных характеристик:

$$A_T = A_{\text{вх}} \times A_{\text{вых}}.$$

Полосовые фильтры позволяют реализовать заданную АЧХ с подавлением за полосой пропускания до 40—70 дБ с коэффициентом прямоугольности формы $1,2 \div 1,5$. Фильтры рассчитаны на центральные частоты от 10^7 до 10^9 Гц с полосой пропускания от 10^{-1} до 4×10^5 Гц.

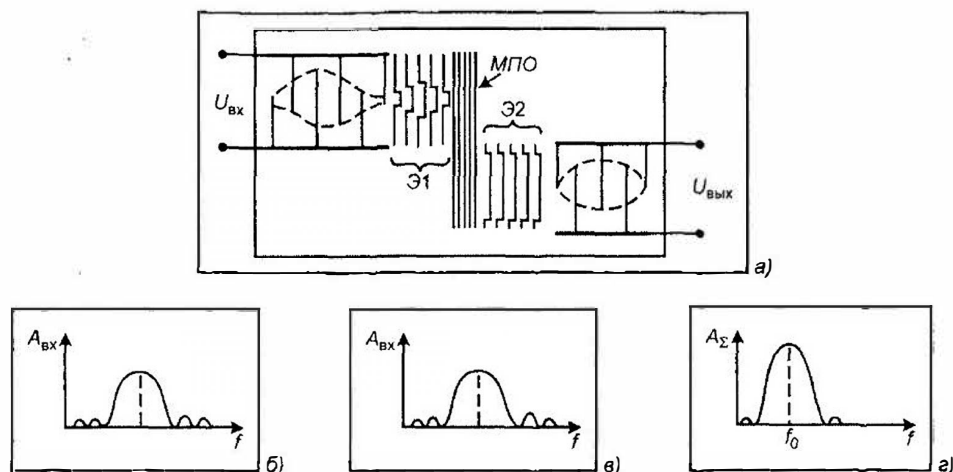


Рис. 2.19. Фильтр с использованием ПАВ-резонаторов: а — общая схема; б — АЧХ входного ВШП; в — АЧХ выходного ВШП; г — суммарная АЧХ

Микроэлектронная технология производства полосовых фильтров позволяет легко интегрировать их в радиотехнические цепи и системы, тиражировать без подстройки параметров.

Системный анализ показал, что улучшение характеристик полосовых фильтров возможно за счет снижения вклада физических эффектов второго порядка:

- ☐ многократные отражения ПАВ от краев ВШП;
- ☐ сигналы тройного прохождения, обусловленные преобразователями;
- ☐ электроакустическая регенерация вследствие постоянного чередования электрических и акустических полей;
- ☐ дифракция пучка ПАВ и его отклонение от направления распространения;
- ☐ отражение от краев звукопровода;
- ☐ ложные сигналы, вызываемые объемными, продольными и поперечными модами;
- ☐ изменение скорости ПАВ из-за поверхностной загрузки звукопровода;
- ☐ дисперсия скорости звука в металлизированных областях;
- ☐ искажение фазового фронта ПАВ вследствие различия прохождения в центральной и периферийной областях;
- ☐ краевые эффекты взаимодействия соседних электродов;
- ☐ паразитные электростатическая и электромагнитная связи входного и выходного преобразователя;
- ☐ искажение характеристик внешними цепями согласования и т. п.

Полосовые фильтры имеют достаточно широкий диапазон частот и полос пропускания (рис. 2.20). Поэтому они нашли применение в качестве фильтров промежуточной частоты для телевизионных приемников, в системах связи и радиолокации.

Дисперсионные фильтры представляют собой функциональное линейное устройство, предназначенное для получения задержки, зависимой от частоты и имеющее модулиро-

ванную импульсную характеристику. Устройства со сжатием импульса находят широкое распространение в радиолокационных системах. В этих системах увеличение длительности зондирующего импульса позволяет увеличить чувствительность за счет увеличения отношения сигнал/шум. При этом необходимо сузить полосу пропускания приемника и снизить мощность шума на входе.

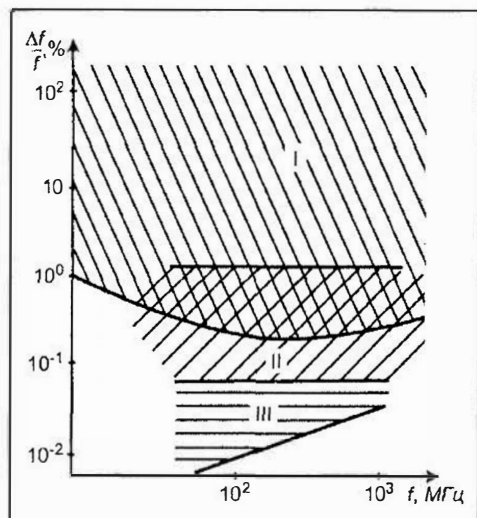


Рис. 2.20. Рабочий диапазон полосовых фильтров:
1 — широкополосные трансверсальные фильтры;
2 — узкополосные фильтры на резонаторах;
3 — узкополосные фильтры на термостабилизированных резонаторах

Однако большая длительность импульса ведет к ухудшению разрешающей способности по дальности и невозможности различать две близко находящиеся цели. Это противоречие можно разрешить методом сжатия импульсов с использованием дисперсионных фильтров.

В системе со сжатием импульса используются ЛЧМ-сигналы с большой базой, другими словами, произведение длительности импульса на полосу пропускания имеет большую величину. Основное свойство ЛЧМ-сигнала состоит в том, что форма его огибающей при амплитудной модуляции воспроизводит форму модуля его спектральной плотности в частотной области. Идея сжатия ЛЧМ-сигнала приведена на рис. 2.21. Сигнал $S(t)$ от точечной цели имеет форму, представленную на рис. 2.21, а и характеризуется увеличением во времени частоты сигнала. Фильтр необходимо изготовить так, чтобы его импульсная характеристика $\Phi(t)$ представляла обращенный во времени сигнал. Тогда частота такого сигнала падает линейно со временем (рис. 2.21, б).

Выходной суммарный сигнал $g(t)$, прошедший дисперсионный фильтр, имеет вид, представленный на рис. 2.21 в. Такой сигнал называется корреляционным пиком, ширина спектра которого является обратной величиной ширины спектра входного сигнала, а ширина выходного пика намного меньше длительности входного сигнала T . Величина T/B называется коэффициентом сжатия, где B — девиация частоты.

Если имеются две точечные цели, то на вход поступают два сигнала, разделенные временным интервалом. В выходном сигнале также будут присутствовать два корреляционных пика. Если оба импульса перекрываются во времени, то их разрешение возможно при условии превышения разницы времени задержки длительности пика T/B . Типичное значение ширины спектра составляет 20 МГц.

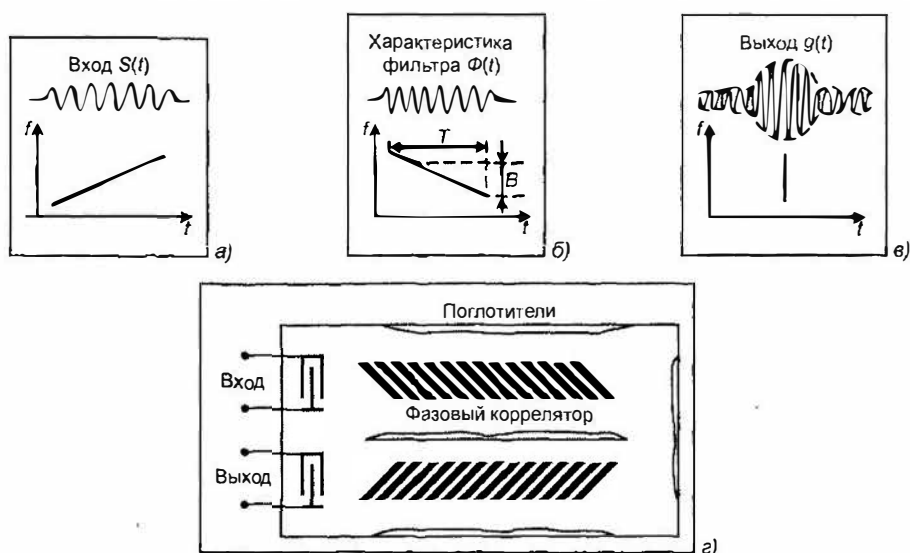


Рис. 2.21. Обработка ЛЧМ-сигналов дисперсионным фильтром: а — входной сигнал $S(t)$; б — характеристика фильтра $\Phi(t)$; в — выходной сигнал; г — конструкция дисперсионного фильтра на отрагательных решетках

На рис. 2.21, г представлена конструкция дисперсионного фильтра. Входной и выходной ВПП практически идентичны. В качестве устройства управления ПАВ используются отрагательные неэквидистантные решетки из канавок, шаг которых меняется от $d_{\max} = V_a / 2f_{\min}$ до $d_{\min} = V_a / 2f_{\max}$. Каждая из решеток меняет направление волн на 90° .

В фильтрах ЛЧМ-сигналов используется амплитудное взвешивание для подавления помех. С этой целью изменяется глубина канавок, которая пропорциональна коэффициенту отражения. Канавки травятся методом ионной бомбардировки. Такая технология весьма трудоемка. ЛЧМ-фильтры позволяют обрабатывать сигналы длительностью $T \sim 10^{-4}$ с и девиацией $B = \Delta f \approx (3 \div 5) \times 10^8$ Гц.

2.2.3. Генераторы на ПАВ

Акустоэлектронный генератор является активным акустоэлектронным устройством, предназначенным для генерации акустических сигналов.

Исследованы конструкции генераторов с позиций предложенной модели прибора функциональной электроники. Различают два типа генераторов на ПАВ в зависимости от используемого стабилизирующего элемента. В первом типе используется ПАВ-линия задержки, включенная в цепь положительной обратной связи усилителя (рис. 2.22, а).

Линия задержки позволяет возбудить колебания на частотах ω , определяемых соотношением

$$\omega = 2\pi V_a / L.$$

где L — длина линии задержки.

Для получения одномодового режима необходимо строгое выполнение соотношения

$$L = n\lambda,$$

где λ — длина звуковой волны, n — целое число. Частота генераций может быть изменена за счет изменения скорости ПАВ в резонирующей полости или введением фазового сдвига во внешней цепи. Спектральные характеристики генератора на ЛЗ приведены на рис. 2.22, б.

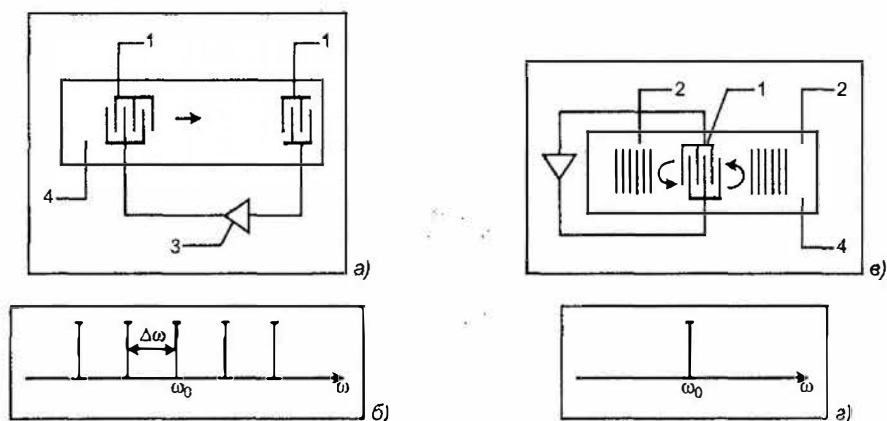


Рис. 2.22. ПАВ-генератор на основе ЛЗ (а) и его спектральная характеристика (б); на основе резонатора (в) и его спектральная характеристика (г); 1 — ВШП; 2 — отражательные структуры; 3 — усилитель; 4 — подложка

В генераторе второго типа в качестве стабилизатора частоты используется ПАВ-резонатор с одним входом (рис. 2.22, в). Усилитель подключается между разнополярными шинами преобразователя. Селективные свойства генератора определяются размерами резонирующей полости и параметрами отражательной структуры. Добротность ПАВ-резонаторов составляет $\sim 10^4$ и это позволяет создавать конкурентоспособные генераторы. В одновходовых генераторах возникают трудности с перестройкой частот. Линейная модуляция в таких генераторах обеспечивается в очень узкой полосе.

У генераторов этого типа шумовые характеристики лучше, чем у генераторов на ПАВ ЛЗ. На рис. 2.22, г приведены спектральные характеристики генератора на ПАВ-резонаторе.

Оба типа генераторов могут работать на основных частотах от 10^7 до 10^9 Гц без использования дополнительных схем умножения частоты.

Анализ показал, что важным преимуществом генераторов на ПАВ является возможность частотной модуляции, а также заданный переход к нужным частотам при условии сохранения стабильности частоты: кратковременной 10^{-9} с $^{-1}$, долговременной 10^{-6} год $^{-1}$ и температурной $\sim 10^{-6}$ °C $^{-1}$ в диапазоне до $+50$ °C. Такие генераторы могут быть использованы в радиолокационных системах, дальномерах на основе доплеровского сдвига частоты, а также в стандартных радиотехнических устройствах.

2.2.4. Усилители

Усилители представляют собой акустоэлектронное устройство, предназначенное для усиления акустических сигналов.

Во всех ранее рассмотренных устройствах функциональной электроники уровень входного сигнала существенно выше уровня выходного, преобразованного сигнала. Это вполне закономерно, т. к. значительная часть энергии затрачивается на преобразование одного

вида сигнала в другой и обратно, прохождение через звукопровод и т. п. Для сохранения уровня сигнала необходимо встраивать усилители, что приводит к значительным технологическим трудностям.

Решение проблемы усиления сигналов можно найти, более пристально изучая физические процессы при распространении динамических неоднородностей. Ранее уже указывалось, что динамические неоднородности в виде ПАВ являются результатом взаимодействия электрических полей с пьезоэлектриками. Распространение акустических волн сопряжено с распространением волн зарядовой плотности. Усиление акустических волн можно производить как за счет подачи на электроды ВШП большего сигнала, так и за счет "перекачки" энергии из волн зарядовой плоскости в акустические волны. Это явление акустоэлектронного взаимодействия, широко используемое в конструкциях усилителей.

В зависимости от типа континуальной среды различают ионное взаимодействие (в металлах), потенциал-деформационное взаимодействие (в полуметаллах и полупроводниках) и пьезоэлектрическое взаимодействие (в пьезополупроводниках). Если к кристаллу пьезополупроводника приложить внешнее постоянное электрическое поле E_0 , создающее дрейф электронов в направлении ПАВ, то усиление будет зависеть от соотношения скорости дрейфа носителей $V_{др}$ и скорости звука V_a (рис. 2.23, а).

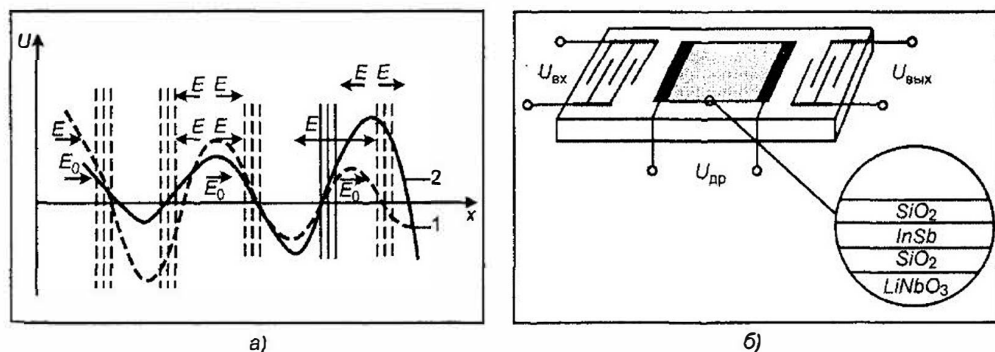


Рис. 2.23. Схема взаимодействия динамических неоднородностей в полупроводнике с пьезоэлектрическими свойствами (а), конструкция усилителя на ПАВ (б); 1 — акустическая волна и зарядовые пакеты в отсутствии электрического поля E_0 ; 2 — то же в присутствии поля

Так, при скорости дрейфа носителей, меньшей скорости распространения ПАВ $V_a > V_{др} = \mu E_0$ (μ — подвижность электронов), энергия волны поглощается электронным газом. При скорости дрейфа, большей фазовой скорости ПАВ $V_{др} > V_a$, происходит перекачка энергии электронов в энергию акустической волны, возрастает ее амплитуда в канале распространения и, соответственно, на выходных ВШП получится усиленный сигнал.

Таким образом, за счет перераспределения энергии в динамических неоднородностях различной физической природы можно получить эффект усиления. Исследовались конструкции на основе этого эффекта с позиций предложенной модели прибора. В частности акустоэлектронный усилитель, схема которого приведена на рис. 2.23, б. С целью обеспечения взаимодействия динамических неоднородностей различной физической природы предложены две континуальные среды.

В одной преимущественно распространяются акустические волны (пьезоэлектрик), в другой — электронные (полупроводник). Совмещение сред осуществляется путем создания

слоистых структур с использованием технологии тонких пленок. В качестве генератора ПАВ используется стандартная конструкция ВШП. Звукопровод выполняется, например, на основе ниобата лития, на поверхности которого на изолирующей пленке диоксида кремния выращивается пленка сурьмянистого индия толщиной $\sim 500 \text{ \AA}$. Вся поверхность покрывается достаточно толстой ($\sim 1000 \text{ \AA}$) защитной пленкой диоксида кремния. Края полупроводниковой пленки снабжены омическими контактами для подвода дрейфового напряжения. В качестве полупроводниковых сред могут быть использованы также CdS , CdSe , ZnS , ZnO , GaAs и др.

Детектором является ВШП, на который поступает усиленная акустическая волна. В качестве устройства управления используется пленка полупроводника в слоистой структуре на тракте распространения сигнала.

Анализ показал, что акустоэлектронные усилители могут работать как в импульсном, так и в непрерывном режимах усиления. Коэффициент электронного усиления лежит в пределах от 10 до 90 дБ, коэффициент шума до 10 дБ на центральных частотах 10^8 Гц с полосой $\sim 10\%$.

Акустоэлектронные усилители выполняются по микроэлектронной технологии и легко монтируются в соответствующие системы.

2.3. Нелинейные устройства

2.3.1. Физические основы

В рассмотренных выше устройствах функциональной акустоэлектроники использовались линейные свойства твердых тел при прохождении акустоэлектронной волны.

Весьма привлекательной для исследователей оказалась область нелинейного взаимодействия акустических волн с твердым телом. В этом случае не выполняется принцип суперпозиции гармонических колебаний, имеет место сложение частот сигналов, взаимодействие электрических и акустических полей в твердом теле.

Нелинейные явления в твердых телах сводятся к двум типам взаимодействия:

- ☐ взаимодействие различных акустических волн;
- ☐ взаимодействие акустических волн с внешним электрическим полем.

При нелинейном взаимодействии акустических волн возникает зависимость фазовой скорости волны от амплитуды и изменение формы гармонического колебания. Искажение формы волны связывается с искажением ее спектра за счет обогащения ее гармониками. Если в твердом теле взаимодействуют несколько акустических колебаний, то в результате их взаимодействия появляются волны комбинационных частот — Фурье-компоненты. При этом должны выполняться условия синхронизма

$$\omega_{\Sigma} = \omega_1 \pm \omega_2,$$

$$\vec{K} = \vec{K}_1 \pm \vec{K}_2,$$

где ω_i и \vec{K}_i — частоты и волновые векторы взаимодействующих волн, ω_{Σ} и \vec{K}_{Σ} — соответственно параметры результирующей волны. На языке квазичастиц — фононов условие синхронизма записывается в виде:

$$\hbar\omega_1 \pm \hbar\omega_2 = \hbar\omega_{\Sigma},$$

$$\hbar\vec{K}_1 \pm \hbar\vec{K}_2 = \hbar\vec{K}_{\Sigma},$$

где $\hbar = \frac{h}{2\pi}$, h — постоянная Планка. Это так называемое явление фонов-фононного взаимодействия.

Анализ взаимодействия акустических волн может быть произведен с помощью метода дисперсионных диаграмм, в котором каждая из взаимодействующих волн представлена парой (ω, \vec{K}_i) , а суммарная волна является суммарным вектором.

При распространении акустической волны в континуальной среде со свободными носителями (полупроводник) движение динамических неоднородностей акустической природы порождает динамические неоднородности электрической природы — волны зарядовой плотности, которые также можно описать уравнениями:

$$n_1 = n_1^0 \exp(j(\omega_1 t - \vec{K}_1 \vec{r}_1))$$

$$n_2 = n_2^0 \exp(j(\omega_2 t - \vec{K}_2 \vec{r}_2)),$$

где \vec{r}_1 и \vec{r}_2 — направления распространения, t — шкала времени.

При нелинейном взаимодействии электрическое поле одной акустической волны может взаимодействовать с носителями, сопровождающими другую волну.

Возникает нелинейная индукция $D_{n\omega}$ и нелинейная составляющая электрического тока

$$J_{n\omega} \sim \partial D_{n\omega} / \partial t$$

или

$$J_{n\omega} = q\mu(E_1 n_2 + E_2 n_1),$$

где q и μ — соответственно заряд и подвижность носителей, E , n — компоненты поля и концентрации носителей волны, соответственно. На распространение акустической волны в нелинейной среде существенное влияние оказывает внешнее электрическое поле.

Например, если в нелинейной среде распространяется акустическая волна с частотой ω_1 и волновым вектором \vec{K}_1 , и на среду воздействует однородное в пространстве поле частотой $\omega_2 = 2\omega_1$ и $\vec{K}_2 = 0$, то образуется новая акустическая волна с частотой $\omega_\Sigma \approx \omega_1$ и вектором $\vec{K}_\Sigma = -\vec{K}_1$. Акустическая волна будет распространяться уже в противоположном направлении (рис. 2.24, а).

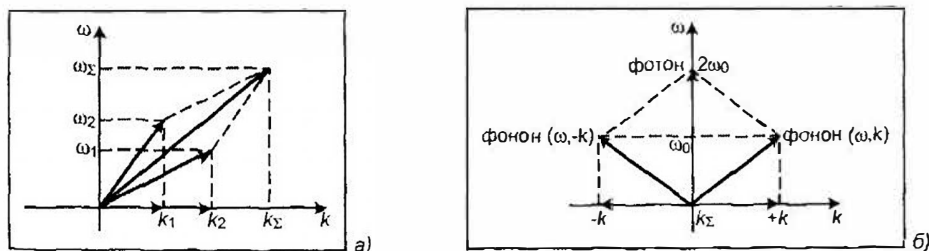


Рис. 2.24. Дисперсионные диаграммы взаимодействия акустических волн с частотами ω_1 и ω_2 и волновыми векторами \vec{K}_1 , \vec{K}_2 (а), параметрическое взаимодействие фононов с электромагнитным полем накачки (б)

Еще одним примером взаимодействия динамических неоднородностей различной физической природы является осуществление нелинейного взаимодействия фононов с элек-

ромагнитным излучением поля накачки. Значения скоростей акустической и электромагнитных волн в твердом теле различаются на пять порядков. Тем не менее, с помощью накачки потоком фотонов удастся осуществить генерацию обратной акустической волны. Такой процесс интерпретируется как распад фотона с частотой ω и волновым вектором \vec{K} на два фонона с векторами $\vec{K}_1 = -\vec{K}_2$ и частотами $\omega_1 = \omega_2 = \omega / 2$.

При взаимодействии акустической волны с электрическим полем может возникнуть нелинейная составляющая электрической индукции. Происходит перераспределение носителей, например, в ловушках и образуется потенциальный рельеф, отражающий моментальное состояние электрического поля. Такая "память" определяется временем жизни носителей в ловушках.

Если подать на среду электрическое поле нужной частоты и ориентации вектора, то "замороженное" электрическое поле в виде потенциального рельефа может быть считано в виде акустических волн, распространяющихся в прямом и обратном направлениях. Это так называемые многофононные процессы, при которых может возникнуть стимулированное трехфотонное эхо, которое можно рассматривать как динамическую неоднородность.

Механизм формирования сигналов фононного эха заключается в подаче ВЧ-импульса ($\sim 10^6 \div 10^8$ Гц) на твердое тело, помещенное в колебательную систему, который генерирует упругие волны тех же частот. Волны разной поляризации (определяемой волновым вектором \vec{K}) распространяются с разными скоростями, и происходит рассогласование по фазе. Через время τ подается второй ВЧ-сигнал, который генерирует волны тех же частот с теми же фазовыми соотношениями, но с противоположными значениями волнового вектора \vec{K} (рис. 2.25). За время $t = \tau$ рассогласованные по фазе волны при распространении в обратном направлении снова окажутся в фазе. За счет обратного пьезоэффекта в момент $t = 2\tau$ возникает электрический сигнал такой же частоты, представляющий собой сигнал двухимпульсного эха. В результате нелинейного взаимодействия между упругими волнами, возбужденными первым импульсом и электрическим полем второго импульса, формируется периодическая структура электрического поля и механического напряжения. Период этой структуры равен периоду волны. Электрическое поле третьего импульса, действующее на такой "электрический" рельеф в момент времени T , генерирует упругие волны с противоположным направлением волнового вектора. Через время τ после подачи третьего импульса эти волны окажутся в фазе и в момент времени $T + \tau$ дадут сигнал стимулированного эха. Записанный "электрический рельеф" или своеобразная голограмма, записанная на дислокациях и дефектах кристаллической решетки может существовать несколько суток и, соответственно, эффект трехимпульсного эха может быть воспроизведен через несколько суток.

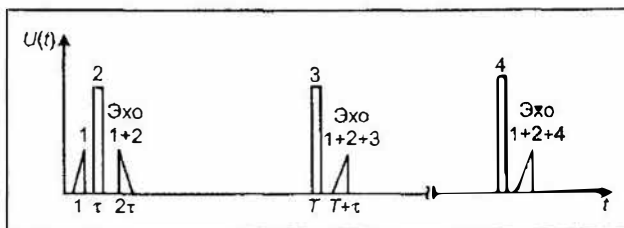


Рис. 2.25. Временная диаграмма формирования фононного эха

Заметим, что аналогично рассмотренным механизмам нелинейного взаимодействия динамических неоднородностей электрической и акустической природы существуют и механизмы нелинейного взаимодействия динамических неоднородностей акустической и магнитной природы. Это может происходить в континуальных средах, обладающих магнитострикционными свойствами.

Особый интерес представляет собой явление обращения волнового фронта (ОВФ), заключающееся в таком преобразовании волнового поля, при котором сохраняется первоначальное пространственное распределение амплитуды и фазы волны при изменении вектора распространения волны на противоположное. Другими словами, явление обращения волнового фронта приводит к инверсии волны во времени.

ОВФ в нелинейных средах может происходить по двум типам четырехволнового взаимодействия: параметрическом и голографическом.

В *параметрическом* типе взаимодействия в результате взаимодействия встречных волн происходит пространственная однородная модуляция среды на удвоенной частоте. В этом случае результатом параметрического взаимодействия переменного возмущения среды и сигнальной волны является обращенный волновой фронт.

В *голографическом* типе амплитуда и фаза сигнальной волны записывается в процессе ее взаимодействия с волной накачки. Частота волны накачки должна соответствовать частоте сигнальной волны. Обобщенная информация о распределении амплитуды и фазы записывается в виде пространственного неоднородного возмущения среды. При считывании информации волной, аналогичной волне накачки и распространяющейся навстречу записывающей волне, происходит "размораживание" — генерация обращенного волнового фронта.

Нелинейные взаимодействия эффективно используются в устройствах обработки и хранения информации — высокопроизводительных процессорах и ЗУ большой емкости.

2.3.2. Конвольверы

В радиоизмерительных и локационных устройствах широко используются методы корреляционной обработки сигналов, заключающиеся в одновременной обработке входного и опорного сигналов путем интегрирования результатов перемножения за время существования сигнала. Функция взаимной корреляции двух сигналов S_1 и S_2 в общем случае имеет вид (часть III, глава 9, формула (9.12)):

$$B_{S_1 S_2}(\tau) = \int_{-\infty}^{+\infty} S_1(t) S_2^*(t + \tau) dt = S_1 \times S_2.$$

Интегральное преобразование можно осуществить методом аналоговой обработки сигналов, используя нелинейное взаимодействие акустических колебаний. Такой процессор сигналов называется *акустическим конвольвером* (рис. 2.26, а). На левый ВШП 1 подается сигнал

$$S_1(t, \omega) = U(t) \exp j(\omega_1 t - \vec{K}_1 \vec{r}),$$

а на правый 2 —

$$S_2(t, \omega) = U(t) \exp j(\omega_2 t - \vec{K}_2 \vec{r}).$$

Синфазное взаимодействие происходит при выполнении условия $\omega_{\Sigma} = \omega_1 + \omega_2$ и $\vec{K}_{\Sigma} = \vec{K}_1 + \vec{K}_2$. Суммарный сигнал может быть зарегистрирован с помощью двух сплош-

ных параметрических металлических электродов 3 длиной L , перекрывающих звукопровод (4).

Выходной сигнал, снимаемый с электродов 3, можно записать как:

$$S_z = \int_{-L/2}^{+L/2} U(t - \frac{z}{V_a}) V(t + \frac{z}{V_a}) dz, \quad (2.6)$$

где V_a — скорость звука, z — координата.

Введем новые переменные

$$\tau = t - \frac{z}{V_a},$$

и тогда уравнение (2.6) может быть переписано в виде:

$$S_c = \int_{-T/2}^{+T/2} U(\tau) V(2t + \tau) d\tau,$$

где $T = L / V_a$ — время интегрирования. Если исходные сигналы $U(t)$ и $V(t)$ имеют прямоугольную огибающую, то их свертка представляет треугольный импульс (рис. 2.26, б). Амплитуда треугольного импульса имеет амплитуду, пропорциональную длительности T_c при условии, что $T_c \leq T$. Если $T_c > T$, то выходной сигнал имеет форму трапеции с пиковой амплитудой, равной амплитуде треугольного импульса при $T = T_c$.

В случае, когда на левый выход поступает сигнал

$$S_1(t, \omega) = U(t) \exp j(\omega_1 t - \vec{K}_1 \vec{r}),$$

а на выходной электрод —

$$S_2(t, \omega) = V(t) \exp j(\omega_1 t - \vec{K}_2 \vec{r}),$$

$$S_2(t, 2\omega) = V(t) \exp(j2\omega_1 t - \vec{K}_2 \vec{r}).$$

В каждой точке пространства в момент времени t амплитуда акустического сигнала на частоте 2ω будет пропорциональна

$$U(t - \frac{z}{V_a}) \cdot V(t).$$

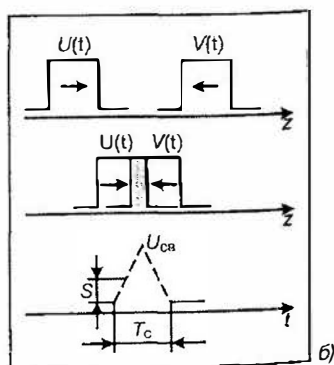
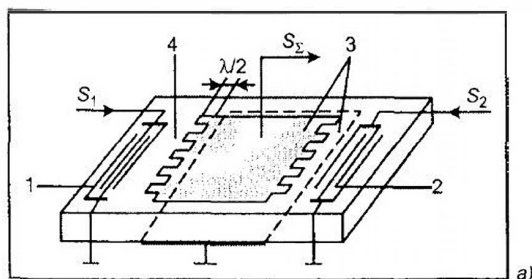


Рис. 2.26. Конвольвер на ПАВ (а) и процесс образования свертки двух сигналов $U(t) * V(t)$ (б)

Появится новая акустическая волна на частоте ω_1 , которая будет распространяться в противоположном направлении (рис. 2.26, а). На левом входном ВПП появится сигнал, который будет представлять функцию акустической корреляции:

$$S_k = \int_{-L/2}^{L/2} U\left(t - \frac{2z}{V_a}\right) \cdot V\left(t - \frac{z}{V_a}\right) dz$$

или при замене переменных $\tau = t - \frac{z}{V_a}$ получаем

$$S_k = \int_{-\infty}^{\infty} U(2\tau - t) \cdot V(\tau) d\tau$$

или

$$S_k = U * V,$$

т. е. происходит обращение сигнала во времени.

Конвольвер как процессор сигналов выполняет несколько функций, например корреляционную обработку сигналов и операцию свертки. В табл. 2.1 приведены виды преобразований.

Конвольвер выполняется по микроэлектронной технологии. В конструкции конвольвера можно найти все пять элементов, характерных для изделия функциональной электроники.

Таблица 2.1. Виды преобразований, осуществляемые взаимодействием акустических волн

Входной сигнал	Выходной сигнал	Результат
Взаимодействие встречных волн		
Произвольный	$\int_{-\infty}^{+\infty} U_1(\tau) U_2(2t - \tau) d\tau$	Сжатая по времени свертка
$U_2(t) = U_1(-t)$	$\int_{-\infty}^{+\infty} U_1(\tau) U_1(\tau - 2t) d\tau$	Сжатая по времени автокорреляция
Генерация обратной волны		
Произвольный	$\int_{-\infty}^{+\infty} U_2(\tau) U_2(2\tau - t) d\tau$	Взаимная корреляция $U_2(t/2), U_1(t)$
$U_3(t) = U_1(2t)$	$\int_{-\infty}^{+\infty} U_1(2\tau) U_1(2\tau - t) d\tau$	Автокорреляция $U_1(t), U_3(t/2)$
$U_3(t) = \sigma(t)$	$U_1(-t)$	Обращение во времени

Информационная емкость конвольвера определяется величиной

$$C_m = B \cdot T_m$$

где B — полоса пропускания ВШП, $T_{\text{и}}$ — время интегрирования ($C_{\text{и}} = 10^3 \div 10^4$). Динамический диапазон определяется уровнем ложных сигналов на выходе и нелинейным насыщением выходного сигнала. Для снижения уровня ложных сигналов применяют параметрический электрод с зубчатыми краями, глубина которых равна половине второй гармоники поверхностной акустической волны.

Конвольверы нашли применение в устройствах обработки радиосигналов для согласованной фильтрации, для получения корреляционных функций, для дискретного Фурье-преобразования в реальном масштабе времени, для кодирования радиосигналов.

2.3.3. Устройства памяти

Акустическое запоминающее устройство

Акустическое запоминающее устройство (АЗУ) предназначено для запоминания сигналов, например в процессе их корреляционной обработки. Операция записи информации осуществляется путем нелинейного взаимодействия акустической несущей волны с опорной акустической или электрической волной. При этом формируется пространственный рельеф из зарядов поверхностных или объемных ловушек. Информация хранится в виде фазовых и амплитудных распределений взаимодействующих волн. Операция восстановления сигнала осуществляется электрическими или звуковыми полями и сводится к "размораживанию" заполненного пространственного рельефа.

Различают импульсные и параметрические методы записи и считывания информации, а также метод их комбинаций.

Импульсный метод записи и хранения информации в АЗУ заключается в создании потенциального рельефа плотности зарядов в приповерхностном слое полупроводника с его последующей фиксацией. Возникновение потенциального рельефа обусловлено поперечным акустоэлектронным эффектом.

Под действием ПАВ происходит перераспределение зарядов в полупроводниковой пластине, находящейся вблизи поверхности звукопровода. В результате между верхним контактом к полупроводниковой пластине и металлическим электродом звукопровода возникает разность потенциалов, пропорциональная интенсивности ПАВ (рис. 2.27, а).

Таким образом, в полупроводниковой пластине принципиально можно записать "образ" волны, проходящей по звукопроводу. Существует несколько способов записи "образа" волны, все сводящиеся к созданию матрицы полупроводниковых диодов. В этом случае устройство АЗУ принимает вид, представленный на рис. 2.27, б. Фрагмент матрицы с диодами Шоттки приведен на рис. 2.27, в. Диоды Шоттки имеют время восстановления при прямом смещении 0,1—1,0 нс и позволяют записывать в АЗУ сигналы с полосой до нескольких сот мегагерц.

Матрица р—п-диодов, формируемых на основе переходов "монокристаллический кремний—поликристаллический кремний" (МК — ПК), представлена на рис. 2.27, г. При нулевом и обратном смещении диодов Шоттки время их восстановления лежит в пределах 10^{-3} — 10^{-1} с, и они не успевают реагировать на ВЧ-сигналы с частотой ~ 100 МГц. Весь модулируемый заряд находится в объеме кремния в слое толщиной 1 мкм.

При приложении импульса смещения к электроду подложки для смещения диодов в прямом направлении время восстановления становится малым, и диоды уже реагируют на ВЧ-сигналы.

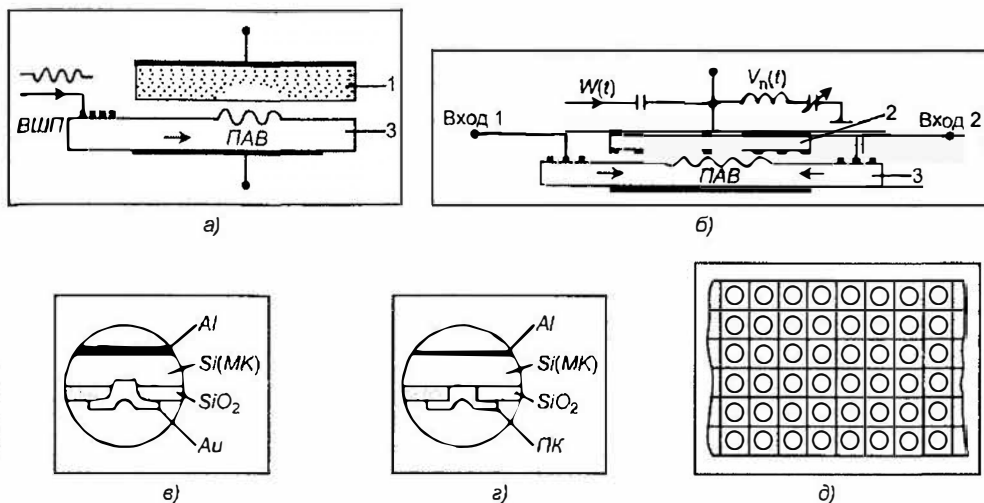


Рис. 2.27. Схема АЗУ: а — возникновение поперечного акустоэлектронного эффекта; б — устройства АЗУ; в — фрагмент с диодами Шоттки; г — фрагмент с диодами типа МК — ПК; д — фрагмент общего вида матрицы; 1 — полупроводниковая пластина; 2 — матрица диодов; 3 — пьезоэлектрическая пластина

Диоды заряжаются пропорционально локальной напряженности поля. Полный заряд матрицы формируется ПАВ и импульсом смещения и соответствующим образом меняется по поверхности. Матрица позволяет записывать и считывать информацию 10^5 раз. Итак, диодная матрица на поверхности звукопровода "фотографирует" потенциальный рельеф в приповерхностном слое полупроводника. Зафиксировать этот рельеф в диодной матрице можно при условии подачи сразу на все диоды короткого импульса отрицательного напряжения. Диоды откроются, сместятся в обратном направлении, ток мгновенно возрастет с напряжением из-за малого значения времени переключения, а диоды зарядятся в строгом соответствии с мгновенным значением электрического поля ПАВ, зафиксировав мгновенное распределение рельефа.

Если на входные ВШП подается сигнал $S(t) \cdot \exp(j\omega t)$, то возбуждается поперечное поле

$$E_1 \approx S(t - z/V_a) \exp(j\omega t - \vec{k} \cdot \vec{z}),$$

которое формирует распределение зарядов памяти, пропорциональное входному сигналу

$$Q_a(z) \sim S(-\frac{z}{V_a}) e^{i\vec{k} \cdot \vec{z}}.$$

Это выражение является функцией корреляции входного и записываемого сигналов. Для адекватного отображения электрического поля волны необходимо, во-первых, чтобы на длину акустической волны приходилось, как минимум, два диода (на каждую половину), и, во-вторых, чтобы длительность заряжающего импульса смещения на диоды была короче периода входного сигнала $t_3 \leq \frac{\pi}{\omega}$.

На рис. 2.28, а приведена дисперсионная диаграмма, соответствующая импульсному методу записи сигнала. Возбужденная сигналом $S(t)$ ПАВ (ω_1, \vec{k}_1) взаимодействует с одно-

родным ВЧ-полем $W(t)$ ($\omega_2, \vec{k}_2 = 0$) при условии $\omega_1 = \omega_2$ и формирует в матрице диодов "образ" в виде распределенных зарядов

$$Q_s(z) = e^{ikz} \int S(t - \frac{z}{V}) W(t) dt.$$

Пределы интегрирования определяются временным интервалом процесса.

Функция $Q_s(z)$ является функцией корреляции. При импульсном методе считывания на параметрический электрод подается короткий импульс, который мгновенно открывает диод и, разрушая заряд памяти, снимает локальные деформации поля. В пьезоэлектрике это равносильно возбуждению двух ПАВ, распространяющихся в противоположных направлениях, одна из которых является копией исходной волны, а другая инвертирована по времени (рис. 2.28, б). Если вместо короткого импульса на параметрический электрод подается считывающий сигнал вида $R(t)e^{i\omega t}$, то на левый вход придет ПАВ в виде сигнала корреляции

$$U_1(t) \approx e^{i\omega t} \int S(\tau - t) R(\tau) d\tau \approx e^{i\omega t} S \times R,$$

а на правый вход — сигнал свертки:

$$U_2(t) \approx e^{i\omega t} \int S(t - \tau) R(\tau) d\tau \approx e^{i\omega t} S \otimes R.$$

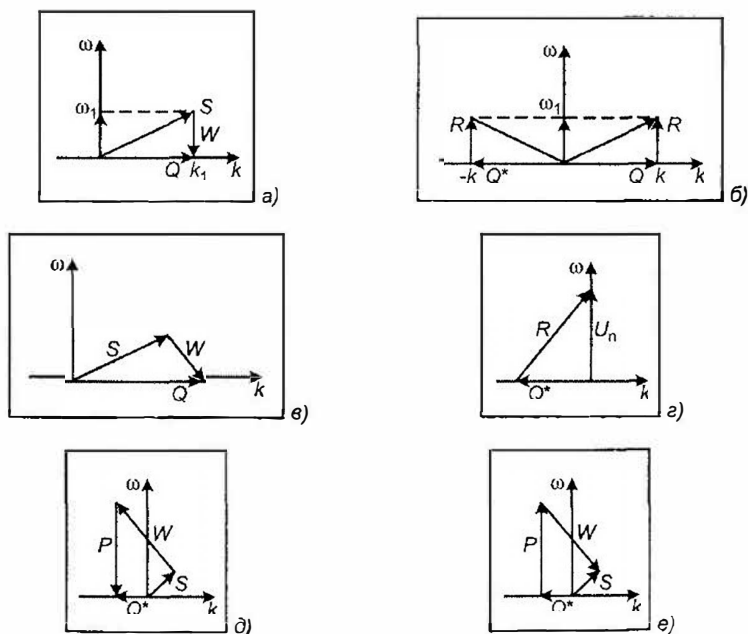


Рис. 2.28. Дисперсионные диаграммы, соответствующие импульсному методу записи (а) и считывания (б), параметрическому методу записи (в) и считывания (г), с сигналом накачки (д, е)

Пределы интегрирования определяются временем прохождения волнами области взаимодействия.

Параметрический метод записи и хранения информации также основан на нелинейном взаимодействии волн. Сигнал памяти получается при взаимодействии ПАВ, формируемых сигналами $S(t)$ со входа 1 и $W(t)$ со входа 2. Возникающая в результате нелинейного взаимодействия составляющая электрического тока формирует соответствующую зарядку диодов dQ_s , описываемую выражением:

$$Q_s(z) \cong e^{2jkz} \int S\left(t - \frac{z}{V_a}\right) W\left(t - \frac{L-z}{V_a}\right) dt.$$

Это выражение отражает функцию корреляции сигналов $S(t)$ и $W(t)$. Для записи всех спектральных составляющих сигнала $S \pm \Delta S$ необходимо найти компромисс между длительностью и шириной полосы обрабатываемых в АЗУ сигналов. Дисперсионная диаграмма параметрического метода записи приведена на рис. 2.28, в.

Метод параметрического считывания основан на нелинейном смешивании в полупроводнике "замороженных" полей статического заряда памяти с полем считывающего импульса.

Если на вход 1 подать сигнал $R(t - z/V) \exp(j\omega t - kz)$, то на параметрическом электроде получим сигнал корреляции входного и считывающего сигналов:

$$U_n(t) \approx e^{j\omega t} \int Q_s(z) R\left(t - \frac{z}{V}\right) dz \approx e^{j\omega t} S \times R.$$

Если же на вход 2 подать этот же сигнал $R(t)$, то на параметрическом электроде получим сигнал, пропорциональный сигналу свертки $e^{j\omega t} S * R$. При методе параметрической записи и считывания может быть получена функция корреляции. Дисперсионная диаграмма процесса параметрического считывания приведена на рис. 2.28, г. В ряде функциональных устройств по обработке сложных радиосигналов на базе АЗУ могут быть применены импульсные методы записи с параметрическим считыванием и наоборот.

Заметим, что рассмотренные трехволновые взаимодействия в АЗУ подобны процессам записи и воспроизведения информации в голографии. В обоих случаях в наличии опорная и предметная волны. В АЗУ роль опорной волны выполняет либо ВЧ-сигнал на параметрическом электроде или импульсная ПАВ на одном из входов или параметрическом электроде. Записанный сигнал, как и в голографии, несет информацию об амплитудном и фазовом распределении входного сигнала. Восстановление исходного сигнала и есть процесс считывания, роль фотоэмulsionи отводится матрице полупроводниковых диодов. Другими словами, процессы записи, хранения и считывания информации в АЗУ представляются как голография динамических неоднородностей. Причем динамические неоднородности представляют собой либо поверхностную акустическую волну, либо ВЧ-электромагнитный сигнал. Как и в голографии, каждый бит информации "размазывается" по всей диодной матрице и запоминается тысячами диодов. Поэтому надежность АЗУ достаточно высока в отличие от ЗУ, реализованными схемотехническими методами.

Все рассмотренные процессы относятся к категории трехволновых. Реализованы методы четырехволнового взаимодействия с использованием сигнала накачки $P(t)$. Сигнал $P(t)$ подается на параметрический электрод с частотой

$$\omega_n = \omega_1 \pm \omega_2,$$

где ω_1 и ω_2 — частоты двух взаимодействующих ПАВ ($\omega_1 \neq \omega_2$). В этом случае распределение заряда памяти будет иметь вид "тройной корреляции":

$$Q_s(z) = \int_{-T/2}^{T/2} S(t - \frac{z}{V_a}) W(t - \frac{L-z}{V_a}) P(t) dt,$$

где $T = L / V_a$ — время интегрирования.

Схема записи приведена на дисперсионной диаграмме (рис. 2.28, д). Хранение осуществляется аналогично — в виде распределенного заряда. Считывание записанной информации можно провести с помощью генерации обратной волны при накачке (рис. 2.28, е). Можно считать информацию путем сканирования двумя встречными ПАВ. В обоих случаях входной сигнал представляет собой тройную свертку или корреляцию. Метод тройной свертки находит применение в гидролокации, являясь более дешевым по сравнению с цифровыми методами. Акустические ЗУ находят широкое применение в различных устройствах для обработки сложных радиосигналов, например, для создания линий задержек, когда требуются большие времена задержки сигнала.

Основное же применение АЗУ нашли в качестве корреляторов согласованной фильтрации со встроенной памятью в радиолокационных системах. Устройства с АЗУ использовались для сжатия ЛЧМ-сигналов большой длительности и доплеровской обработки ясных радиолокационных сигналов.

Технико-экономические исследования показывают, что используемые конструкции и устройства на основе АЗУ целесообразно применять, если их типовые характеристики лежат в следующих пределах: емкость памяти $\sim 10^5$ бит, длительность выборки информации $\sim 2 \times 10^{-4}$ с, скорость ввода-вывода информации $\sim 10^8$ бит/с, время хранения информации $\sim 10^{-1}$ с, а плотность хранения $10^5 \div 10^6$ бит/см².

Характеристики АЗУ открывают широкие перспективы их использования в системах обработки как аналоговых, так и цифровых сигналов.

Экзотические устройства акустической памяти

Рассмотренные динамические неоднородности акустической природы позволяют прогнозировать появление новых экзотических типов памяти.

Одним из таких устройств хранения информации может стать память на основе фононного (электроакустического) эха. Можно реализовать режим динамической памяти на двухимпульсном эхо, а также режим квазистатической (долговременной) памяти на основе трехимпульсного эхо. Такие устройства найдут применение в радиотехнических устройствах, поскольку работают они в диапазоне частот до 10 ГГц, а время хранения сигнала достигает 10^6 с.

Перспективным направлением развития акустической памяти являются ферроакустические устройства, в которых используется нелинейное взаимодействие динамических неоднородностей акустической и магнитной природы в континуальных средах типа металлических ферромагнетиков. Взаимодействие между динамическими неоднородностями позволяет создать в континуальной среде пространственный рельеф остаточной намагниченности, соответствующий сигналу. Применение двух типов физических носителей позволяет уменьшить помехи. Считывание информации производится путем обратимого изменения намагниченности при распространении механического напряжения в акустической волне. Ферроакустические ЗУ позволяют обрабатывать цифровые и аналоговые сигналы. Плотность записи информации достигает значения 10^{12} бит/см³. Ферроакустические устройства памяти могут работать в условиях мощных дестабилизирующих факторов, они высоконадежны и технологичны.

2.3.4. Фурье-процессоры

Акустоэлектронные *Фурье-процессоры* (АЭФП) представляют собой аналоговые преобразователи, использующие акустоэлектронные процессоры для вычисления преобразования Фурье дискретного

$$S_k = \sum_{n=0}^{N-1} \bar{S}_k \exp\left(j \frac{2\pi nk}{N}\right),$$

где $k = 0, 1, \dots, N-1$,

и непрерывного (интегрального) типов

$$S(\omega) = \int_{-\infty}^{+\infty} H\left(\frac{t}{T_0}\right) S''(t) \exp(-j\omega t) dt,$$

где $H(t/T_0)$ — прямоугольная функция окна шириной T_0 с центром $t = 0$.

Различают параллельные и интерференционные процессоры, а также процессоры на алгоритмах ЛЧМЗ-преобразований.

Резонансные (фильтровые) ФП характеризуются высоким быстродействием и позволяют обеспечить высокую вероятность обнаружения сигналов в реальном масштабе времени. Схемное и конструктивное решение этого вида ФП основывается на акустоэлектронных устройствах частотной селекции, в частности на фильтрах ПАВ с линейной фазовой характеристикой. Набор ПАВ-фильтров с собственными частотами ω , позволяет произвести анализ спектральной плотности исследуемого сигнала в дискретной форме.

Интерференционный ФП представляет собой одноканальное рециркуляционное устройство и предназначен для формирования и измерения комплексного спектра сигналов в реальном масштабе времени с полосой анализа до десяти килогерц. Структурно этот тип ФП состоит из акустоэлектронной линии задержки и технологически совместимых усилителей, смесителя, генератора, полосового фильтра, выполненных в микроэлектронном исполнении. При воздействии на вход интерференционного ФП гармонического сигнала $S(t) = \cos(2\pi f_c t + \varphi)$ задержанный в рециркуляторе сигнал сдвигается по частоте на $f_{\text{сдв}} = 1/\tau$. Происходит синфазное накопление сигнала в момент времени

$$t = F\tau_s / B + q\tau,$$

соответствующий неизвестной частоте

$$f_c = K/\tau_s + F,$$

где F — частоты, q — число циркуляции, τ_s — время запаздывания, $T_a = q\tau_s$.

Условие синфазности выполняется для всех частот $F \in [K/\tau_s(K+1)\tau_s]$. Отклик интерференционного ФП после q циркуляций в интервале $\tau \in [T_a; (T_a + \tau_s)]$ имеет вид радиоимпульса $U_q(\tau)$, огибающая и фаза отклика которого при больших q соответствует амплитудно- и фазочастотным спектрам входного сигнала $U_q(\tau) \equiv S(\omega)$.

Фурье-преобразователи на основе ЛЧМЗ-преобразований представляют особый интерес в силу относительной простоты реализации, технологичности, малых габаритов, высокого быстродействия и широких функциональных возможностей.

Этот тип Фурье-преобразователей имеет алгоритм ЛЧМЗ- и ЛЧМ-преобразований, позволяющий вычислить дискретное преобразование Фурье (ДПФ) и интегральное преобразование Фурье (ИПФ) путем использования модифицированных уравнений в виде:

$$S_k = \exp\left(-j \frac{\pi k^2}{N}\right) \sum_{n=0}^{N-1} S_n'' \exp\left(-j \frac{\pi n^2}{N}\right) \exp\left(j \frac{\pi(k-n)^2}{N}\right),$$

$$S(\omega) = S(\beta t) = \exp\left(-j \frac{\beta t^2}{2}\right) \int_{-\infty}^{+\infty} S(\tau) \exp\left(-j \frac{\beta \tau^2}{2}\right) \exp\left(j \frac{\beta(t-\tau)^2}{2}\right) d\tau,$$

где $\beta = \frac{\omega}{t}$.

В этом случае вычисление спектра сводится к следующим операциям:

1. Перемножение выборок сигнала S_n или непрерывного сигнала $\bar{S}(t)$ во временной области с ограниченным во времени сигналом $\bar{U}_{21} = \exp(j_2 \Theta)$ с линейно-частотной модуляцией.
2. Свертка полученного результата в фильтре, отклик которого $\bar{U}_\phi = \exp(j \Theta)$.
3. Перемножение входного сигнала фильтра с ЛЧМ-сигналом $\bar{U}_{22} = \exp(-j \Theta)$ в частотной области.

Для алгоритма ДПФ величина $\Theta = \frac{\pi n^2}{N}$, а для алгоритма ИПФ $\Theta = \frac{\beta t^2}{2}$.

Алгоритм может быть записан в виде:

$$S(\omega) = U_{22} \{U_\phi^+ * [S''(t) U_{21}^-]\},$$

и носит название перемножение — свертка — перемножение (П — С — П). Знаками "+" и "-" обозначены положительный и отрицательный наклоны характеристик группового времени запаздывания устройства.

Преобразование Фурье можно получить по схеме С — П — С с иной последовательностью выполнения операций:

$$S(\omega) = \{[S(t) * U_{\phi 1}^+]\} * U_{\phi 2}^+.$$

Аналоговые АЭФП могут использоваться и для получения обратного преобразования Фурье (ОПФ) и применением схем П — С — П и С — П — С. Величины $U_2(t)$ и $U_\phi(t)$ представляют собой импульсные отклики ЛЧМ, соответственно генератора и фильтра. АЭФП формируются на уже рассмотренных процессорах: дисперсионных линиях задержки, конвольверах, генераторах ЛЧМ-сигналов и т. п.

На примере АЭФП можно проследить преимущества устройств функциональной электроники для обработки информации.

Многие задачи обработки сигналов решаются цифровыми процессами, позволяющими осуществлять быстрое преобразование Фурье (БПФ). Это сложные и энергопотребляющие устройства. Разработаны и комплексы БИС, позволяющие осуществлять БПФ. Они уникальны и дороги. В любом случае цифровые процессоры имеют характеристики, определяемые точностью и динамическим диапазоном входящих в них аналогово-цифровых преобразователей. Использование в вычислительном модуле акустоэлектронных устройств позволяет упростить устройство, повысить его быстродействие и надежность.

Сравнительные характеристики показывают, что на базе АЭФП можно создать электронно-вычислительный модуль (ВМ) с 2048 точками преобразования, цифровой аналог

(ЦА) — 2048; скорость обработки информации в ВМ — 250 МГц, в ЦА — 250 Мбайт/с; выходная точность ВМ составляет ~ 1%, а в ЦА — 7 разрядов.

Заметим, что АЭФП выполняют операции, которые можно отнести к функциям высшего порядка. Перечислим основные из них:

- ☐ синтез частот в диапазоне 500 МГц с малым временем переключения при смене частот;
- ☐ демодуляция частотно-манипулированных сигналов;
- ☐ спектральная обработка сигналов с большой фазой;
- ☐ программируемая согласованная фильтрация;
- ☐ преобразование временного масштаба;
- ☐ центральная обработка сигналов — нелинейная обработка для детального исследования сигналов;
- ☐ интегральные преобразования Фурье, Гильберта, Френеля, Лапласа, Меллина.

Все это свидетельствует о перспективности разработки и использования акустоэлектронных Фурье-преобразователей.

Анализ показал, что акустоэлектронные процессоры являются ярким примером интеграции функций в приборах и устройствах обработки информации.

2.4. Акустоэлектроника в системах и средствах связи

В приспосабливаемых устройствах и системах связи в диапазоне частот от 1 МГц до 10 ГГц широко применяются акустоэлектронные устройства. К таким устройствам относятся спутниковое, кабельное, цифровое, сотовое телевидение, телевидение высокой четкости, а также подвижные, спутниковые, тропосферные и радиорелейные линии связи.

Условно можно отметить четыре группы акустоэлектронных устройств, применяемых в системах и устройствах связи.

1. Дуплексор для абонентских станций подвижных систем связи; резонаторные фильтры для односторонних и двусторонних пейджеров; резонаторы для радиоиндексации автомобилей; резонаторные фильтры для защиты от несанкционированного доступа.
2. Фильтры с малым вносимым затуханием для оконечных каскадов приемопередающих устройств систем и средств связи; фильтры промежуточной частоты с малым вносимым затуханием; линии задержки с малым вносимым затуханием для маломощных приемников.
3. Фильтры для цифрового телевидения и цифровой радиосвязи; дисперсионные линии задержки; линии задержки для кодированного разделения каналов.
4. Конвольверы для широкополосных систем и средств связи.

Следует отметить, что акустоэлектронные устройства на ПАВ-приборах имеют перспективу улучшения массогабаритных и энергетических характеристик. Намечались следующие пути совершенствования акустоэлектронных устройств.

В области материалов — это использование тетрабората лития лангасита — новых перспективных материалов.

В области технологии:

- ☐ развитие технологии изготовления термокомпенсированных пленочных структур на пьезокристаллах с высоким коэффициентом электромеханической связи;
- ☐ переход на технологию получения субмикронных топологических структур;
- ☐ развитие технологии получения интегральных слоистых на основе алмазоподобных пленок;
- ☐ разработка методов автоматизированного проектирования акустоэлектронных устройств на основе новых технологий.

Контрольные вопросы

1. Что такое функциональная акустоэлектроника?
2. Какие физические эффекты относятся к акустоэлектронным?
3. Какие динамические неоднородности акустоэлектронной природы вы знаете?
4. Каким способом можно возбудить ПАВ?
5. Как связана частота ПАВ с топологией ВПП?
6. Как осуществляется управление ПАВ в акустическом канале?
7. Как устроена линия задержки на ПАВ?
8. Какие типы линий задержки на ПАВ вы знаете?
9. Опишите конструкцию и работу дисперсионной линии задержки.
10. Какие устройства частотной селекции вы знаете?
11. Как устроен фильтр на ПАВ-резонаторах?
12. Как устроен дисперсионный фильтр и каковы его характеристики?
13. Какие физические явления положены в основу работы усилителя на ПАВ?
14. Какие типы нелинейного взаимодействия вы знаете? Объясните на дисперсных диаграммах.
15. Опишите конструкцию и принципы работы конвольвера.
16. Как устроено ЗУ на ПАВ?
17. Что такое Фурье-процессор и как он устроен?

Рекомендуемая литература

1. Гуляев Ю. В. Акустоэлектронные устройства для систем связи и обработки информации. В кн. Проблемы современной радиотехники и электроники. Под ред. акад. В. А. Котельникова. — М.: Наука, 1980.
2. Кочемасов В. Н., Долбня Е. В., Соболев П. В. Акустоэлектронные Фурье-процессоры. — М.: Радио и связь, 1987.
3. Морган Д. Устройства обработки сигналов на поверхностных акустических волнах. — М.: Радио и связь, 1990.
4. Речинский В. И. Акустоэлектронные компоненты. — М.: Радио и связь, 1987.
5. Щука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

3. Функциональная диэлектрическая электроника

3.1. Физические основы

Функциональная *диэлектрическая* электроника представляет собой направление в функциональной электронике, в котором изучаются явления и эффекты в активных диэлектриках, а также возможность создания приборов и устройств обработки и хранения информации на основе динамических неоднородностей электрической, магнитной или электромагнитной природы.

3.1.1. Динамические неоднородности

В активных диэлектриках в основном используются динамические неоднородности различной физической природы: домены, квазичастицы, неоднородности фоторефрактивной и электрооптической природы.

Особый интерес представляют сегнетоэлектрические домены. Они представляют собой области однородной спонтанной поляризации P_S (рис. 3.1, а). Размеры доменов составляют несколько микрон и разделены переходной областью или доменной границей толщиной в несколько ангстрем (рис. 3.1, б). Поляризация P зависит от напряженности электрического поля E нелинейным образом. При циклическом изменении E от $+E_K$ до $-E_K$ вектор поляризации меняется от $+P_S$ до $-P_S$, где E_K — коэрцитивная сила. Функция $P_S = f(E_K)$ имеет вид петли гистерезиса (рис. 3.1, в). Петля гистерезиса является отличительной чертой динамических неоднородностей сегнетоэлектрической природы. Время электронной переполяризации составляет $\sim 10^{-15}$ с.

Спонтанной поляризацией обладают и пьезоэлектрики, но направление вектора P_S в них строго фиксировано и поэтому не обнаружено динамических неоднородностей типа доменов.

В сегнетоэлектриках-сегнетоэластиках наблюдается спонтанная локальная деформация в результате изменения внешних условий. Спонтанная деформация сопровождается изменением знака при воздействии электрического поля.

В сегнетоэлектриках наблюдаются фотостимулированные фазовые переходы 1-го и 2-го рода, при которых генерируются *фазоны* — макроскопические квазичастицы. Фазоны представляют собой возбуждение в гетерофазной системе из смеси фазовых состояний одного и того же вещества, связанной с образованием вокруг заряженной частицы флуктуации концентрации одной из компонент системы, создающей для частицы потенциальную яму. Самосогласованный радиус фазона R_0 при данной температуре и заданной плотности поверхностной энергии пропорционален числу локализованных электронов n : $R_0 \sim n^{2/3}$ ($2R_0 \approx 10$ Å). Фазоны можно рассматривать как подвижные пакеты квазичастиц с эффективной массой M . Модель фазона предполагает существенное влияние фазового перехода 1-го рода на электропроводность. Такие пакеты квазичастиц — фазонов — можно, по-видимому, рассматривать как динамические неоднородности в сегнетоэлектриках.

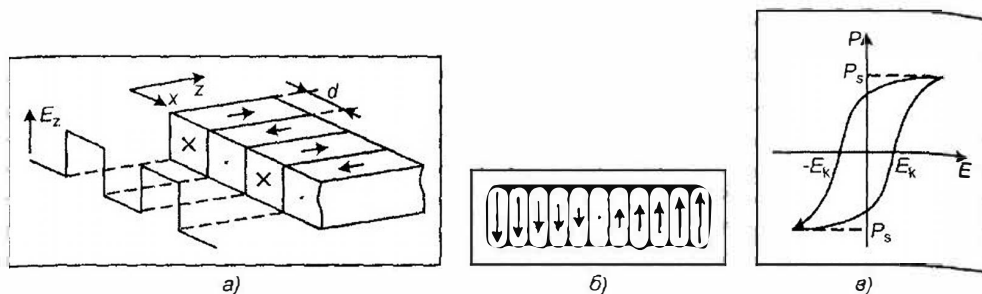


Рис. 3.1. Слоистая доменная структура и распределение в ней полей (а), структура сегнетоэлектрической доменной стенки (б), петля гистерезиса переполаризации домена (в)

В сегнетоэлектриках же в области фотостимулированных фазовых переходов типа "порядок—беспорядок", образуются макроскопические квазичастицы — *флуктоны*, которые тоже можно рассматривать в качестве динамических неоднородностей. Флуктоны возбуждаются в гетерофазной системе и представляют собой флуктуации концентрации одной из компонент системы, создающие для частицы потенциальную яму, перемещающуюся вместе с частицей.

В сегнетомагнитных кристаллах одновременно существуют магнитная и электрическая дипольные структуры и соответственно векторы спонтанной поляризации P_s и спонтанной намагнитченности M_s . В элементарных ячейках кристаллографических структур сегнетомагнетиков должны находиться частицы, обладающие как электрическим дипольным моментом, так и магнитным. Сосуществование электрических и магнитных дипольных структур определяется принципом симметрии в различных кристаллографических структурах. Вещества определенной симметрии обладают магнитоэлектрическим эффектом. Он состоит в том, что при наложении электрического поля появляется пропорциональная полю намагнитченность, а при наложении магнитного поля появляется электрическая поляризация, пропорциональная этому полю. В общем случае для таких сред справедливы соотношения:

$$D = \varepsilon E + vH, \quad B = \mu H + v'E,$$

где v и v' — магнитоэлектрические тензоры, ε — диэлектрическая проницаемость, μ — магнитная проницаемость вещества.

Представляют интерес для их использования в качестве динамических неоднородностей *экситоны* и *поляритоны*. Экситоны представляют собой электронное возбуждение в кристалле диэлектрика или полупроводника, распространяющееся по кристаллу в виде волны, но не связанное с передачей электрического заряда или массы. При освещении кристалла могут возникнуть поляритоны, представляющие собой квантовые квазичастицы, возникающие при взаимодействии экситона и оптического фонона с фотоном. Эти экзотические квазичастицы способны переносить информационный сигнал и могут быть использованы в устройствах функциональной электроники.

3.1.2. Континуальные среды

В качестве континуальных сред в функциональной диэлектрической электронике используются, как правило, активные диэлектрики. Под активными диэлектриками будем понимать диэлектрические материалы, в которых могут быть возбуждены динамические неод-

народности и которые предназначены для процессов генерации, усиления, преобразования и хранения информационных сигналов.

Активные диэлектрики существенно отличаются от пассивных, которые применяются в основном в качестве электроизоляционных материалов.

Свойства активных диэлектриков зависят от воздействия на них определенных физических полей: механических, тепловых, электрических, магнитных. Результатом взаимодействия структуры диэлектрика и полей является генерация динамических неоднородностей определенной физической природы. Активные диэлектрики, используемые в качестве континуальных сред в приборах и устройствах функциональной диэлектрической электроники, обладают широким набором свойств, определяемых типом воздействующих физических полей: механических, тепловых, электрических, магнитных, оптических.

В результате анализа континуальных сред удалось систематизировать свойства диэлектриков в зависимости от воздействующих на них физических полей.

На рис. 3.2 схематично представлены информационные поля, отражающие различные физические взаимодействия в активных диэлектриках. В соответствии с принятой системой классификации пронумеруем различные свойства активных диэлектриков в зависимости от типов информационных полей.

Область, ограниченная координатами 1.1—2.1, 1.3—2.1, 1.1—2.3, 1.3—2.3 представляет собой термоупругие свойства диэлектриков.

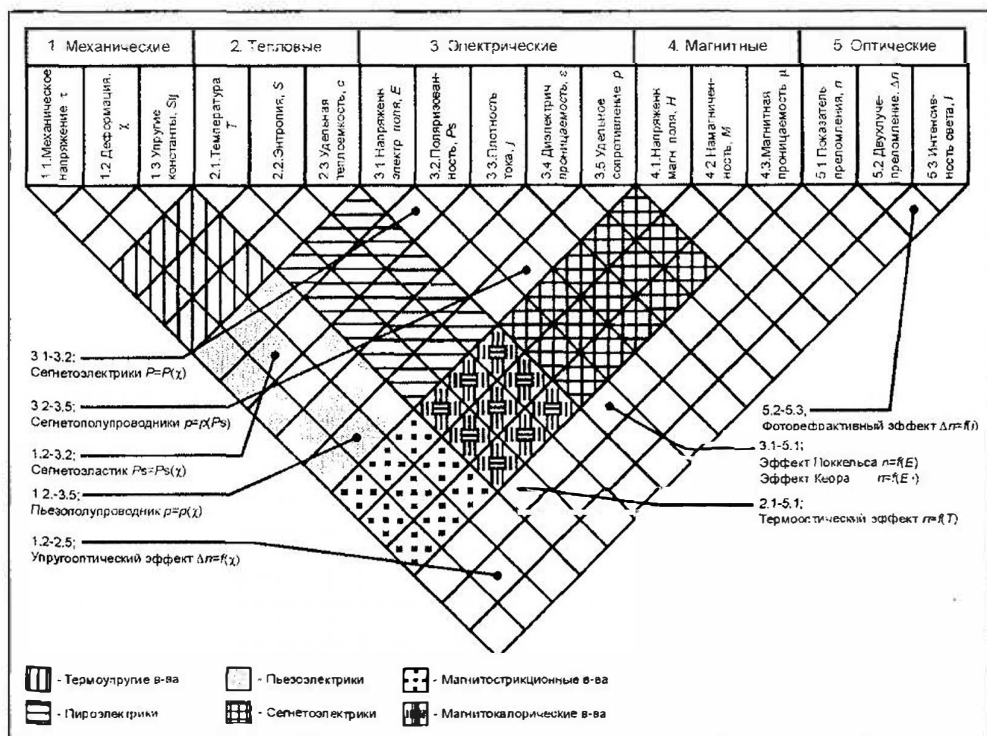


Рис. 3.2. Системный анализ свойств активных диэлектриков

Пьезоэлектрики располагаются в области, ограниченной координатами 1.1—3.1, 1.3—3.1, 1.1—3.5, 1.3—3.5. Их основным свойством является наличие поляризации, направление и уровень которых могут быть изменены при внешних воздействиях. Различают прямой пьезоэлектрический эффект, заключающийся в поляризованности P диэлектрика в зависимости от механического напряжения σ , описываемый соотношением

$$P = d\sigma,$$

где d — пьезомодуль.

В обратном пьезоэлектрическом эффекте деформация χ линейно зависит от поляризованности P в соответствии с соотношением

$$\chi = gP,$$

где

$$g = \frac{d}{\epsilon_0(\epsilon - 1)},$$

здесь ϵ — диэлектрическая проницаемость диэлектрика, ϵ_0 — константа. Заметим, что пьезоэлектрические материалы анизотропны. Их пьезоэлектрические, диэлектрические и упругие свойства описываются набором компонент пьезомодулей d_{ij} , упругих констант S_{ij} , диэлектрических проницаемостей ϵ_{ij} по разным направлениям.

Каждая ячейка описываемой информационной области обладает рядом замечательных свойств и определяет тип пьезоэлектриков. Сегнетоэластики занимают ячейку с координатами 1.2—3.2 и описываются зависимостью $P = P(\chi)$. Пьезополупроводники располагаются в ячейках с координатами 1.2—3.5 и, как правило, представляют собой пленки CdS, ZnS, ZnO с низким значением ϵ (< 10) и с выраженной зависимостью $\rho = \rho(\chi)$.

Пирозлектрики представляют собой класс активных диэлектриков, в которых происходит поляризация при однородном по объему изменению температуры. Этот класс располагается в информационной области с координатами 2.1—3.1; 2.3—3.1; 2.1—3.5; 2.3—3.5. Спонтанная или остаточная поляризация P зависит от температуры $P = P(T)$, и для количественного описания вводят *пирокоэффициент* p , определяемый выражением

$$p = \frac{dP}{dT}.$$

Все пирозлектрики обладают обратным электрокалорическим эффектом, потому что их температура изменяется при поляризации.

Особый интерес для функциональной электроники представляют собой сегнетоэлектрические континуальные среды, имеющие доменную структуру. Каждый домен обладает спонтанной поляризацией, уровень и направление которых могут быть изменены внешними полями, например электрическим. В отличие от пирозлектриков, у которых направление вектора поляризации P строго фиксировано, у сегнетоэлектриков суммарная поляризация зависит от поляризации отдельных доменов. На рис. 3.2 сегнетоэлектрики занимают область с координатами 3.1—3.2.

В зависимости от вида поляризации различают ионные, дипольные и несобственные сегнетоэлектрики. В *ионных сегнетоэлектриках* ячейка имеет вид куба типа ABO_3 , по вершинам которого расположены ионы типа $A(Ba^{2+}, Pb^{2+}, K^{2+}, \dots)$, в центре ионы типа $B(Ti^{4+}, Nb^{4+})$, а в центрах граней ионы кислорода O^{2-} . Если ионы типа B находятся в центре элементарной ячейки (рис. 3.3, а), то центры тяжести положительных и отрицательных ионов совпадают и общий дипольный момент равен нулю. Смещение иона $B(Ti^{4+}, \text{напри-})$

мер) из центра кислородного октаэдра приводит к разбалансировке и возникновению дипольного момента P , направленного в сторону смещения (рис. 3.3, б).

Спонтанная поляризация в сегнетоэлектриках такого типа является следствием упорядоченного смещения ионов и представляет собой фазовый переход типа смещения.

К ионным сегнетоэлектрикам относятся структуры типа перовскита (CaTiO_3), титанат бария (BaTiO_3), титанат свинца (PbTiO_3), ниобат калия (KNbO_3) и т. п.

В элементарной ячейке дипольных сегнетоэлектриков содержатся атомы с двумя положениями равновесия, электрическая поляризация которой и определяется взаимным положением атомов. На рис. 3.3, в представлено взаимное расположение ионов при положительном и отрицательном направлениях дипольного момента (рис. 3.3, г). Фазовый переход в дипольных сегнетоэлектриках, происходящий в точке Кюри, называется переходом типа "порядок—беспорядок".

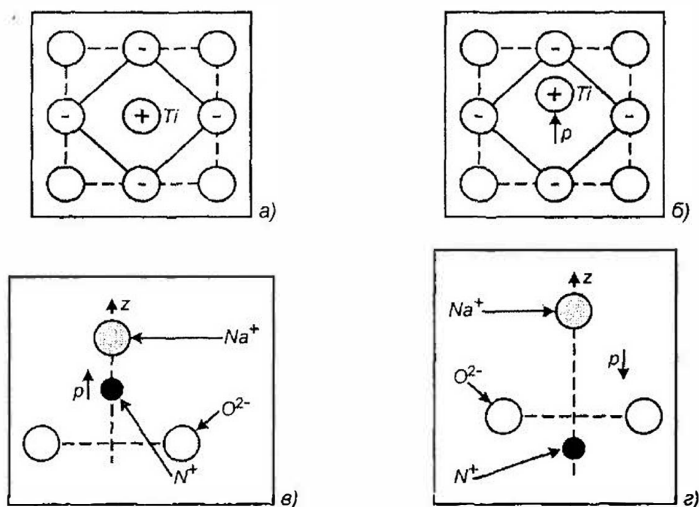


Рис. 3.3. Элементарные ячейки сегнетоэлектриков:
ионный тип в неполяризованном (а) и поляризованном (б) состояниях;
дипольный тип при положительном (в) и отрицательном (г) дипольных моментах

К дипольным сегнетоэлектрикам относятся сегнетова соль ($\text{NaKC}_4\text{H}_4\text{O}_6 \cdot 4\text{H}_2\text{O}$), триглицинсульфат ТГС ($\text{NH}_2\text{CH}_2\text{COOH}$) $_3 \cdot \text{H}_2\text{SO}_4$, нитрид натрия (NaNO_2) и т. п.

К особой группе относятся несобственные сегнетоэлектрики, в которых спонтанная поляризация обусловлена, например, деформацией кристаллической решетки при фазовом переходе. К этому типу относятся молибдат гадолиния $[\text{Cd}_2(\text{MoO}_4)_3]$, лангбейниты, $[\text{Ti}_2\text{Cd}_2(\text{SO}_4)_3]$, фторбериллат аммония $[(\text{NH}_4)_2\text{BeF}]$ и т. п.

Кристаллы, в соседних ячейках которых одноименные ионы смещены в противоположных направлениях, называются *антисегнетоэлектриками*. Спонтанная поляризованность антисегнетоэлектриков равна нулю. При наложении сильного электрического поля антисегнетоэлектрик может переходить в сегнетоэлектрик с параллельной ориентацией диполей. Отличительной особенностью антисегнетоэлектриков является наличие двойных петель диэлектрического гистерезиса.

К антисегнетоэлектрикам относятся кристаллы цирконата свинца (PbZnO_3), ниобата натрия (NaNbO_3), гафната свинца (PbHfO_3), дигидрофосфата аммония ADP ($\text{NH}_4\text{H}_2\text{PO}_4$) и т. п. Сегнетоэлектрики обладают довольно большим удельным сопротивлением ($\rho > 10^9$ Ом·м). Однако существуют такие сегнетоэлектрики, у которых значение ρ соответствует полупроводникам. Можно искусственно снизить значение ρ , например, ионным легированием и получить новый тип сегнетоэлектрика — *сегнетополупроводник*. К ним относятся вещества BiFeO_3 , PbTiO_3 и др. Эти вещества занимают ячейку с координатами 3.2—3.5 (рис. 3.2).

Весьма перспективными являются полимерные сегнетоэлектрические пленки, нанесенные на поверхность кремния. Пленки наносятся методом литья из раствора винилиденфторида $\text{CH}_2=\text{CF}_2$ и трифторэтилена $\text{CH}_2=\text{F}_2-\text{CHF}-\text{CF}_2$ в ацетоне с последующим осаждением при температуре + 20...22 °С. Толщина пленок лежит в пределах от 6 до 20 мкм.

На пересечении информационных полей электрических и магнитных свойств диэлектриков (см. рис. 3.2) располагаются *сегнетомагнитные кристаллы*. Отличительной особенностью структуры сегнетоэлектриков является сосуществующие магнитная и электрическая взаимопроникающие решетки. Они образованы частицами, несущими электрические и магнитные дипольные моменты.

В диэлектриках этого класса обнаружен магнитоэлектрический эффект. Его сущность состоит в том, что в веществах определенной симметрии при наложении электрического поля появляется намагниченность, пропорциональная полю, и при приложении магнитного поля появляется электрическая составляющая, пропорциональная полю.

Сегнетомагнетики перспективны для устройств функциональной электроники второго поколения, где возможна интеграция эффектов и явлений за счет присутствия в среде спонтанной поляризации P_s и спонтанной намагниченности M_s . К сегнетомагнетикам прежде всего относятся перовскитные соединения типа ферровольфрамат $[\text{Pb}(\text{Fe}_{2/3}\text{W}_{1/3})\text{O}_3]$ и ферроннобат свинца $[\text{Pb}(\text{Fe}_{1/2}\text{Nb}_{1/2})\text{O}_3]$, а также феррит висмута (BiFeO_3), ренийсодержащие перовскиты ($\text{Sr}_2\text{CoReO}_6$, SrNiReO_6), сегнетомагнитные борациды ($\text{Mg}_3\text{B}_7\text{O}_{13}\text{Cl}$) и т. п.

Континуальные среды, соответствующие ячейкам информационного поля с координатами 1.1—4.1, 1.3—4.1, 1.3—4.3, 1.1—4.3 (см. рис. 3.2), обладают магнитострикционными свойствами, а расположенные в информационном поле 2.1—4.1, 2.3—4.1, 2.1—4.3, 1.1—4.3 — магнитокалорическими.

Менее изучены свойства континуальных сред на пересечении оптических взаимодействий с механическими, тепловыми, электрическими и магнитными.

В ячейках с координатами 1.2—5.1 и 1.2—5.2 среды обладают упругооптическим эффектом, с координатами 2.1—5.1 и 2.1—5.2 — термооптическим эффектом (см. рис. 3.2).

Свойства сред, расположенных в ячейке 3.1—5.1, описываются эффектом Поккельса — $n = f(E)$ и эффектом Керра — $n = f(E^2)$. Среда, обладающая фоторефрактивным эффектом $\Delta n = f(I)$, расположена в ячейках с координатами 5.1—5.3. Проведенный системный анализ диэлектрических континуальных сред позволяет более эффективно осуществлять их выбор для перспективных процессоров и запоминающих устройств.

3.1.3. Генераторы динамических неоднородностей

Генерирование динамических неоднородностей в диэлектрических континуальных средах осуществляется физическими полями. Динамические неоднородности в устройствах

диэлектрической электроники, как правило, не перемещаются в континуальной среде. Информационный сигнал передается за счет изменения локальных физических свойств среды, прежде всего за счет локальных полей. При приложении электрического поля с помощью статических неоднородностей соответствующей конфигурации к сегнетоэлектрику в нем генерируется неоднородность в виде домена определенной поляризации. При приложении электромагнитного поля к сегнетоэлектрикам (например, KDP, ADP и др.) можно получить динамические неоднородности, изменяющие оптическую плотность среды и способствующие линейному преобразованию падающего излучения. В любом случае реализуется элемент конструкции — *генератор динамических неоднородностей*.

3.1.4. Другие элементы приборов

В приборах и устройствах диэлектрической электроники управление динамическими неоднородностями, как правило, носит статический характер и осуществляется с помощью меняющихся физических полей, в зависимости от типа выбранной динамической неоднородности.

Детектором динамических неоднородностей является, как правило, локальная область, изменяющая свои свойства под действием физических полей. Считывание информации можно осуществлять и сканирующим детектором. В этом случае имеет смысл говорить о распределенном детекторе.

3.2. Приборы и устройства функциональной диэлектрической электроники

3.2.1. Слоистые структуры

В приборах диэлектрической функциональной электроники используются, как правило, *слоистые структуры*. Слой активного диэлектрика применяется для хранения или обработки информации, а ввод и детектирование информации осуществляется в других слоях с соответствующими статическими неоднородностями.

Например, в структурах "сегнетоэлектрик—полупроводник" используется эффект поля. В этом случае свойства полупроводника определяются состоянием поляризации сегнетоэлектрика.

Если сегнетоэлектрик деполаризован ($P = 0$), то на границе раздела "сегнетоэлектрик—полупроводник" n -типа нет изгиба зон (рис. 3.4, а). Собственный уровень Ферми $E_{\text{ск}}$ представляет собой прямую линию, а основные носители смешаны в полупроводнике с неосновными. При поляризации сегнетоэлектрика ($-P$) у границы раздела в полупроводниковом слое образуется слой накопления основных носителей. Происходит изгиб зон, включая уровень Ферми, и при этом поверхностная плотность больше объемной плотности (рис. 3.4, б).

Если переполаризовать сегнетоэлектрик ($+P$), то у границы раздела в полупроводнике образуется слой обеднения основными носителями. Зоны изогнутся в противоположную сторону и поверхностная плотность будет меньше объемной плотности зарядов (рис. 3.4, в).

В таких структурах информация, соответствующая одному из направлений вектора спонтанной поляризации, может быть считана по изменению поверхностного потенциала полупроводника. Время перезаписи информации определяется временем переполаризации.

В таких слоистых средах большое значение имеют электрические поля в сегнетоэлектриках. Поле спонтанной поляризации должно быть меньше коэрцитивного. В этом случае энергия переключения знака поляризации меньше затрачиваемой энергии. Одновременно должно соблюдаться условие, при котором уровень коэрцитивного поля меньше уровня электрического пробивного поля. Это условие позволит производить переполяризацию сегнетоэлектрика без его разрушения.

Управляемость поверхностным потенциалом полупроводника в структуре типа "металл—сегнетоэлектрик—полупроводник" (МСЭП-структура) существенным образом зависит от качества границы раздела, от фиксированного в сегнетоэлектрике заряда, заряда, захваченного на поверхностных состояниях $Q_{сз}$, и заряда, захваченного на поверхностных состояниях $Q_{хх}$.

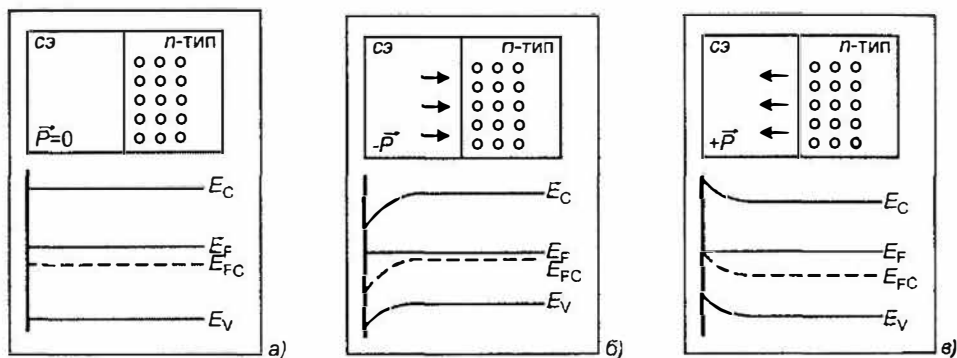


Рис. 3.4. Изменение свойств полупроводника при изменении поляризации сегнетоэлектрика

Весьма перспективной является структура "сегнетоэлектрик—фотополупроводник". Такая структура помещается между оптически прозрачными проводящими электродами. При локальном освещении полупроводника его сопротивление уменьшается, и все напряжение между проводящими электродами падает на слой сегнетоэлектрика. В освещенных областях сегнетоэлектрик переполяризуется. Возникающее пространственное распределение поляризации сегнетоэлектрика соответствует распределению освещенности и позволяет осуществить оптическое считывание.

3.2.2. Устройства памяти

Использование диэлектрических материалов при создании запоминающих устройств весьма перспективно для создания новых приборов, характеризующихся энергонезависимостью, высоким выходным сигналом, устойчивостью к воздействию перегрузок и радиации. По функциональному назначению эти устройства близки к репрограммируемым полупроводниковым ЗУ (РПЗУ). Одними из первых устройств этого типа были сегнетоэлектрические ЗУ, изготовленные на поликристаллической сегнетокерамике типа цирконата-титаната свинца (РЗТ-керамика).

Запоминающее устройство изготовлялось на основе керамической пластины 2 в форме диска или прямоугольника, фрагмент которого изображен на рис. 3.5. Термическим напылением в вакууме на внешнюю поверхность наносились металлические электроды в форме секторов или прямоугольников 1. С другой стороны наклеивался сплошной электрод из латуни или инвара 3, который служил общим электродом и для еще одной пла-

стины PZT-керамики 4 со сплошным электродом 5, присоединенной снизу. Такая слоистая структура из двух типов сегнетоэлектриков с электродами между ними представляет собой твердотельную память на диэлектрике.

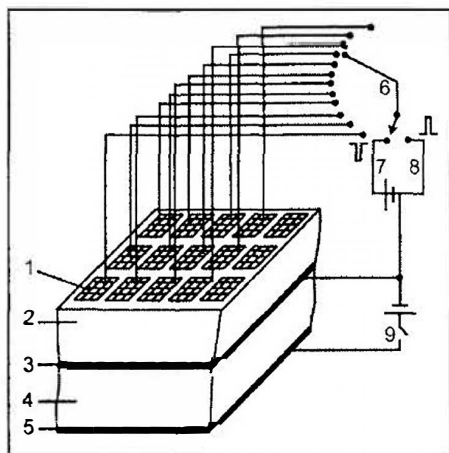


Рис. 3.5. Фрагмент устройства памяти на сегнетоэлектрической PZT-керамике

Первоначально сегнетоэлектрик имеет остаточную поляризацию, например, отрицательную, при которой вектор остаточной поляризации $-\vec{P}_2$ направлен вниз. Этому состоянию памяти можно приписать хранение во всех разрядах информационных нулей. Это же состояние можно сформировать путем переполаризации секторов (разрядов) памяти, подавая отрицательный импульс напряжения с помощью коммутаторов 6 и 7.

Программирование запоминающего устройства происходит путем записи двоичной единицы информации в соответствующий разряд памяти ($+P_2$). С этой целью туда необходимо подать положительный импульс напряжения соответствующей комбинацией коммутаторов 6 и 8. Амплитуда и длительность положительного импульса определяется процессом локальной переполаризации сегнетокерамики. После снятия напряжения записи заданное направление поляризации сохраняется в силу гистерезисных свойств сегнетокерамики.

Процесс считывания двоичной информации основывается на пьезоэлектрических свойствах сегнетокерамики. Так, если к электродам 3 и 5 сегнетокерамики 4 приложить напряжение с помощью коммутатора 9, то возникнет механическая деформация. Импульс деформации от изгибающего элемента 4 пройдет через сегнетокерамику 2 и индуцирует на секторных электродах заряд, соответствующий ранее записанной двоичной информации. Этот заряд может быть считан как соответствующий ему потенциал с помощью коммутатора 6 и, таким образом, расшифрована ранее записанная информация. Такая память на изгибающем элементе требует довольно больших входных сигналов (~ 15 В), способных вызвать значительные механические деформации и, соответственно, большие значения выходных сигналов (~ 2 В).

В процессе совершенствования конструкции сегнетокерамику 2 и 4 изготавливают из нескольких различных составов, с разными значениями коэрцитивной силы, что позволяет оптимизировать соотношения между амплитудами импульсов записи и считывания информации, выбирать резонансный или нерезонансный режим работы.

Такие достаточно надежные и помехоустойчивые устройства могут использоваться для задания времени срабатывания. Описанные ЗУ отличаются простотой в обращении, имеют малую стоимость, простую конструкцию, высокую радиационную стойкость. Со временем эта конструкция стимулировала развитие идеи создания пьезоэлектрической вычислительной среды (ПЭВС) с функциями запоминающих устройств. При этом использовались два типа динамических неоднородностей — сегнетоэлектрические домены для хранения информации и акустические волны для обработки информации.

Важно отметить, что в этом случае решается принципиально новая технологическая задача: активная диэлектрическая континуальная среда в виде твердых тел заменяется аналогичной пленочной структурой. При этом необходимо воспроизвести сегнетоэлектрические свойства в тонких континуальных слоях. Сложность заключается в том, что физическая природа сегнетоэлектричества формируется дальнедействующими силами дипольного происхождения. Радиус действия таких сил определяется корреляционной длиной. В этом случае толщина слоев сегнетоэлектрика должна на порядок превышать корреляционную длину и составлять 100—500 нм. При этом необходимо достаточно точно воспроизвести стехиометрии трех-, четырех- и более компонентного состава. В этих пленках должны быть сохранены физические свойства объемного материала. При этом одновременно уменьшаются управляющие электрические и механические напряжения, оптические и акустические потери, улучшаются другие потребительские свойства.

Вопросы получения промышленных пленок сегнетоэлектриков являются специальной, технологической проблемой. Информационным значениям "0" или "1" соответствует разность фаз считываемого сигнала в 2π , что соответствует направлениям поляризации $\pm P_2$ в отдельных ячейках памяти.

Разработанные *пьезокерамические матрицы* (ПКМ) функционируют на принципе импульсного или резонансного считывания информации и напоминают конструкцию, рассмотренную на рис. 3.5.

Считывание информации происходит без ее разрушения с тактовой частотой до 1 МГц, определяемой временем переполяризации (~ 50 мкс). Допускается 10^6 — 10^8 циклов перезаписи информации. Разработаны интегральные ПКМ емкостью более 512 бит с напряжением перезаписи ~ 4 В, работающие в интервале температур 80—200 °С, выдерживающие ударные нагрузки до 3×10^4 г и отличающиеся высокой радиационной стойкостью.

Следует отметить возможность увеличения на порядок плотности записи информации без увеличения числа ячеек памяти. В основе лежит физический принцип переключения поляризации из состояния $-P_2$ в состояние $+P_2$. Из-за отсутствия абсолютного порога переключения можно один бит информации, например "1", разложить на амплитудный многоуровневый аналоговый сигнал и воспользоваться законами многозначной логики. Так, в ПКМ достаточно легко обеспечить 20 уровней $\pm \Delta P_2$ во всем диапазоне $\pm P_2$ с разрешением цифроаналогового преобразования 10^{-1} В. Одним из основных требований к сегнетоэлектрическим материалам таких ПКМ является обеспечение коэффициента прямоугольности петли диэлектрического гистерезиса $0,85 \leq K_{пр} = P_2 / P_s < 1$.

Весьма перспективным материалом для ЗУ являются сегнетоэлектрические пленки на основе нитрата калия (KNO_3). Они отличаются малым временем переключения $\sim 10^{-9}$ с, высокой тактовой частотой $\sim 10^8$ Гц, малыми геометрическими размерами доменов, позволяющими получить размер ячейки ~ 20 мкм² и, соответственно, достичь степени интеграции 10^8 — 10^{10} бит/см². Емкостные элементы функциональной электроники нашли перспективное применение в устройствах памяти схемотехнической электроники.

Известно, что одним из существенных недостатков существующих схем памяти является отсутствие энергозависимости. Разработано несколько подходов к созданию энергонезависимых схем памяти. Лучший результат дает технология осаждения на поверхность тонкой пленки цирконата титаната свинца (ЦТС или PZT), с помощью которой формируются сегнетоэлектрические конденсаторы. Такие конденсаторы соответствующим образом поляризуются и, имея высокое значение диэлектрической проницаемости ($\epsilon \approx 1200$), обеспечивают определенное состояние ЗУ при частом отключении питания.

Керамика цирконат титанат свинца (PZT) обладает высоким удельным сопротивлением (изолятор), термической и химической стойкостью, высокой механической прочностью. Температура фазового перехода PZT-керамики составляет 330 °C, и поэтому такие схемы могут работать в температурном диапазоне – 180 ... + 320 °C. Пробивное напряжение таких схем составляет ~ 40 В, диэлектрическая проницаемость выше $1,2 \times 10^3$, что позволяет резко повысить плотность интеграции по сравнению с известными схемами ДОЗУ.

Таким образом, динамические неоднородности в виде сегнетоэлектрических доменов, встроенных в устройства схемотехнической электроники, позволяют получить их новое качество, существенно улучшить выходные характеристики.

3.2.3. Процессоры

Значительный интерес представляют процессоры для обработки больших информационных массивов, реализованные на принципах функциональной электроники. Это обусловлено тем, что такие устройства позволяют обрабатывать информацию в аналоговом виде, одномоментно либо весь массив, либо его часть. Результирующая информация может быть преобразована в цифровую форму. Возможна операция выделения разностной информации между двумя и более временными состояниями информационного массива.

Процессоры можно реализовать на слоистых структурах типа "сегнетоэлектрик—фотополупроводник". С обеих сторон такой структуры напыляются проводящие пленки, причем со стороны фотополупроводника пленка должна быть оптически прозрачной. При формировании на слое фотополупроводника оптического информационного массива на слоистую структуру одновременно подается "записывающий" импульс напряжения. В освещенных участках сопротивление фотополупроводника резко уменьшается, и все приложенное напряжение падает на слое сегнетоэлектрика. Под засвеченными местами сегнетоэлектрик переполаризуется. В темных местах сопротивление фотополупроводника велико и все напряжение падает на него. Сегнетоэлектрик под темными местами не переполаризуется. Возникает пространственное распределение поляризации сегнетоэлектрика, соответствующее распределению освещенности.

Таким образом, оптический информационный массив преобразован в электрически заряженный информационный массив, который может быть легко считан. Стирание записанной информации может быть осуществлено путем засветки всей поверхности фотополупроводника с одновременной подачей импульса напряжения, противоположного по знаку записывавшему импульсу. Это позволяет перевести сегнетоэлектрик в исходное состояние. Такая структура является основой для создания оптических процессоров сигналов с промежуточным запоминанием информации, систем отображения информации.

Оптоэлектронные свойства материалов определяются характером персориентации доменов над воздействием поляризующего электрического поля. Домены образуют зерна, размер которых определяется характеристиками материала. Так для цирконаттитаната свинца, легированного лантаном (ЦТСЛ, PLZT), размер зерна составляет несколько мик-

рон и включает в себя до десятка сегнетодоменов. Простейший тип переориентации доменов под действием электрического поля приводит к изменению вектора поляризации на противоположное (180°). Однако режим поворота доменов каждого элемента сопряжен с трудностями нормирования матрицы таких элементов с множеством электродов.

Устройство на основе метода преимущественной ориентации доменов с помощью механического напряжения, создаваемого деформацией пластины, получило название **ФЕРПИК** (FERPIC, FERroelectric PICture). Его структура приведена на рис. 3.6. Проведем анализ работы процессора в терминах модели прибора функциональной электроники.

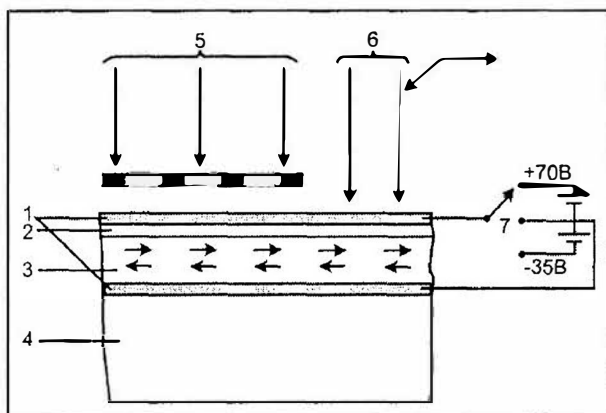


Рис. 3.6. Устройство обработки информации типа ФЕРПИК: 1 — прозрачные проводящие электроды; 2 — фотопроводник; 3 — сегнетокерамика; 4 — прозрачная подложка; 5 — облучение через маску; 6 — облучение сканированием; 7 — переключатель режимов работы; 8 — источники напряжения

В качестве основной континуальной среды используется сегнетоэлектрик 3 в виде [ТСЖ]-керамики $[(\text{Pb}_{0,99}\text{Zr}_{0,65}\text{Ti}_{0,35})_{0,93}\text{La}_{0,07}\text{O}_3]$ со средним размером зерна $1,0\text{—}1,5\text{ мкм}$. В такой среде динамические неоднородности представляют собой субмикронные сегнетоэлектрические домены. Континуальная среда подвергается смещению деформацией, что позволяет сориентировать домены вдоль оси напряжения. Континуальная среда сопряжена, с одной стороны, фотополупроводником 2, с другой — прозрачным проводящим электродом 1 и прозрачной подложкой 4, несущей механические нагрузки. Прозрачный электрод 1 нанесен и с другой стороны фотопроводящего слоя.

Запись информационного массива осуществляется путем освещения устройства через транспарант 5, либо сканированием модулируемого по амплитуде (или по растру) лазерного луча. Управление записью осуществляется полями. На керамическую пластину подается напряжение $+70\text{ В}$. В местах засветки уменьшается сопротивление фотопроводящего слоя, и величина напряженности электрического поля достаточна для поворота доменов на 90° . По окончании сканирования информационный массив записан в виде распределения доменов различных направлений поляризации. В этом случае можно говорить, что одновременно произошло детектирование информационных сигналов.

Заметим, что оптическое пропускание устройства типа ФЕРПИК зависит от состояния доменов.

Записанную информацию можно считать, соединив электроды (поставив переключатель 7 в среднее положение и осветив устройство пучком монохроматического света). При этом достигается разрешение $\sim 40\text{ лин/мм}$, эффективность считывания $\sim 10\%$ и контраст $10:1$.

Стирание информационного массива можно осуществить путем освещения всей поверхности и одновременной подачей на слоистую структуру отрицательного напряжения (-35 В). При этом динамические неоднородности в виде сегнетоэлектрических доменов возвращаются в исходное состояние. Это устройство обработки информации в русской терминологии называют еще "Сегнот" — сегнетоэлектрическое отображение.

В дальнейшем эта конструкция процессора стимулировала появление аналогичных устройств на основе слоистых структур "сегнетокерамика—фотопроводник". К ним относятся, например, КЕРАМПИК (CERAMPIC, CERAMic PICture), континуальной средой в котором служит крупнозернистая керамика $(\text{Pb}_{0,99}\text{La}_{0,77}\text{Zr}_{0,65}\text{Ti}_{0,35})_{0,93}\text{O}_3$ с размером зерен $\sim 3,0$ мкм.

В процессоре такого типа в качестве континуальной среды используется сегнетокерамика толщиной 250 мкм, обладающая свойством локально деформироваться и, таким образом, распределение поверхностной деформации соответствует распределению переполаризованных участков. Амплитуда деформации составляет 0,5 мкм и шлирен-методом можно получить увеличенное изображение в отраженном свете с разрешением до 100 лин/мм.

Слоистая структура может применяться в качестве оптических процессоров, в частности для преобразования входного некогерентного изображения в выходной когерентный сигнал, что используется в голографических ЗУ. Весьма перспективно применение таких структур в процессорах сложения и вычитания информационных массивов. В этих устройствах используется способность сегнетоэлектрика к устойчивому частичному переключению в пределах от $-P_S$ до $+P_S$ под действием электрического поля.

В устройстве, аналогичном известной конструкции ФОТОТИТУСа, операции сложения изображений можно применять для увеличения отношения сигнал/шум в процессе обработки изображений. Операцию вычитания изображений можно использовать для обращения изображения — преобразования позитивного изображения в негативное, дифференцирования изображения, выделения слабоконтрастных контуров.

Слоистые структуры типа "сегнетоэлектрик—фотополупроводник" можно использовать в основе устройств памяти, оптических процессоров, устройств отображения информации.

Контрольные вопросы

1. Что такое функциональная диэлектрическая электроника?
2. Какие типы динамических неоднородностей диэлектрической природы вы знаете?
3. Что такое сегнетоэлектрический домен? Какими свойствами он обладает?
4. Опишите устройство памяти на основе сегнетокерамики.
5. Опишите конструкцию и принцип работы процессора типа ФЕРПИК.

Рекомендуемая литература

1. Барфут Дж., Тейлор Дж. Полярные диэлектрики и их применение. — М.: Мир, 1981.
2. Кузьминов Ю. С. Сегнетоэлектрические кристаллы для управления лазерным излучением. — М.: Наука, 1982.
3. Шука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

4. Функциональная полупроводниковая электроника

4.1. Физические основы

Функциональная полупроводниковая электроника является направлением в функциональной электронике, в котором изучаются взаимодействия динамических неоднородностей в полупроводниковых континуальных средах с физическими полями, а также возможность создания приборов и устройств для преобразования и хранения информации.

Одним из основных отличий приборов и устройств полупроводниковой функциональной электроники от аналогичных устройств полупроводниковой схемотехнической электроники является способ обработки информации. В приборах и устройствах функциональной электроники для обработки и хранения информации используются динамические неоднородности различной физической природы, в приборах же схемотехнической микроэлектроники — различные схемные решения на основе транзисторных структур.

Если в устройствах функциональной электроники развивается принципиально новый подход к методам обработки и хранения информации, связанный с использованием динамических неоднородностей в качестве носителей информационного сигнала, то в традиционных устройствах полупроводниковой электроники используются либо классические схемотехнические решения, известные со времен дискретной электроники, либо новые технические решения, использующие уникальные возможности микроэлектронной технологии, обязательно связанные с созданием в строго регламентированных технологических процессах различных статических неоднородностей.

4.1.1. Динамические неоднородности в полупроводниках

В состоянии термодинамического равновесия электроны проводимости и дырки в полупроводниках движутся хаотически. Направленное движение, или дрейф, электроны и дырки приобретают под действием внешнего электрического поля \vec{E} . Скорость дрейфа определяется выражением

$$V_{др} = \mu \vec{E},$$

где μ — подвижность носителей заряда, значение которой лежит в пределах от 10^5 до 10^{-3} см²/В·с. Носители могут объединяться в группы, называемые *зарядовыми пакетами*. Величина зарядового пакета определяется количеством сконцентрированных носителей. Его скорость перемещения в полупроводнике — соответствующими электрическими полями. Зарядовые пакеты, состоящие из электронов или дырок, представляют собой динамические неоднородности полупроводниковой природы. Величина зарядового пакета определяется физическими параметрами среды и топологией емкостной структуры:

$$Q = SC_{зд}U_m,$$

где S — площадь затвора, $C_{зд} = \epsilon_d \epsilon_0 / h_d$ — удельная емкость подзатворного диэлектрика, ϵ_d и ϵ_0 — соответственно диэлектрические проницаемости окисла и вакуума, U_m — мак-

симальное напряжение на затворе, h_d — толщина подзатворного диэлектрика. Верхнее предельное значение поверхностной плотности зарядов оценивается величиной

$$Q_{\text{max}} = 5 \times 10^{-7} \text{ Кл} \cdot \text{см}^{-2} = 3,13 \times 10^{12} \text{ электронов/см}^{-2},$$

а нижнее —

$$Q_{\text{min}} = 5 \times 10^{-10} \text{ Кл} \cdot \text{см}^{-2} = 3,13 \times 10^9 \text{ электронов/см}^{-2}.$$

Критический уровень величины зарядового пакета, соответствующий значению логической единицы, оценивается величиной

$$Q_{\text{крит}} = 10^5 \text{ электронов/см}^2.$$

Если приложить к полупроводнику внешнее магнитное поле, перпендикулярное к направлению дрейфа носителей, то под действием силы Лоренца носители отклоняются в соответствии с известным эффектом Холла.

В сильных электрических полях напряженностью $10^2 + 10^3$ В/см наблюдается изменение распределения носителей по энергии в сторону увеличения средней энергии. Изменяется подвижность, время свободного пробега, коэффициент диффузии. Разогрев носителей способствует возникновению горячих электронов и приводит к отклонению от закона Ома

$$E = \rho I,$$

где I — плотность тока, E — напряженность электрического поля, ρ — удельное сопротивление полупроводника. Возникает отрицательное дифференциальное сопротивление (ОДС) или отрицательная дифференциальная проводимость (ОДП). Это явление связано с переходом электронов, стимулированных электрическим полем зоны проводимости, из низкоэнергетической долины, в которой их подвижность относительно велика, в высокоэнергетические долины с меньшими значениями подвижностей носителей. Флуктуации концентрации свободных носителей в среде с объемным отрицательным удельным сопротивлением генерируют пространственный заряд.

Различают два типа полупроводников, обладающих объемной неустойчивостью. У некоторых полупроводников удельное сопротивление при определенных значениях напряженности поля резко и нелинейно увеличивается (рис. 4.1, а). Приборы на таких полупроводниках имеют N -образную вольтамперную характеристику и относятся к полупроводникам с ОДП (рис. 4.1, б). В полупроводниках такого типа возникают электрические домены, движущиеся по направлению тока (рис. 4.1, в). Эти домены характеризуются тем, что их поле значительно больше, а концентрация носителей в них меньше, чем в остальной части полупроводниковой среды.

В этих полупроводниках возникает внутренняя положительная обратная связь по напряжению. Домены получили название доменов Ганна и также представляют собой динамические неоднородности. Линейные размеры таких доменов лежат в пределах нескольких микрон, скорость перемещения составляет $\sim 10^5$ м/с, электрическое поле в нем достигает значения 10^5 В/см. Если домен не образуется, то в полупроводниковой континуальной среде могут возникнуть волны пространственного заряда.

В другом типе полупроводников наблюдается падение удельного сопротивления с ростом напряженности электрического поля. Этот тип полупроводников относят к полупроводникам с ОДС. Приборы, изготовленные на основе таких полупроводников, имеют S -образную характеристику с положительной обратной связью по току (рис. 4.1, г). В приборах с S -образными ВАХ возникают динамические неоднородности с высокой плотностью тока, называемые *токовыми шнурами* (рис. 4.1, е). Токовые шнуры могут

быть использованы в качестве динамических неоднородностей только при условии, если плотность тока в шнуре существенно меньше значения пробоя полупроводника и последующего необратимого процесса его разрушения.

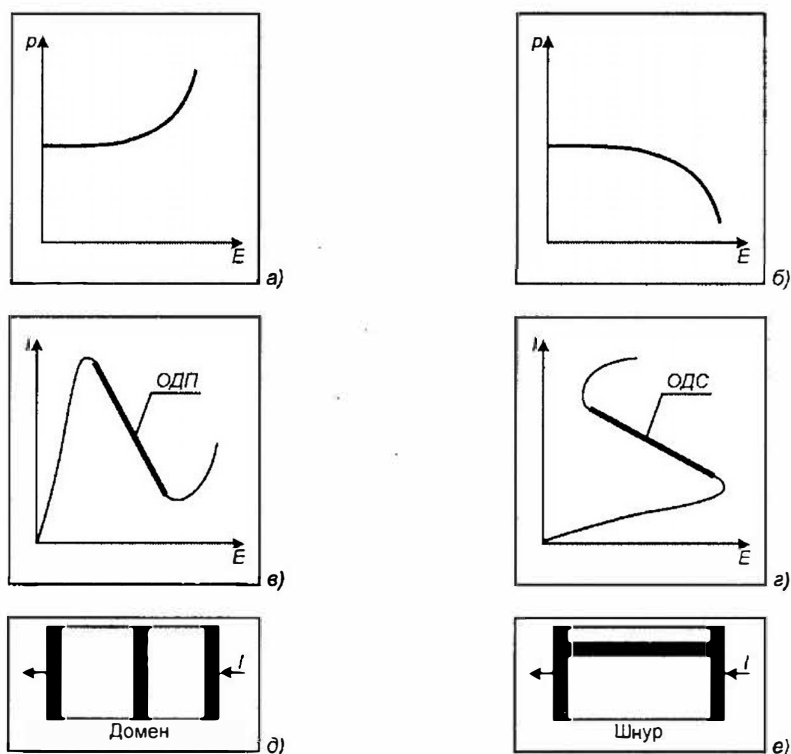


Рис. 4.1. Зависимость удельного дифференциального сопротивления от напряженности электрического поля для двух типов полупроводников: а, в — для приборов с доменами и N -образным ВАХ; б, г — для приборов с токовыми шнурами и S -образными ВАХ. Формирование домена сильного поля (д) и токового шнура высокой плотности (е)

В полупроводниках, обладающих пьезоэлектрическими свойствами (Te, Se, CdS, CdSe, ZnO, GaAs, InSb), в результате электрон-фононного взаимодействия поток носителей становится интенсивным излучателем упругих волн. Частным случаем электрон-фононного взаимодействия является акустоэлектронное взаимодействие. В результате обмена энергией между потоком дрейфующих электронов в электрическом поле в направлении распространения звуковой волны акустоэлектронное взаимодействие зависит от соотношения скорости дрейфа $V_{др}$ и скорости звука c . Если выполняется соотношение $V_{др} = \mu E_0 < c$, то волна поглощается электронным газом, а если $V_{др} > c$ — электроны отдают свою кинетическую энергию акустической волне. Как уже ранее описывалось, такие акустические волны также являются динамическими неоднородностями.

В электронной системе полупроводника, обладающего объемной отрицательной дифференциальной проводимостью, когда характерные масштабы изменения параметров среды, приложенного поля и концентрации носителей заряда существенно превосходят длины

релаксации энергии и импульса электронов, возникает возмущение электронной плотности. Такие волны называются *волнами пространственного заряда* (ВПЗ). Так для GaAs в сильном ($\sim 3,5 \times 10^3$ В/см) электрическом поле возникают волны пространственного заряда, движущиеся со скоростью $\sim 10^5$ м/с. Фазовая скорость ВПЗ совпадает со скоростью дрейфа носителей заряда $V_{др}$. Дисперсия ВПЗ незначительна вплоть до частоты 50 ГГц.

Волны пространственного заряда могут усиливаться при распространении. Этими волнами можно управлять, меняя вектор скорости дрейфа носителей. Длина когерентности в тонких полупроводниковых слоях может на два порядка превышать длину волны ВПЗ. Такие свойства ВПЗ как динамических неоднородностей делают их весьма перспективными в устройствах функциональной электроники.

В плазме полупроводника, находящегося в магнитном поле, могут возникнуть *геликоны* — слабо затухающие низкочастотные электромагнитные волны. Геликоны могут взаимодействовать со звуковыми волнами.

В плазме твердых тел могут возникать кванты плазменных колебаний — *плазмоны*, вызываемые созданием электрического поля за счет флуктуаций плотности заряда. В свою очередь электрическое поле в плазме вызывает ток носителей, стремящийся восстановить локальную электронейтральность. Вследствие инерционности носителей, проскакивающих положение равновесия, возникают кванты коллективных колебаний со спектром, определяемым зонной структурой. Плазмон является квазичастицей, способной переносить информационный сигнал.

В полупроводниковых кристаллах могут распространяться *фононы* — квазичастицы, соответствующие волне смещения атомов (ионов) и молекул из положения равновесия. Фононы взаимодействуют друг с другом, с другими квазичастицами, а также с дефектами кристаллической решетки. С помощью фононов осуществляется связь всех квазичастиц твердотельной континуальной среды с окружающей средой. В аморфных средах понятие фононов вводится для длинноволновых акустических колебаний.

Интерес вызывают *поляроны*. Это квазичастицы, состоящие из носителей в полупроводнике (электрон или дырка) и фонона, образованного деформацией и поляризацией кристаллической решетки за счет движения в полупроводнике электрона проводимости (дырки). Такая составная квазичастица движется по кристаллу как единое целое, она способна переносить заряд. Эффективная масса полярона значительно превышает эффективную массу электрона. Поляроны могут служить носителями заряда в кристалле. Возможно образование *биполяронов*, представляющих собой связку двух электронов в общей деформационной яме. По своей природе поляроны близки к *флуктуонам*.

В полупроводниковых континуальных средах наблюдаются и другие квазичастицы. Речь идет об *экситонах*, представляющих собой связанное состояние электрона проводимости и дырки. Такая квазичастица электрически нейтральна и не способна переносить заряд и массу. Различают по два типа экситонов.

Экситоны Френкеля возбуждаются в узлах кристаллической решетки и благодаря межмолекулярным взаимодействиям они распространяются по кристаллу в виде волны.

Экситоны Ванье — *Мотта* представляют собой водородоподобное связанное состояние электрона и дырки. Время жизни экситонов достаточно мало. Они исчезают в результате рекомбинации с излучением фонона и фотона, а также при захвате дефектами решетки. Именно это обстоятельство затрудняет использование экситонов в качестве динамических неоднородностей в устройствах функциональной электроники.

Если экситонный газ освещать, то в результате взаимодействия экситонов с энергией ε и фотонов частоты $\nu = \frac{\varepsilon}{h}$ рождается составная квазичастица — *поляритон*. Свойства поляритонов отличаются от свойств экситонов и фотонов. Их наличие в полупроводниках существенно влияет на оптические спектры последних.

При дискретном туннелировании одного электрона возникают импульсы тока I , определяемые соотношением

$$\int I \cdot dt = e \cong 1,6 \times 10^{-19} \text{ А} \cdot \text{с}.$$

Такой импульс тока может быть использован в качестве динамической неоднородности.

Перечисленные типы динамических неоднородностей являются наиболее заслуживающими внимание и не исчерпывают весь их перечень.

4.1.2. Континуальные среды

Полупроводниковые среды занимают промежуточное положение по проводимости электрического тока между диэлектриками и металлами. При комнатной температуре их удельное сопротивление лежит в пределах от 10^{-6} до 10^9 Ом·м. Такой огромный диапазон (15 порядков) значений определяется рядом специфических свойств полупроводников. Прежде всего, удельная проводимость зависит от количества примесей. Одна тысячная процента примеси может изменить проводимость полупроводника на 4—5 порядков. Полупроводники обладают отрицательным температурным коэффициентом сопротивления в большом диапазоне температур. К полупроводникам относится большой ряд простых веществ из II, III, IV, VI, VII групп таблицы Менделеева: германий, кремний, селен, теллур, бор, углерод, фосфор, сера, сурьма, мышьяк, серое олово, йод, а также соединения различных типов: $A^I B^V$, $A^I B^{VI}$, $A^I B^{VII}$, $A^{II} B^{IV}$, $A^{II} B^V$, $A^{II} B^{VI}$, $A^{II} B^{VII}$, $A^{III} B^V$, $A^{III} B^{VI}$, $A^{III} B^{VII}$, $A^{IV} B^{IV}$, $A^{IV} B^V$, $A^{IV} B^{VI}$, $A^V B^{VI}$, $A^V B^{VI}$, $A^{VI} B^{VI}$, $A^{VII} B^{VI}$, $A^I B^{VII} B_2^{VI}$, $A^I B^V B_2^{VI}$ и т. д.

Различают некоторые классы полупроводниковых континуальных сред. Элементарные полупроводники представляют собой кристаллическую решетку. Например, в германии и кремнии, являющимися основными материалами, атомы образуют кристаллическую решетку типа алмаза. Каждый атом взаимодействует с четырьмя ближайшими соседями, с каждым из которых связан ковалентной связью. Содержание посторонних примесей не должно превышать $10^{-7}\%$, а дозируемое введение донорных (P, As, Sb) или акцепторных (B, Al, Ga, In) примесей меняет их тип проводимости в достаточно широких пределах. Монокристаллы можно выращивать методами Чохральского или бестигельной зонной плавки, получать в процессе эпитаксиального выращивания тонких кристаллических слоев, а также в процессе лазерной рекристаллизации поликристаллов.

Соединения типа $A^{III} B^V$ (GaAs, InAs, GaSb, InSb, GaP, InP) представляют собой твердые растворы. Они характеризуются ковалентной связью с некоторой долей ионной составляющей. Возможно образование тройных и более сложных растворов ($Ga_x Al_{1-x} As$, $GaAs_{1-x} P_x$ и т. п.). Электрические свойства таких полупроводников могут меняться путем легирования донорными (Sn, Te, Se, S) или акцепторными (Zn, Cd, Mg) примесями. Монокристаллы этого класса получают методом Чохральского, зонной кристаллизацией, кристаллизацией из газовой или жидкой фазы, молекулярно-лучевой эпитаксией.

Соединения типа $A^{II} B^{IV}$ (ZnS, CdS, HgSe, ZnSe, ZnTe) представляют собой твердые растворы. Кристаллическая структура имеет ковалентно-ионную связь. Физические свойства

определяются содержанием точечных дефектов, обусловленных отклонениями от стехиометрии состава. Аналогичными свойствами обладают соединения типа $A^{IV}B^{VI}$ (PbS, SnS, SnSe, PbTe, SnTe).

Соединения элементов VI группы с элементами I—V групп образуют кристаллы, в которых преобладает ионная связь. Если элементы этой группы связаны с переходными или редкоземельными элементами, то такие соединения обладают магнитными свойствами. Некоторые из таких полупроводников обладают ферромагнитными свойствами (EuO, EuS, $CdCr_2Se_4$), другие — антиферромагнитными (EuTe, EuSe, NiO).

Тройные соединения $A^{II}B^{IV}C_2^V$ ($CdSnAs_2$, $CdGeAs_2$, $ZnSnAs_2$) обладают ферромагнитными и (или) сегнетоэлектрическими свойствами.

Аморфные полупроводники представляют собой класс сильно легированных веществ — компенсированный полупроводник со смешанной ионно-ковалентной связью. К этому классу относятся халькогениды свинца: галенит (PbS), клаусталит (PbSe), алтаит (PbTe), халькогенидные стекла ($As_{21}Ge_{30}Se_{21}Te_{18}$), оксидные стекла ($V_2O_5-P_2O_5$).

Органические полупроводники характеризуются наличием в молекулах сопряженных связей. К этому классу веществ относятся ароматические соединения (например, нафталин и др.), природные пигменты (хлорофил, каротин и др.), ион-радикальные соли, молекулярные комплексы с переносом заряда. Это могут быть монокристаллы, поликристаллы или аморфные порошки.

Определенный интерес вызывают полупроводники, в которых проявляются механизмы оптической нелинейности.

Качественно новый класс полупроводниковых материалов и континуальных сред представляют собой *сверхрешетки* — синтезированные структуры, в которой на носители заряда, помимо кристаллической решетки, действует дополнительный потенциал с периодом, существенно превышающим постоянную решетки. Параметры этого потенциала можно менять и тем самым управлять энергетическим спектром носителей. Тип сверхрешетки определяется технологией изготовления. *Композиционные* сверхрешетки формируются путем чередования тонких слоев различных материалов с близкими значениями постоянных решетки. *Легированные* сверхрешетки формируются из слоев с дырочной и электронной проводимостями в объеме одной полупроводниковой среды. *Спиновые* сверхрешетки создаются технологией селективного легирования полупроводника магнитными и немагнитными примесями. Наиболее совершенные сверхрешетки формируют при использовании молекулярно-лучевой эпитаксии.

Свойства полупроводниковых континуальных сред могут быть достаточно полно описаны с помощью квантовой теории энергетического спектра электронов, зонной теории.

4.1.3. Генераторы динамических неоднородностей

Процесс генерирования динамических неоднородностей в полупроводниковой континуальной среде определяется конкретными физическими явлениями и эффектами, характерными для данного их типа, и данной среды. Рассмотрим некоторые генераторы динамических неоднородностей. Генераторы зарядовых пакетов в полупроводниковых структурах могут быть нескольких типов. В приборах и устройствах функциональной электроники используется нестационарное состояние полупроводниковых структур. Например, в МДП-конденсаторе при подаче напряжения на металлическую обкладку возникает электрическое поле. Это электрическое поле вытесняет основные носители в глубь полупроводника N (напряжением на затворе U , площадью металлического затвора S).

В результате процесса термогенерации в полупроводнике образуются электронно-дырочные пары и под воздействием электрического поля генерируются те заряженные частицы, которые имеют знак противоположный напряжению на металлической обкладке. Они же заполняют обедненную область. Заряженные частицы одноименного знака с основными носителями полупроводника отталкиваются этим полем в глубь среды. Однако процессы термогенерации инерционны ($0,1 \div 0,5$ с) и не используются в качестве основного механизма создания динамических неоднородностей типа зарядовых пакетов.

Генерацию зарядовых пакетов целесообразно осуществлять методом инжекции — экстракции зарядов через прямосмещенный р—п-переход. Одно из конструктивных решений генератора зарядовых пакетов приведено на рис. 4.2. Здесь рассмотрен случай, когда используется подложка из кремния р-типа с зарядовыми пакетами, состоящими из электронов. Практически аналогичная ситуация реализуется для подложки из n-типа.

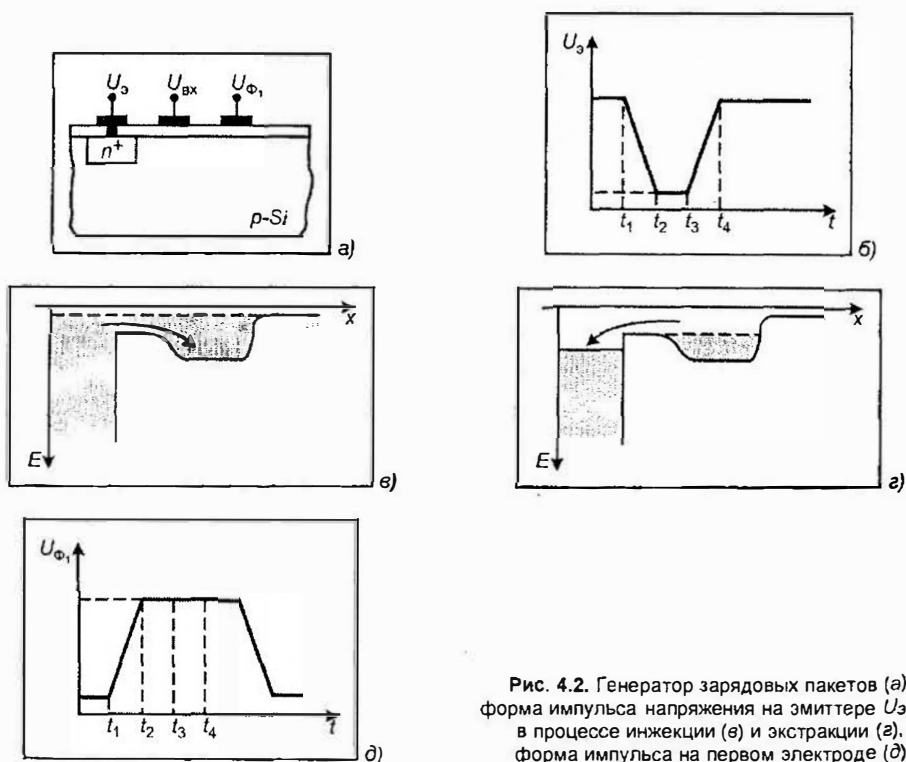


Рис. 4.2. Генератор зарядовых пакетов (а), форма импульса напряжения на эмиттере U_3 (б), в процессе инжекции (в) и экстракции (г), форма импульса на первом электроде (д)

Сильно легированная область n-типа является источником электронов и находится под потенциалом U_3 . На входной электрод (затвор) подается потенциал. Потенциал на первом электроде $U_{Ф1}$ больше потенциала на входном электроде $U_{вх}$. Происходит экстракция или ток носителей заряда из высоколегированной области перетекает под электроды в другой области полупроводника. Ток течет только во время действия импульса на эмиттирующем электроде. Это время выбирается таким, чтобы успела заполниться потенциальная яма под электродом $U_{Ф1}$.

Величина накопленного заряда определяется как:

$$Q_{\text{н}} = SC_{\text{а}}(U_{\text{ф}} - U_0),$$

где $C_{\text{а}}$ — удельная емкость МОП-конденсатора, S — площадь затвора, $U_{\text{ф}}$ — фазовый потенциал на затворе, U_0 — пороговое напряжение, необходимое для создания канала в полупроводнике.

Такой процесс называют еще режимом *залива* заряда через барьер. Он характерен тем, что в исходном состоянии элемента весь заряд находится в эмитирующей области. Между этой областью и первой ячейкой существует потенциальный барьер, формируемый импульсом напряжения. Заполнение ячейки осуществляется как бы подъемом дна и заливкой через канал, образованный под входным затвором (рис. 4.2, в).

Когда зарядовый пакет полностью сформирован, дно опускается и часть заряда, превышающая емкость зарядового пакета, снова стекает в эмиттерную область. Это наступает, когда потенциалы входной и первой ячейки сравниваются (рис. 4.2, г).

Генерацию зарядовых пакетов можно осуществить и оптическим путем. Различают четыре основных способа генерации оптической информации:

- ☐ непосредственный оптический ввод со стороны подложки;
- ☐ оптический ввод со стороны оптических электродов;
- ☐ ввод с использованием фоточувствительного слоя;
- ☐ ввод путем фотоэлектронного преобразования оптической информации.

С этой целью используются внешний и внутренний фотоэффекты. Основным физическим механизмом генерации зарядовых пакетов является внутренний фотоэффект, возникающий при поглощении фотонов и генерации носителей заряда. Различают собственный внутренний фотоэффект, заключающийся в образовании одновременно электронов в зоне проводимости и дырок в валентной зоне. Этот эффект реализуется в случае, когда энергия фотона больше ширины запрещенной зоны ($h\nu > E_{\text{з}}$).

Примесный внутренний фотоэффект возникает в случае, когда энергия фотона меньше ширины запрещенной зоны, но больше энергии возбуждения примеси. В этом случае при поглощении фотона возбуждаются доноры и акцепторы в примесях, и образуется один носитель: электрон в зоне проводимости или дырка в валентной зоне.

Фотогенерированные в полупроводниковой среде носители собираются в элементе накопления или в фотодиоде и являются по существу генераторами динамических неоднородностей в устройствах подобного типа (рис. 4.3). Модель накопителя можно рассматривать как модель фотодиода с глубиной р—п-перехода, равной нулю. Генерируемые в обедненной области носители в ней не рекомбинируют и попадают в потенциальную яму.

Если носители генерируются в квазинейтральной области, то они диффундируют к границе обедненной области. Под действием электрического поля они падают в потенциальную яму, частично рекомбинируя. Общий фототок определяется суммой тока в обедненной области $I_{\text{об}}$ и диффузионного тока $I_{\text{диф}}$ и может быть записан в виде:

$$I_{\text{ф}} = I_{\text{об}} + I_{\text{диф}}.$$

Зарядовый пакет формируется интеграцией тока за промежуток времени освещения τ и определяется пропусканием электродной системы T , падающим потоком излучения Φ_0 , частотой ν , коэффициентом поглощения α , толщиной обедненной области $d_{\text{об}}$, диффузионной длиной основных носителей L . В итоге будет сформирован зарядовый пакет величины:

$$Q_{zn} \approx q\Phi_0 \cdot T \cdot \tau \cdot Sh\nu \left[1 - \exp \frac{-\alpha d_{об}}{1 - \frac{\alpha L_n}{1 + L_n}} \right]. \quad (4.1)$$

Для значений $d_{об} = 10$ мкм, $L_n = 100$ мкм и для видимого светового диапазона выражение в квадратной скобке будет близко к единице.

Тогда имеем

$$Q_{zn} \approx q\Phi_0 T Sh\nu.$$

Квантовый выход η_ϕ определяется отношением фототока к поглощенному потоку излучения:

$$\eta_\phi = \eta_{об} + \eta_{diff}.$$

Чувствительность определяется произведением коэффициента пропускания T на общий квантовый выход η_ϕ .

Фотодиодная ячейка также может служить генератором зарядовых пакетов и ее использование предпочтительнее в коротковолновой части спектра ($0,4 \div 0,45$ нм).

На рис. 4.3 приведены конструкции фотогенераторов зарядовых пакетов. Генератор типа МОП-накопителя (рис. 4.3, а) представляет собой фотодиод и три МОП-затвора.

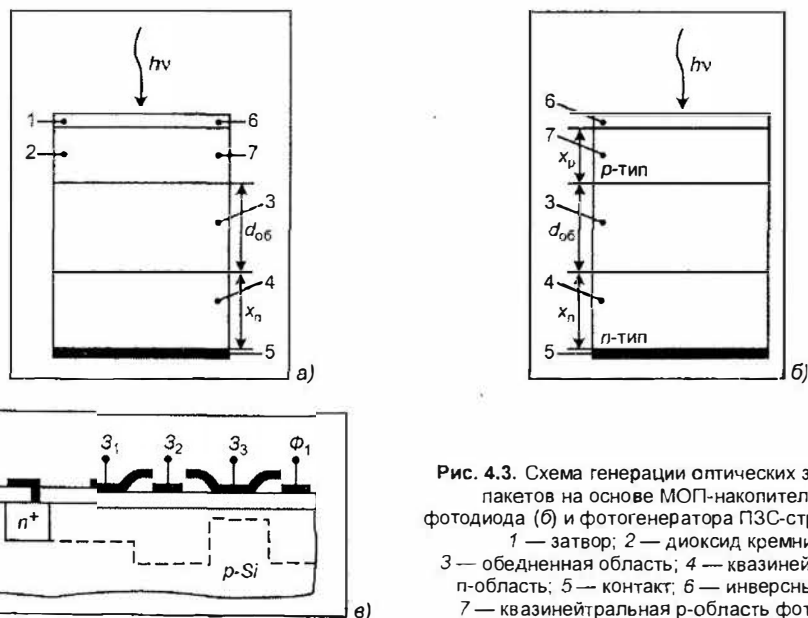


Рис. 4.3. Схема генерации оптических зарядовых пакетов на основе МОП-накопителя (а), фотодиода (б) и фотогенератора ПЗС-структуры (в): 1 — затвор; 2 — диоксид кремния; 3 — обедненная область; 4 — квазинейтральная п-область; 5 — контакт; 6 — инверсный слой; 7 — квазинейтральная р-область фотодиода

Первый разделительный затвор 3_1 служит для формирования потенциального барьера и зарядового пакета в процессе разделения носим фотогенерирования электронно-дырочных пар.

Второй затвор 3_2 является накопительным и формирует величину зарядового пакета. По окончании периода накопления заряда и окончательного формирования зарядового пакета

на разрешающий затвор Z_2 подается высокий потенциал и зарядовый пакет перетекает в фазовый элемент шины Φ_1 .

Напряжение на разделительном затворе Z_1 понижается, и между накопительными элементами и элементами фазового напряжения создается потенциальный барьер. Начинается следующий период накопления, в процессе которого весь процесс генерирования зарядового пакета повторяется и он перемещается в регистр ПЗС.

Такие генераторы характеризуются высокой интегральной чувствительностью, линейным преобразованием падающего излучения в зарядовый пакет, удовлетворительной скоростью параллельной передачи зарядовых пакетов, удовлетворительной скоростью параллельной передачи зарядовых пакетов из секции накопления в ПЗС-регистры, улучшением частотно-контрастной характеристики.

Конструкция генераторов неоднородностей в виде электрического домена (домена Ганна) весьма проста (рис. 4.4, а). На отрицательном электроде кристалла двухдолинного полупроводника, например арсенида галлия, возникает неоднородное распределение концентрации электронов в виде дипольного слоя. Такой полупроводник имеет N -образную характеристику. Слой образуется между слоями электронов и дырок. Между этими слоями создается дополнительное поле ΔE , которое добавляется к внешнему электрическому полю. Поле внутри домена становится больше, чем вне его (рис. 4.4, б). Если дифференциальное сопротивление меньше нуля, то ток уменьшается с ростом поля внутри диполя. Растет падение напряжения на дипольном слое и домен стабилизируется. Распределение зарядов представлено на рис. 4.4, в. Так как напряженность поля вне домена меньше, чем внутри, то новые домены не образуются. Домен образован электронами проводимости и движется в направлении их дрейфа со скоростью, близкой к скорости дрейфа $V_{др}$.

Новый домен может образоваться только после аннигиляции домена у анода. Размер домена составляет $10 \div 20$ мкм. Домены могут генерироваться в полупроводниках как с электронным, так и с дырочным типом проводимости. Форма колебаний тока представлена на рис. 4.4, г, а частота следования колебаний f пропорциональна отношению V/l , где V — скорость домена, l — длина образца.

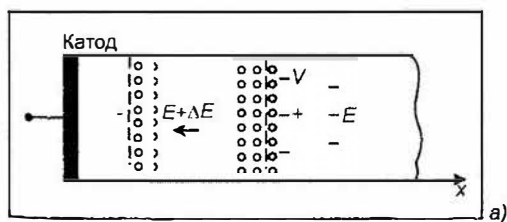
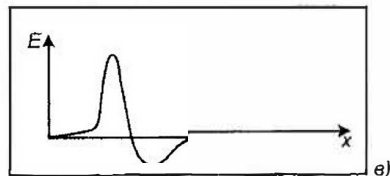
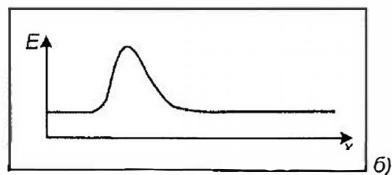


Рис. 4.4. Генератор доменов Ганна (а), распределение электрического поля E (б), объемного заряда ρ (в) в домене



Генерирование динамических неоднородностей в виде токовой неустойчивости возможно и в полупроводниках с S -образной вольтамперной характеристикой. В полупроводниковой структуре ток по сечению распределен неравномерно. Происходит образование об-

ластей с большей плотностью тока, чем в остальном сечении. Это так называемые токовые шнуры. Это явление характерно для всех приборов с ВАХ S -типа независимо от физического механизма их работы. В некоторых приборах это явление относится к паразитным, поскольку не используется полная площадь прибора. Увеличение тока через полупроводниковую структуру в области шунтирования происходит при постоянном напряжении. Шнур может перемещаться от места своего образования под действием физических полей различной природы. Подвижность шнура зависит от напряженности полей и может быть использована для переноса информационного сигнала.

В полупроводниковой структуре, обладающей S -образной ВАХ, распределенный p — n -переход находится одновременно в двух пространственно разделенных состояниях. Эти состояния определяются разными потенциалами смещения: прямое смещение под нелинейным контактом и обратное смещение на остальной части структуры. Получился бисмещенный переход с инжекционной неустойчивостью или, сокращенно, БИСПИН. Токовая неустойчивость возникает после достижения определенного порога генерации. Порог можно понизить освещением полупроводника или путем пропуска небольшого тока через распределенный p — n -переход.

Волны пространственного заряда генерируются в полупроводнике с объемной отрицательной дифференциальной проводимостью с помощью двух электродов. Электроды располагаются в области распространения волн и обеспечивают однородное электрическое поле, что является условием возникновения отрицательной дифференциальной проводимости (рис. 4.5, а). Значение напряженности электрического поля E лежит в пределах $(3—12) \times 10^3$ В/см. Обеспечение таких значений напряженности поля на больших длинах волн является самостоятельной задачей. Частотная зависимость фазовой скорости приведена на рис. 4.5, б. Длина когерентности волн пространственного заряда на несколько порядков может превышать длину волны. Величиной и вектором фазовой скорости волн пространственного заряда можно управлять путем изменения вектора скорости дрейфа электронов, другими словами, направлением электрического поля.

Таким образом, в основе всех рассмотренных конструкций генераторов динамических неоднородностей лежит идея использования электрических полей. Возбуждаемые динамические неоднородности можно отнести к электрической природе.

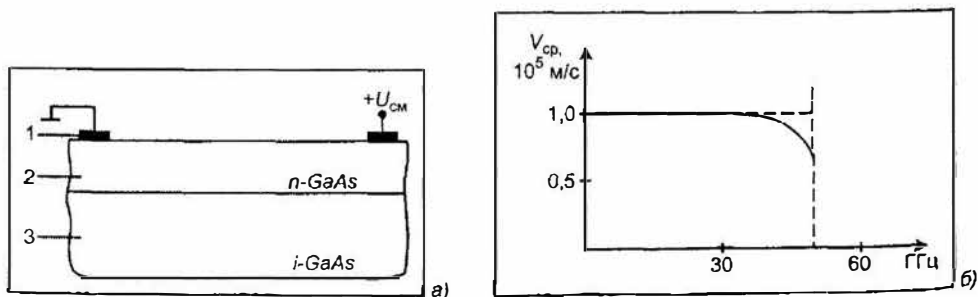


Рис. 4.5. Генератор волн пространственного заряда (а)

и частотная зависимость фазовой скорости волн (б): 1 — металлический контакт, 2 — легированный слой ($n \approx 10^{16} \text{ см}^{-3}$); 3 — собственный полупроводник; $U_{см}$ — напряжение смещения

Если же поместить полупроводник с разными концентрациями носителей тока в постоянное магнитное поле \vec{H} , то возникают спиральные волны — *геликоны*. Магнитное поле должно быть достаточно сильным, чтобы радиус орбиты носителей R был много меньше

длины свободного пробега носителей $R \ll l$. Геликоны представляют собой динамические неоднородности уже электромагнитной природы.

4.1.4. Устройства управления динамическими неоднородностями

Существуют различные способы управления распространением динамических неоднородностей в полупроводниках, но все они основываются на эффекте воздействия поля на динамические неоднородности. Рассмотрим некоторые примеры.

Подавая на каждый последующий электрод цепочки МОП-конденсаторов напряжение выше предыдущего, можно осуществить процесс перетекания заряда во все более глубокую потенциальную яму. Однако такой процесс не может быть долгим потому, что легко может быть получено пробивное напряжение. Для перемещения зарядовых пакетов удобно использовать импульсы напряжения трапецеобразной формы (рис. 4.6, б).

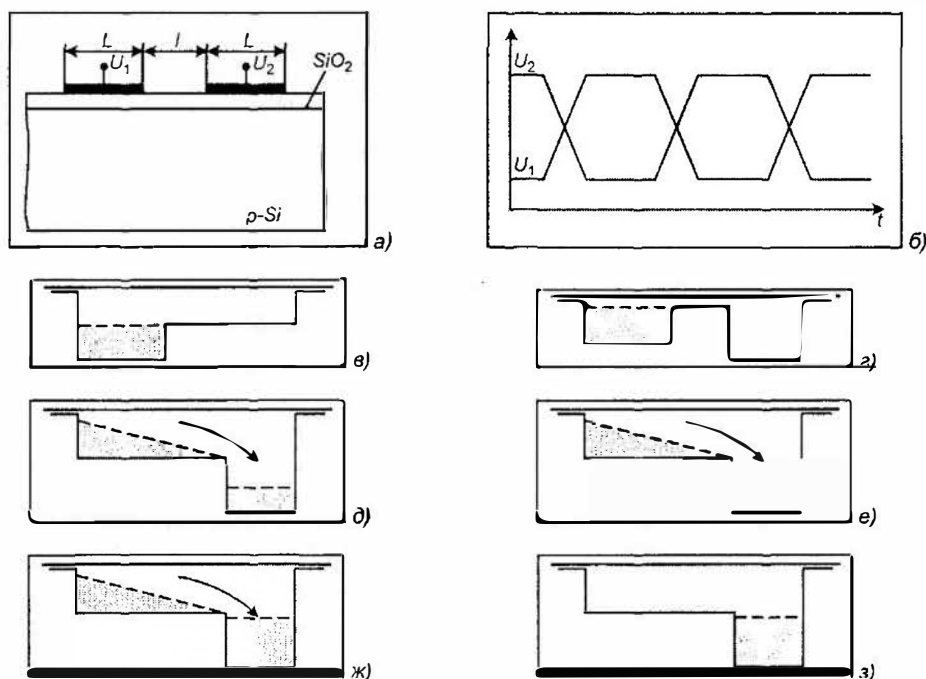


Рис. 4.6. Зарядовая связь между двумя МОП-конденсаторами: а — сечение МОП-конденсаторов; б — импульсы, подаваемые на электроды; в — зарядовый пакет в левой потенциальной яме; г — сформированная пустая потенциальная яма; д — перенос зарядового пакета самоиндуцированным дрейфом; е — перенос зарядов краевым полем; ж — перенос зарядов диффузией; з — зарядовый пакет в правой потенциальной яме

В момент подачи импульса напряжения на электрод 1, на электроде 2 напряжение мало. С уменьшением амплитуды тактового импульса U_1 увеличивается амплитуда импульса U_2 . Соответственно потенциальная яма под электродом 1 уменьшается, а под электродом 2 увеличивается (рис. 4.6 в, г). Зарядовый пакет из-под электрода 1 перемещается под

электрод 2. Возможно обратное перемещение заряда в случае обратной последовательности подачи напряжения, перенос зарядов осуществляется в три стадии.

В начальный момент, когда яма полностью заполнена и зарядовый пакет имеет максимальную величину, перенос происходит из-за электростатического расталкивания инверсионных электронов. Как только часть носителей попадает в потенциальную яму под второй электрод, возникает градиент плотности заряда и соответственно градиент электрического поля. Именно это поле вызывает ток носителей или самоиндуцированный дрейф (рис. 4.6, *д*). Величина заряда в первой яме уменьшается по гиперболическому закону. Время перетекания заряда составляет порядка 10^{-9} с.

С уменьшением зарядового пакета силы расталкивания ослабевают и самоиндуцированный дрейф замедляется. Начинается дрейф под действием краевого поля, обусловленного разностью потенциалов под соседними электродами. Величина краевого поля зависит от амплитуды тактовых импульсов. Дно потенциальной ямы под первым электродом наклоняется в сторону перетекания заряда. Угол наклона пропорционален напряженности краевого поля (рис. 4.6, *е*). По гидравлической аналогии это соответствует наклону дна сосуда. Основная часть зарядового пакета ($\sim 90\%$) переносится именно с помощью этого механизма (рис. 4.6, *ж*).

Это медленная стадия процесса переноса. Эффективность переноса определяется величиной

$$\eta = 1 - \frac{Q(T)}{Q_{\Sigma}},$$

где $Q(T)$ — величина зарядового пакета, успевшего перетечь за период такта, Q_{Σ} — величина полного зарядового пакета. При малом значении η происходит искажение информационного сигнала, что недопустимо в устройствах. Оценки показывают, что допустимая неэффективность переноса $\varepsilon = 1 - \eta \approx 10^{-5}$.

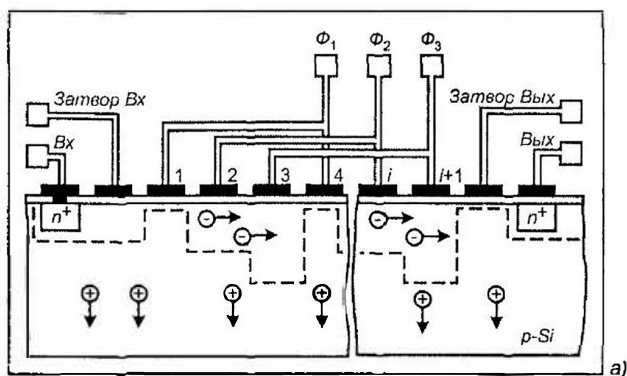
В момент минимального значения импульса U_1 яма под электродом 1 схлопнется, в то время как на электроде 2 будет максимальное значение амплитуды импульса, а под электродом будет локализован практически весь зарядовый пакет (рис. 4.6, *з*). Цепочка таких МОП-конденсаторов с зарядовой связью образует сдвиговый регистр, предназначенный для транспортировки и хранения зарядовых пакетов.

Управлять процессом передачи зарядов можно, управляя величиной на затворах зарядово-связанных МОП-конденсаторов. Различают несколько типов ПЗС-регистров, различающихся количеством фаз.

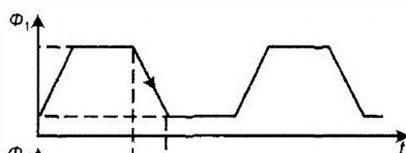
На рис. 4.7 представлена схема накопления и переноса зарядовых пакетов в трехфазном ПЗС. Последовательность импульсов на фазах Φ_1 , Φ_2 , Φ_3 имеет периодический характер и трапецевидную форму. Их временная диаграмма представлена на рис. 4.7, *в*, *г*, *д*. Импульсы следуют с некоторым временным перекрытием так, чтобы фронт последующего по времени импульса нарастал бы раньше, чем начался спад импульса предыдущей фазы. Заметим также, что импульсы имеют некоторое постоянное смещение $U_{см}$ (1—3 В), обеспечивающее постоянное обеднение поверхности основными носителями. Отсутствие такого напряжения смещения приводило бы к потерям величины зарядового пакета вследствие рекомбинации электронов с дырками. Напряжение на фазах ПЗС-структуры колеблется в пределах 10—20 В.

Формируя симметричные либо асимметричные топологии структуры, можно создать также одно-, двух-, четырехтактные сдвиговые регистры. Рассмотренный трехтактный ПЗС-регистр относится к первому типу, в котором направленность переноса зарядового

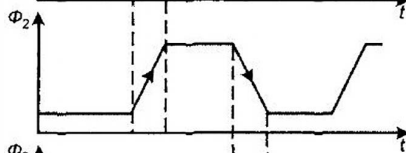
пакета обеспечивается индуцированными потенциальными барьерами. Эти барьеры формируются электрическими полями со стороны, противоположной переносу зарядового пакета.



б)



в)



г)

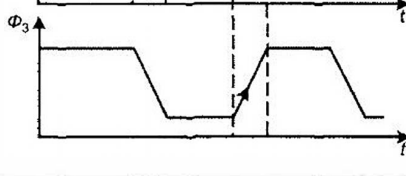


Рис. 4.7. Накопление и перенос зарядовых пакетов в трехфазном ПЗС (а) и диаграммы управляющих импульсов на фазах Φ_1 (б), Φ_2 (в), Φ_3 (г)

Направленность переноса зарядовых пакетов можно обеспечить с помощью технологически встроенных зарядовых барьеров. Такие конструкции относятся ко второму типу сдвиговых регистров.

Асимметричное распределение потенциала, обеспечивающее направление и управляемый перенос зарядовых пакетов, можно получить неоднородным распределением примесей под электродами, а также изменением толщины слоя диэлектрика.

В процессе управляемого переноса информационного сигнала существенную роль играют каналы переноса, которые можно классифицировать по признаку локализации.

Различают:

- *поверхностный канал*, расположенный у границы "диэлектрик—подложка", который обеспечивает высокую зарядовую емкость, простоту технологии. К недостаткам поверхностного канала следует отнести невысокое быстродействие, большие шумы, низкую эффективность переноса;

- *объемный канал*, расположенный на глубине 0,2 мкм, называют мелким, на глубине до 0,5 мкм — глубоким. Достоинством этого канала является более высокое, чем у поверхностных каналов, быстродействие, высокая эффективность переноса из-за устранения влияния поверхностных ловушек и поверхностного рассеяния носителей. К недостаткам относится небольшая зарядовая емкость и усложнение технологии производства;
- *перистальтический канал*, состоящий из неглубокого канала с высокой степенью легирования и находящегося под ним глубокого канала с небольшой степенью легирования. Такая комбинация позволяет достичь оптимального сочетания быстродействия и зарядовой емкости. Недостатком является существенное усложнение технологии производства;
- *многоканальные структуры* содержат много каналов переноса, позволяющие переносить носители зарядов противоположного знака. Такая сложная конструкция сложна в изготовлении, но расширяет функциональные возможности ПЗС-регистра;
- *канал с резистивным электродом* является типом канала, в котором перенос заряда осуществляется под действием в канале электрического поля переменной напряженности. Перенос осуществляется механизмом дрейфа носителей. Это является ограничивающим фактором на число электродов в регистре при достаточно простой его конструкции.

Направленный перенос зарядовых пакетов обеспечивается также *стоп-канальными* областями. Эти области формируются легированием подложки примесями того же типа проводимости, что и зарядовые пакеты, но с большей концентрацией (рис. 4.8). Стоп-каналы легируются до такой степени, чтобы состояние обеднения в них не достигалось при максимальном напряжении смещения.

С помощью стоп-канальных областей можно соединять и разделять каналы переноса, проводить другие операции управления.

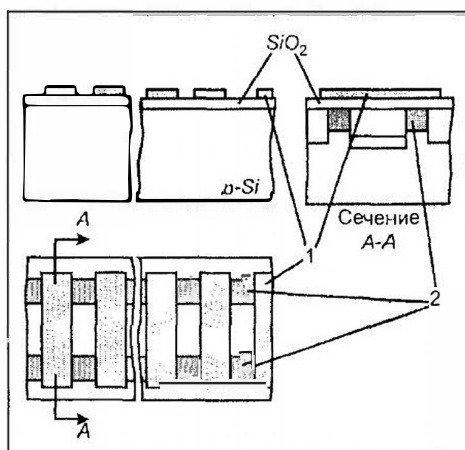


Рис. 4.8. Формирование стоп-канальных областей в ПЗС-регистре:
1 — металлический электрод;
2 — стоп-канальные области n^+ -типа проводимости

Рассмотрим другой пример — управление динамическими неоднородностями волновой природы. Величина и направление фазовой скорости волн пространственного заряда задается с помощью электрического статического тянущего поля $U_{ст}$. Это поле формируется между истоковым и стоковым омическими контактами в эпитаксиальном арсенид-гал-

лиевом слое n -типа проводимости. Управлять формой выходного высокочастотного сигнала можно с помощью встречно-штыревых преобразователей (ВШП). Их конфигурация позволяет отслеживать объемный заряд в полупроводниковой пленке, индуцирующий выходной сигнал.

Весьма перспективными являются полупроводниковые структуры, работа которых основана на управлении путем интерференции электронных волн. Если сгенерировать динамические неоднородности в виде электронных волн с высокой фазовой когерентностью на протяжении длины их распространения, то электрическим полем можно модулировать электронную волну.

В области детектора наблюдается конструктивная или деструктивная интерференция двух электронных волн в зависимости от напряжения на устройстве управления (рис. 4.9). Трудность реализации таких приборов заключается в малой длине когерентности электронных волн в твердом теле.

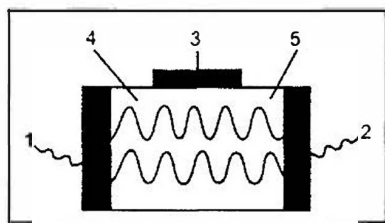


Рис. 4.9. Интерференция электронных волн в твердом теле: 1 — генератор; 2 — детектор электронных волн; 3 — электрическое устройство управления фазой электронной волны; 4 — когерентные электронные волны; 5 — электронные волны с противоположными фазами

4.1.5. Детекторы динамических неоднородностей

Детекторы динамических неоднородностей в полупроводниковых приборах функциональной электроники предназначены для считывания информации. При этом информационный сигнал может быть либо аналоговым, либо цифровым. В основе работы детекторов лежат различные процессы взаимодействия динамических неоднородностей с веществом или полями. Это могут быть различные электрические явления наведения тока или образования заряда, появление потенциала или наведение электрического или магнитного импульсного поля, возбуждение атомов или молекул. Электрически нейтральные динамические неоднородности могут быть зарегистрированы по вторичным заряженным частицам или продуктам взаимодействия, возникающим в результате взаимодействия с веществом.

Все детекторы динамических неоднородностей, как правило, являются элементами двоичной логики и выделяют логические ноль и единицу. Детекторы многозначной логики используются в более сложных комбинациях физических процессов и явлений.

Рассмотрим некоторые примеры детекторов динамических неоднородностей полупроводниковой природы. Простейшим является плавающий обратносмещенный p — n -переход (рис. 4.10).

С этой целью на выходе ПЗС-структуры создается плавающая диффузионная область (ПДО). Зарядку диффузионной области до плавающего потенциала производят путем подачи обратного смещения E на последнюю диффузионную область. Соединение осуществляется за счет создания индуцированного канала в транзисторе сброса T . С этой целью на его затвор Z_m подается импульс смещения, в течение которого плавающая диффузионная область заряжается от источника E (рис. 4.10, б) Транзистор сброса запирается

и ПДО остаются заряженной и изолированной (4.10, а). Приход зарядового пакета из-под затвора 3 вызовет изменение потенциала ПДО на величину ΔU , способную открыть выходной транзистор $T_{\text{вых}}$ (рис. 4.10, з). Это приведет к возникновению в выходной цепи токового импульса. После процесса считывания ПДО очищается от уже протектированного зарядового пакета путем подачи импульса на затвор и стока всего заряда.

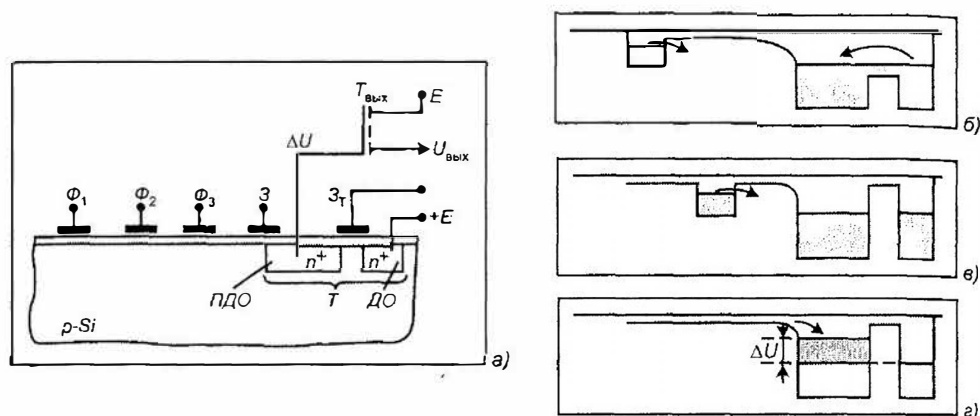


Рис. 4.10. Конструкция детектора зарядовых пакетов на основе ПДО (а), процесс зарядки ПДО до плавающего потенциала (б), ждущий режим ПДО (в), детектирование зарядового пакета и возникновение импульса тока (г)

Существуют и более сложные детекторы зарядовых пакетов, определяющих в системах многозначного информационного сигнала величину и знак заряда.

Детектирование динамических неоднородностей в виде волн пространственного заряда осуществляется с помощью металлического электрода, нанесенного на поверхность пленочных структур. Электрод располагается вблизи стока, на который подается напряжение смещения, вызывающее дрейф электронов от истока к стоку (от катода к аноду). Детектирующий электрод образует с полупроводниковой пленкой контакт типа барьера Шоттки. Такой емкостной зонд изолирован от пленки по постоянному току обедненной областью и способен детектировать, в том числе высокочастотные сигналы.

4.2. Приборы и устройства функциональной полупроводниковой электроники

4.2.1. Аналоговые процессоры на ПЗС-структурах

Аналоговые приборы предназначены для обработки информационных сигналов, заданных в виде непрерывных функций. Особый интерес представляют ПЗС-структуры.

Приборы с зарядовой связью используются в различных системах для хранения и обработки цифровой и аналоговой информации.

ПЗС позволяют осуществлять обработку цифровой и аналоговой информации, а именно: аналоговое и цифровое суммирование, деление и усиление сигналов; возможность неразрушающего считывания с необходимыми весовыми коэффициентами; возможность многократного ввода и вывода зарядовых пакетов.

Различают цифровые, аналоговые и фоточувствительные ПЗС.

К аналоговым ПЗС следует отнести линии задержки (ЛЗ), фильтры, аналоговые процессоры.

Линии задержки выполняют функцию задержки сигналов. Для приборов ПЗС линия задержки или регистр сдвига являются базовым элементом (рис. 4.11, а). Различают последовательные, параллельные и последовательно-параллельные линии задержки. В ЛЗ с последовательной организацией зарядовые пакеты последовательно сдвигаются через все N элементов ПЗС (рис. 4.11, б). Время задержки определяется тактовой частотой f_T и числом элементов N :

$$t_d = \frac{N}{f_T}.$$

Максимальная ширина полосы пропускания аналогового сигнала равна половине значения тактовой частоты $\Delta f = f_T / 2$. Такая простая организация ЛЗ при $N \approx 10^2 \div 10^3$ позволяет получить небольшую задержку сигнала.

Для получения большей задержки используется последовательно-параллельная организация ЛЗ (рис. 4.11, в). Информационный поток последовательно вводится с частотой f_T во входную последовательную секцию и после ее заполнения вся строка параллельно сдвигается в нижний регистр. В выходной секции зарядовые пакеты в последовательной форме передаются на выход, воспроизводя исходный сигнал.

Линия задержки на ПЗС-структурах позволяет регулировать задержку сигнала в пределах звуковых частот до частот телевизионных и радиолокационных сигналов.

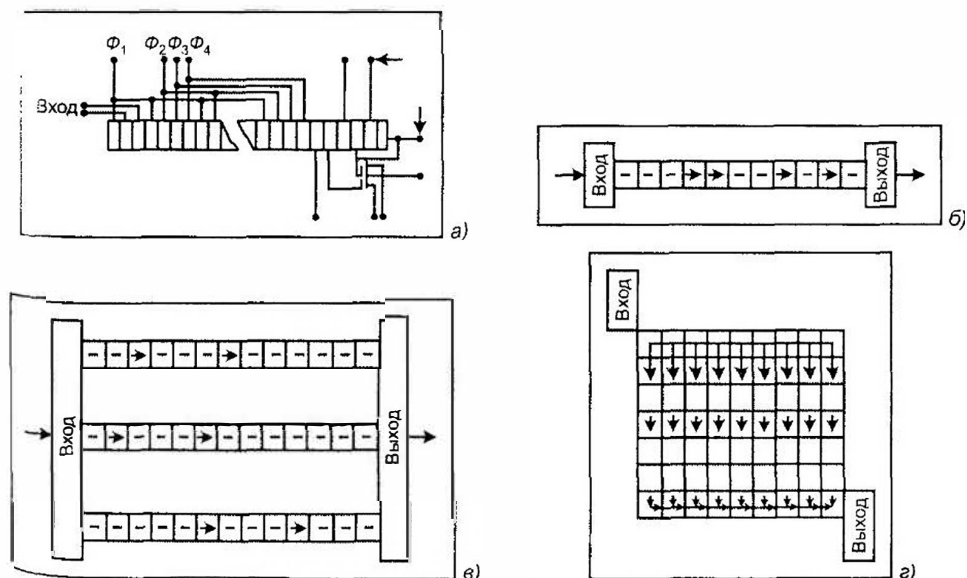


Рис. 4.11. Конструкция линий задержек на ПЗС

Фильтры на ПЗС являются дискретно-аналоговым прибором и служат для частотного деления сигналов, определенного полосой пропускания. Входной сигнал в них дискре-

тизируется во времени или представляется совокупностью значений (отсчетов, выборок). Выборки обрабатываются в фильтрах по определенным правилам.

Различают трансверсальные, рекурсивные фильтры, а также корреляторы, в которых сигнал обрабатывается по определенным правилам.

Трансверсальные фильтры на ПЗС реализуются на основе ЛЗ с отводами, имеющими различные весовые коэффициенты в соответствии с требуемой полосой пропускания (рис. 4.12, а). Взвешивание сигналов в фильтрах осуществляется с помощью разделения элементов на две части. Величина зарядового пакета делится пропорционально площади электродов. Суммирование взвешенных сигналов осуществляется с помощью тактовых шин, на которые подаются идентичные тактовые импульсы. Протекающие токи считываются раздельно, и разность сигналов усиливается дифференциальным усилителем (ДУ). Такая конструкция называется трансверсальным фильтром и имеет точность взвешивания 1—2%. Трансверсальный фильтр может производить операцию свертки $U_1(t) * U_2(t)$ входного сигнала $U_1(t)$ с импульсной характеристикой фильтра $U_2(t)$. Трансверсальный фильтр не имеет обратных связей.

Такой фильтр получается путем соединения выхода трансверсального фильтра через цепь обратной связи со входом. Рекурсивные фильтры на ПЗС содержат несколько ЛЗ и внешние обрамляющие схемы.

Корреляторы предназначены для перемножения двух аналоговых сигналов, например, входного сигнала $U(t)$ и опорного $U(t+r)$, т. е. $U(t) \times U(t+r)$. Коррелятор конструктивно состоит из двух линий задержки на ПЗС с отводами, подключенными к умножителям и сумматору (рис. 4.12, б). Один входной сигнал вводится в ЛЗ₁, и там хранится, другой — в ЛЗ₂. Сигналы с соответствующих вводов перемножаются, суммируются и на выходе формируется свертка двух сигналов.

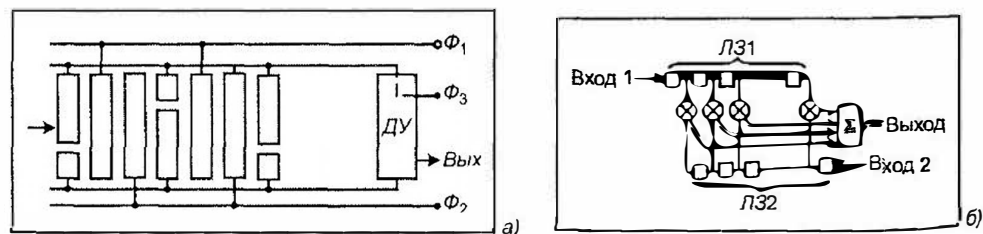


Рис. 4.12. Использование линий задержки в качестве фильтра (а) и коррелятора (б)

Тенденция развития фильтров на ПЗС будет идти по пути расширения универсальности, улучшения характеристик, а также максимальной интеграции электронного обрамления.

На аналоговых устройствах можно построить сложные многофункциональные системы обработки аналоговых сигналов — *процессоры*. Аналоговые процессоры являются высокопроизводительными, имеют малые габариты, массу и низкую потребляемую мощность, высокую надежность, характеризуются высокой степенью функциональной интеграции. К аналоговым процессорам относятся, например, программируемые трансверсальные фильтры. У таких фильтров весовые коэффициенты программируются с помощью микропроцессора.

Фоточувствительные приборы с зарядовой связью (ФПЗС) являются весьма перспективными процессорами.

ФПЗС является изделием функциональной электроники, предназначенным для преобразования оптического изображения в электрический сигнал, действие которого основано на формировании и переносе зарядовых пакетов под действием света по поверхности или внутри полупроводника.

Различают линейные и матричные ФПЗС.

В *линейных ФПЗС* фоточувствительные элементы расположены в один ряд. За один период интегрирования линейный ФПЗС воспринимает изображение и преобразует в электрический (цифровой) сигнал одну строку оптического изображения.

В режиме накопления информации в потенциальных ямах ФПЗС накапливаются фотогенерированные зарядовые пакеты, а затем эти фоточувствительные элементы используются для транспортировки зарядовых пакетов к выходу. Эта конструкция линейных ФПЗС проста, но в ней происходят искажения сигналов за счет засветки в процессе сканирования (рис. 4.13, а).

Целесообразно разделить функции накопления и сканирования. С этой целью вводят дополнительный регистр сдвига (рис. 4.13, б) и разрешающий затвор. Фотогенерированные зарядовые пакеты формируются в фоточувствительных элементах секции накопления. Затем на разрешающий затвор поступает отпирающее напряжение, и весь массив информации параллельно переносится в соответствующие элементы регистра сдвига, защищенного от света. На следующем периоде накопления формируется новая конфигурация зарядовых пакетов, а предыдущие пакеты передаются на выход. В таком приборе устраняется влияние засветки.

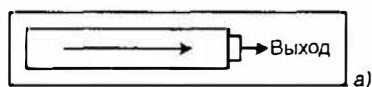
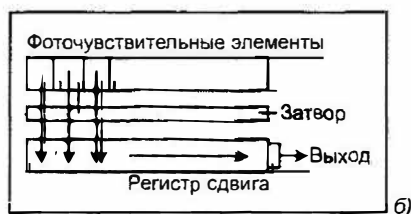


Рис. 4.13. Конструкции линейных ФПЗС:
а — без разделения областей накопления и передачи;
б — с дополнительным регистром сдвига



Линейные ФПЗС используются в системах, где движется объект либо сам прибор. В настоящее время разработаны линейные ФПЗС, имеющие 2048 элементов разложения с шагом 12 мкм, интегральную чувствительность $\sim 5 \times 10^2 \text{ В} \cdot \text{м}^2 / \text{Дж}$, динамический диапазон 60—70 дБ на частоте вывода информации 20 МГц.

Матричные ФПЗС представляют собой фоточувствительные приборы с переносом заряда, в которых фоточувствительные элементы организованы в матрицу по строкам и столбцам.

За один период интегрирования матричный ФПЗС преобразует в электрический сигнал один кадр оптического изображения. В этом приборе зарядовые пакеты передаются на выход путем перемещения потенциальных ям, вызываемого периодическим изменением амплитуды управляющих импульсов.

В конструкции ФПЗС с кадровой организацией различают секцию накопления, секцию хранения, входные и выходные регистры сдвига (рис. 4.14). Секции накопления и хранения образуются трехфазными поликремниевыми электродами, имеющими, например, 288 строк и 360 столбцов каждая. В малокадровом режиме изображение можно проектировать на обе секции. Все регистры имеют идентичные входные устройства, работающие по методу уравнивания потенциалов инжекции — экстракции.

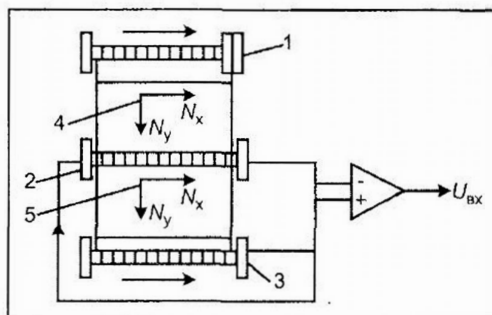


Рис. 4.14. Схема матричной ФПЗС с межкадровым вычитанием:
1, 2, 3 — регистры;
4 — секция накопления;
5 — секция хранения

Обычно обработка видеосигналов ведется внешними устройствами на дискретных элементах. В приборах с зарядовой связью заложены возможности обработки видеосигнала непосредственно в кристалле. В этом случае можно говорить о встроенном процессоре, который может производить межкадровую обработку видеосигнала, подчеркивание контуров, устранение точечных дефектов, распознавание образов. Другими словами, процессор представляет собой устройство функциональной электроники, осуществляющее процесс одномоментной обработки большого информационного массива в реальном масштабе времени.

С целью выделения информации о движущихся объектах на сложном стационарном фоне стараются подавить изображение его стационарных частей в обоих полукадрах.

Режим работы ФПЗС состоит из следующих циклов. Информация в виде i -го кадра накапливается в секции накопления, а затем переносится в секцию хранения; информация $(i + 1)$ кадра накапливается в секции накопления и происходит синхронная передача зарядов i -го и $(i + 1)$ -го кадров на выход через регистры 2 и 3.

Таким образом, матричная структура запоминает оба кадра видеoinформации и одновременно выводит их на дифференциальный усилитель, в котором происходит поэлементное межкадровое вычитание. Операция межкадрового вычитания может быть представлена в виде:

$$\Delta U^i(t) = U^i(t + nT_K) - U^i(t), \quad (4.2)$$

где $U^i(t)$ — видеосигнал i -го элемента текущего кадра; $U^i(t + nT_K)$ — видеосигнал i -го элемента, задержанный на время, кратное времени кадра T_K ; n — целое число; $\Delta U^i(t)$ — видеосигнал межкадровой разности i -го элемента. Таким образом, реализован режим обработки разностной информации в большом информационном массиве.

Существует также *фоточувствительный прибор с зарядовой инжекцией* (ФПЗИ), в котором перемещение зарядового пакета происходит внутри фоточувствительного элемента с последующей инжекцией в подложку или область стока заряда. Такие структуры имеют следующие достоинства: практическое отсутствие потерь передачи; возможность использования в качестве фоточувствительных элементов фотодиодов, имеющих большую по сравнению с МДП-структурами чувствительность; возможность организации произвольной выборки любого элемента или группы элементов.

Изготовление ФПЗС можно осуществлять по двум технологиям.

В приборах с поверхностным каналом перенос сигнальных зарядовых пакетов осуществляется на границе раздела "окисел — полупроводник". В этом типе ФПЗС легко управлять переносом, однако взаимодействие зарядовых пакетов с поверхностным состоянием при-

водит к захвату и эмиссии носителей. Все это приводит к снижению эффективности переноса носителей при низких уровнях освещенности.

В технологии с объемным каналом переноса зарядовых пакетов предусмотрено заглубление канала путем его ионного легирования примесью с проводимостью, противоположной подложке. Этот технологический прием существенно увеличивает эффективность переноса, но несколько снижает управляемость.

В последнее время весьма перспективны ФПЗС для цветного телевидения, конструктивно оформленные в виде строчных или строчно-кадровых матриц с числом элементов 500×500 и совмещенных с цветокодирующим фильтром.

На основе ФПЗС разрабатывается серия многоэлементных приемников изображения средневолнового и длинноволнового ИК диапазона длин волн. В их основе гибридные структуры с интегральными элементами считывания и предпроцессорной аналоговой обработкой видеосигнала на кристалле.

Тепловизионные системы на основе ИК ПЗС по пороговой температурно-пространственной чувствительности и быстродействию не уступают лучшим образцам современных тепловизоров.

4.2.2. Цифровые процессоры на ПЗС-структурах

Цифровые ПЗС предназначены для обработки сигналов в виде дискретных функций, и их можно разделить на структуры для арифметико-логической обработки информации и структуры для хранения информации — запоминающие устройства.

В цифровые ПЗС можно включить регистры сдвига, логические и арифметические устройства, запоминающие устройства. В них информация представляется двумя уровнями зарядов, локализованных в потенциальной яме под затвором. Состояние логической единицы определяется максимальным зарядовым пакетом в яме, состояние логического нуля — отсутствием заряда или величиной фонового заряда. Следует заметить, что цифровые устройства ПЗС являясь устройствами динамического типа, т. е. необходима регенерация информации.

К основным аналоговым операциям относятся сложение, вычитание, умножение, умножение на весовой коэффициент, дифференцирование, интегрирование. Операцию аналогового сложения в зарядовой форме можно произвести различными способами, зависящими от конструкции ПЗС. В схемах параллельного сложения зарядовые пакеты величин Q_1 , Q_2 и Q_3 передаются по n параллельным каналам ПЗС (рис. 4.15, а). Суммирующий элемент должен иметь зарядовую связь с элементами цепочек. Суммирование осуществляется в процессе слияния всех зарядов q_i из параллельных цепочек в потенциальную яму суммирующего элемента, т. е.

$$q_1 + q_2 + \dots + q_n = Q_{\Sigma} < C_{\Sigma}(U_{\phi} - U_{cm}), \quad (4.3)$$

где Q_{Σ} — суммарный заряд, который должен уменьшаться в потенциальной яме суммирующего элемента емкостью C_{Σ} .

В схеме последовательного сложения между суммирующими последними тактовыми электродами расположен экранирующий электрод, создающий потенциальный барьер и препятствующий обратной передаче зарядов из суммирующего элемента (рис. 4.16, б). На суммирующий затвор подается импульс Φ_{Σ} , длительность которого больше периода следования тактовых импульсов. На срезе распределения фазы Φ_i очередной зарядовый па-

кет перетекает через экранирующий элемент в потенциальную яму суммирующего элемента и складывается с ранее переданными зарядами. Зарядовая емкость суммирующего элемента определяется условием:

$$C_{\Sigma}(U_{\phi} - E_3) \geq \sum_{i=1}^n q_i. \quad (4.4)$$

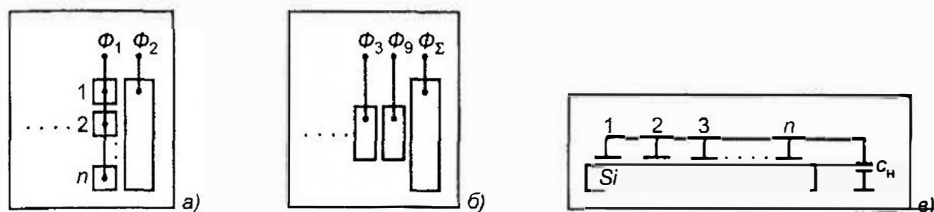


Рис. 4.15. Схемы суммирования зарядовых пакетов: а — параллельный; б — последовательный; в — на плавающих затворах

Аналогичным способом зарядовые пакеты можно суммировать в плавающей диффузионной области. В каждый элемент поступает определенный зарядовый пакет. На общей шине находится емкостная нагрузка. Используя принцип суперпозиции, можно вычислить изменение потенциала при суммировании зарядового пакета на общей шине нагрузочной емкости.

Полное изменение потенциала при поступлении зарядовых пакетов во все элементы определяется выражением:

$$\Delta U = \sum_{i=1}^n \Delta U_i = -\frac{1}{C_{\Sigma}} \sum_{i=1}^n q_i. \quad (4.5)$$

Таким образом, выходной сигнал на плавающей шине будет пропорционален сумме зарядовых пакетов. Этот прием позволяет производить суммирование без разрушения зарядового пакета.

Операцию вычитания аналоговых сигналов можно производить различными способами (рис. 4.16). Одним из способов является формирование пакета заряда, пропорционального разности напряжений на электродах (рис. 4.16, а). В этом случае используются перекрестные связи затворов, которые управляются напряжениями аналоговых сигналов U_1 и U_2 . Если $U_1 > U_2$, то в верхнюю цепочку элементов заряд не поступает вследствие более глубокой потенциальной ямы в первом элементе верхней цепочки по сравнению со второй. Таким образом, после этапа экстракции электроны будут экстрагированы обратносмещенной n^+ -областью истока. В нижнюю цепочку будет введен зарядовый пакет величиной

$$q_2 = S \cdot C \cdot (U_1 - U_2),$$

$$q_1 = 0.$$

Для случая $U_1 < U_2$ заряд вводится только в верхнюю цепочку, т. е.

$$q_1 = S \cdot C \cdot (U_2 - U_1),$$

$$q_2 = 0.$$

Итак, разность двух сигналов $U_2 - U_1$ формируется в нижней цепочке, если разность положительная, и в верхней — при отрицательной разности. При считывании сигналов из

обеих цепочек можно определить как разность потенциалов $U_1 - U_2$, так и знак этой разности.

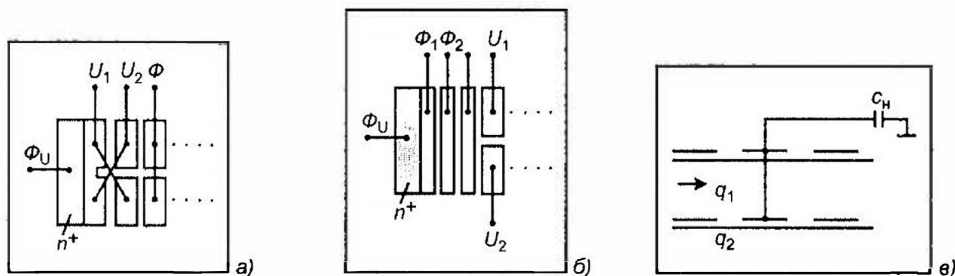


Рис. 4.16. Схемы вычитания зарядовых пакетов: а — с перекрестными связями; б — с зарядом смещения; в — на плавающих затворах

Другой способ предусматривает представление входных сигналов в виде зарядовых пакетов. Разностный сигнал преобразуется в нужную форму: в виде заряда или напряжения (рис. 4.17, б). Во входной элемент инжекции — экстракции вводится фиксированный зарядовый пакет q_0 . Этот заряд делится между двумя элементами, затворы которых находятся под напряжением U_1 и U_2 .

При достаточно медленном спаде напряжения на шине Φ_2 соблюдаются квазистатические условия передачи. Это означает, что поверхностный потенциал шины Φ_2 и электродов U_1 и U_2 одинаковы и $\phi_1 = \phi_2$. Если

$$S_1 = S_2 = S,$$

$$q_1 + q_2 = q_0,$$

то

$$q_1 = 1/2 \cdot [q_0 + S \cdot C_H (U_1 - U_2)],$$

$$q_2 = 1/2 \cdot [q_0 + S \cdot C_H (U_2 - U_1)].$$

Таким образом, в верхней и нижней цепочке электродов будут передаваться заряды, отличающиеся от половинного заряда $q_0/2$ на величину, пропорциональную разности входных сигналов ($U_1 - U_2$). При считывании выходной сигнал будет смещен относительно среднего уровня, соответствующего $q_0/2$ (при $U_1 = U_2$), на величину, пропорциональную $U_2 - U_1$.

Операцию вычитания можно реализовать на плавающих затворах (рис. 4.16, в). При этом используется эффект умножения потенциала затвора при передаче зарядового пакета в элемент с плавающим затвором. Если объединить два плавающих затвора и в один ввести заряд q_1 , а из другого вывести заряд q_2 , то заряд q_1 вызовет понижение потенциала, а выводимый заряд q_2 — повышение потенциала плавающего затвора. Результирующее изменение потенциала плавающей шины будет пропорционально разности зарядовых пакетов, передающихся по верхним и нижним цепочкам.

Операцию дифференцирования можно реализовать путем соединения плавающих затворов двух соседних разрядов (рис. 4.17, а). При этом зарядовый пакет $Q_n(nT)$ из элементов 1 экстрагируется и вызывает увеличение потенциала U_A , одновременно в элемент 2 вводится зарядовый пакет $Q_n((n+1)T)$, соответствующий следующей выборке сигнала. На плавающей шине формируется сигнал, пропорциональный разности двух зарядов или производной передаваемого аналогового сигнала.

Операция интегрирования сигнала является обратной операции дифференцирования и осуществляется суммированием (накоплением) зарядовых пакетов.

На ПЗС-структурах выполняют аналогово-цифровые и цифроаналоговые процессоры, мультиплексоры и демультиплексоры.

Мультиплексор, например, служит для временной задержки и накопления информации (рис. 4.17, б). Тактовая частота подбирается таким образом, чтобы задержка между отдельными входами была равна времени перемещения изображения между входными приемниками. При этом происходит когерентное сложение зарядовых пакетов. Важно отметить, что шумы складываются некогерентно.

Логические ячейки и операции могут быть реализованы с помощью комбинации уже рассмотренных устройств, предназначенных для арифметических операций.

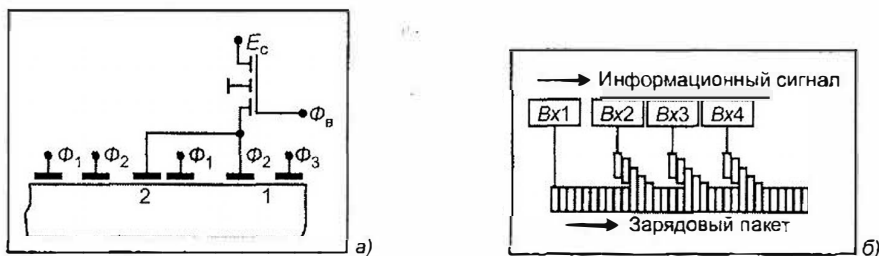


Рис. 4.17. Схемы дифференцирования и умножения сигналов (а) и мультиплексора (б)

Подадим на входные электроды в отдельности или на оба сразу напряжение, соответствующее логической единице. Возникающий зарядовый пакет также будет иметь величину, соответствующую логической единице (рис. 4.17). С любого из входов X_1 или X_2 , благодаря зарядовой связи перемещается под общий электрод и далее индицируется как логическая единица. Если же логическая единица подается сразу на оба входа, то на общем электроде будет удвоенный заряд. Для преобразования такого удвоенного зарядового пакета в зарядовый пакет, соответствующий логической единице, формируется регулируемый сток заряда. Он отделяется от потенциальной ямы под общим электродом с помощью потенциального барьера в виде ионно-легированной области.

Высота барьера выбирается такой, чтобы в потенциальной яме оставался зарядовый пакет, строго соответствующий логической единице. Стоком является обратномещенный р—п-переход.

На рис. 4.18, а представлена логическая ПЗС-ячейка, выполняющая операцию логического сложения (ИЛИ).

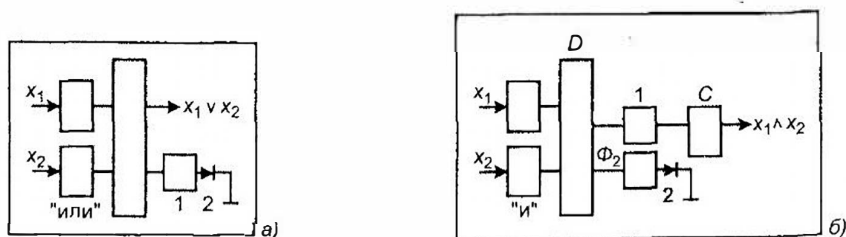


Рис. 4.18. Логические ПЗС-ячейки: а — схема реализации функции ИЛИ; б — функции И

Логическая ячейка для операции логического умножения (И) строится аналогично ячейке типа ИЛИ. За время действия напряжения фазы Φ_1 потенциальные ямы образуются под электродами D и C . Под электрод C заряд попадает только в том случае, если на оба входа X_1 и X_2 подаются одновременно зарядовые пакеты, соответствующие логической единице. По окончании действия фазового напряжения Φ_1 начинает действовать напряжение Φ_2 . Этим напряжением зарядовый пакет перемещается на выход из-под электрода C , а из-под электрода D переводится в обратномсмещенную область и затем экстрасигуруется.

По такому принципу конструируются логические ячейки, выполняющие комбинацию логических функций. Малые размеры логических ячеек позволяют формировать сложные арифметико-логические устройства.

4.2.3. Запоминающие устройства на ПЗС-структурах

Запоминающие устройства на принципах зарядовой связи легко реализуются с помощью ячеек аналоговой памяти. По существу эти ячейки являются конденсаторными МОП-структурами, в которых хранится зарядовый пакет или отсутствует вовсе. В первом случае говорят о хранении логической единицы, в другом — логического нуля. Объединенные в линейки такие ячейки образуют регистры сдвига.

Постоянное запоминающее устройство (ПЗУ) состоит из накопителя с однократной записанной информацией и системы вывода информации на регистрах с зарядовой связью.

Один из вариантов ПЗУ на ПЗС для хранения аналоговых сигналов содержит регистр сдвига, связанный с МОП-конденсаторами, размеры которых соответствуют записанной информации. На рис. 4.19, *а* изображен отдельный элемент такого устройства, где трехтактный регистр P_1 — P_3 непостоянно связан с конденсаторами постоянной памяти K_1 — K_3 . В режиме считывания на выбранный конденсатор подается напряжение и под его электродом накапливается заряд, пропорциональный площади последнего. Затем напряжение подается на выбранный электрод, и заряд передается в регистр. Это выполняется одновременно во всех элементах, так что регистр принимает параллельный код, который затем поступает к выходному усилителю.

Элементы хранения отделены от регистра, подобно фоточувствительным элементам в фотоприемниках со строчной разверткой. Однако они могут быть совмещены с электродами переноса, как показано на рис. 4.19, *б*. При этом информация представлена величиной заряда, встроенного в электродный диэлектрик. Наличие такого заряда приводит к модуляции поверхностного потенциала вдоль канала ПЗС, что позволяет считывать записанный заряд.

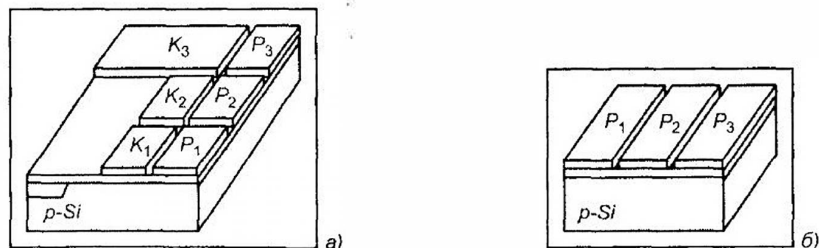


Рис. 4.19. ПЗУ на ПЗС: *а* — элемент памяти вне регистра; *б* — элементы памяти, совмещенные с электродами переноса

Операция считывания протекает в две стадии: вначале производится передача информации из постоянного накопителя в регистры ПЗС, затем — перенос зарядовых пакетов к выходу устройства. Первая стадия зависит от способа представления информации в накопителе, который при соответствующих управляющих сигналах определяет процесс генерации зарядовых пакетов, степень заполнения ими потенциальных ям или распределение зарядов между соседними элементами.

Информация записывается в диэлектрике, что обуславливает различное пороговое напряжение на электродах и одновременно различную управляющую способность. При подаче напряжения на электроды выбранной фазы под ними образуются потенциальные ямы различной глубины. Заполнение их зарядом, за счет естественной термогенерации, создаст зарядовый рельеф, отображающий ранее записанную информацию. Для ускорения заполнения ям "опрашивающие" заряды подводят с помощью самого ПЗС.

Репрограммируемое постоянное запоминающее устройство (РПЗУ) на ПЗС позволяет перезаписывать информацию в процессе работы в соответствии с введенным в него массивом зарядовых пакетов. Из некоторых способов хранения информации перепрограммирование допускает лишь заряд, встроенный в диэлектрик, но сам диэлектрик должен быть усложнен: диэлектрик с захватом заряда и диэлектрик со встроенным плавающим затвором.

Запоминающее устройство с произвольной выборкой по считыванию (ЗУПВ) на ПЗС позволяет считывать информацию из произвольной ячейки матрицы. В их основе лежит матрица ПЗС с дополнительной системой электродов, обеспечивающих произвольную выборку. Как известно, произвольная выборка возможна и в приборах с инжекцией заряда, но зарядовое считывание в матрицах большого объема дает очень слабые выходные сигналы, налагает чрезмерно жесткие требования на ее элементы, в особенности на усилители считывания. Более перспективным представляется использование токового считывания. В этом случае заряд, хранимый элементом, модулирует протекающий в нем ток.

Допустимы различные способы токового считывания при ориентации канала транзистора в различных направлениях по отношению к каналу ПЗС (рис. 4.20): в горизонтальном (а, б) и в вертикальном (в, г).

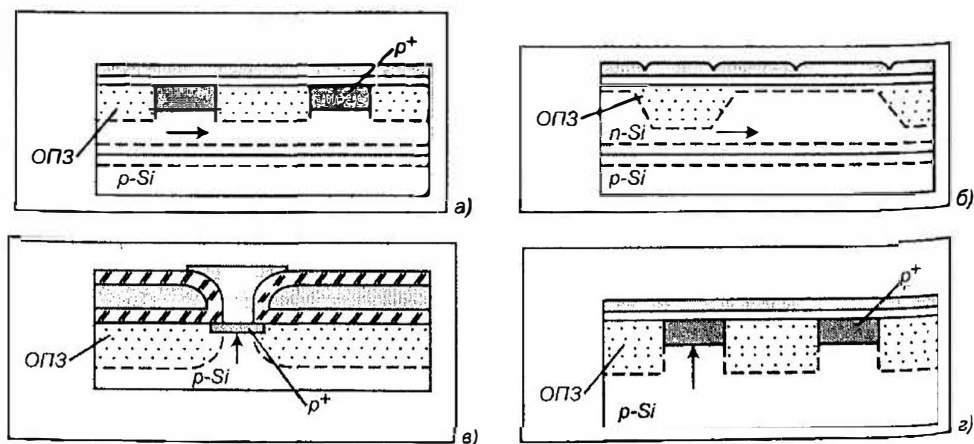


Рис. 4.20. Комбинация элемента ПЗС с транзистором: горизонтальный канал → транзистора поперек (а) и вдоль канала ПЗС (б); вертикальный канал ↑ транзистора со стоком в отверстии электрода ПЗС (в) и в виде изолирующей диффузионной области (г)

Области стока и истока канального транзистора могут быть включены в конструкцию дополнительно, но можно использовать и структуры, входящие в ПЗС-подложку, например, области изолирующей стоп-диффузии. При горизонтальном канале на подложке необходимо наличие слоя противоположного типа проводимости.

Для вертикального канала размеры стока должны быть меньше глубины области пространственного заряда.

Итак, ЗУ строятся на основе регистров сдвига, позволяющих реализовать в ячейках два устойчивых состояния "1" или "0" и перемещать информацию по регистру. По методу поиска адреса ячейки памяти ЗУ делятся на устройства с произвольным и последовательным доступом информации. ЗУ с произвольным доступом позволяют использовать в любой момент времени любую ячейку и имеют число выходов блока выборки, равное числу адресов блока запоминания. ЗУ на основе регистров сдвига на ПЗС относятся к устройствам с последовательным доступом к информации.

Для увеличения времени хранения информации зарядовые пакеты непрерывно циркулируют по ПЗС-структуре путем передачи их с выхода на вход через устройство регенерации, позволяющее компенсировать утечку и рассасывание зарядов.

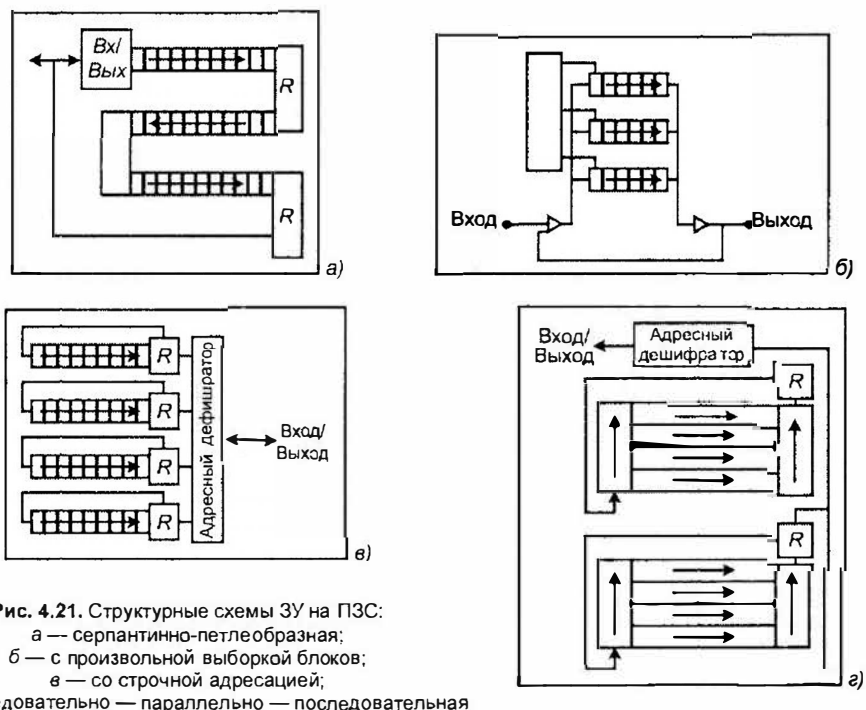


Рис. 4.21. Структурные схемы ЗУ на ПЗС:

а — серпантинно-петлеобразная;

б — с произвольной выборкой блоков;

в — со строчной адресацией;

г — последовательно — параллельно — последовательная

Конструктивно ЗУ организуются по различным схемам (рис. 4.21). Регистры сдвига могут быть объединены в виде серпантинно-петлевой структуры (рис. 4.21, а) с элементами регенерации R . Перспективными также являются структуры ЗУ с произвольной выборкой (рис. 4.21, б) и со строчной адресацией (рис. 4.21, в). Структуры с последовательно — параллельно — последовательной адресацией (ППП) образуют регистр, состоящий

из матрицы параллельно включенных регистров сдвига с последовательным переносом двух регистров — входного и выходного. Информация последовательно заносится в верхний регистр, затем переносится параллельно в матрицу и в ней построчно сдвигается. Достигнув выходного регистра, информация заносится в него, а затем последовательно сдвигается к выходному устройству. Структура типа ППП должна содержать элементы регенерации R , восстанавливающие уровень зарядовых пакетов.

ЗУ на ПЗС большой емкости строятся по блочной структуре или с блочной выборкой.

Основной областью применения ЗУ на ПЗС является буферная память. Так замена буферной памяти на магнитных дисках в ЭВМ позволит резко сократить габариты, потребляемую мощность, другие технические характеристики.

Плотность записи информации в ЗУ на ПЗС может достигать 10^5 бит/см² при скорости записи и считывания 5 Мбит/с. К недостаткам ЗУ на ПЗС следует отнести ограниченное время хранения информации в ЗУ, а также энергозависимость схем.

4.2.4. БИСПИН-приборы

БИСПИН-приборы представляют собой полупроводниковые структуры с БИСмещенным Переходом и Инжекционной Неустойчивостью. Эти приборы преобразуют входной аналоговый сигнал в регулярную последовательность импульсов тока или напряжения.

В качестве непрерывной среды используется полупроводниковая структура с распределенным $p^+ \cdots n$ - или $n^+ \cdots p$ -переходом (рис. 4.22, а).

На поверхности слаболегированной n - или p -области формируют нелинейный A и омический B контакты. В качестве нелинейного контакта может выступать встроенный локальный $p^+ \cdots p$ - или $p^+ \cdots n$ -транзистор.

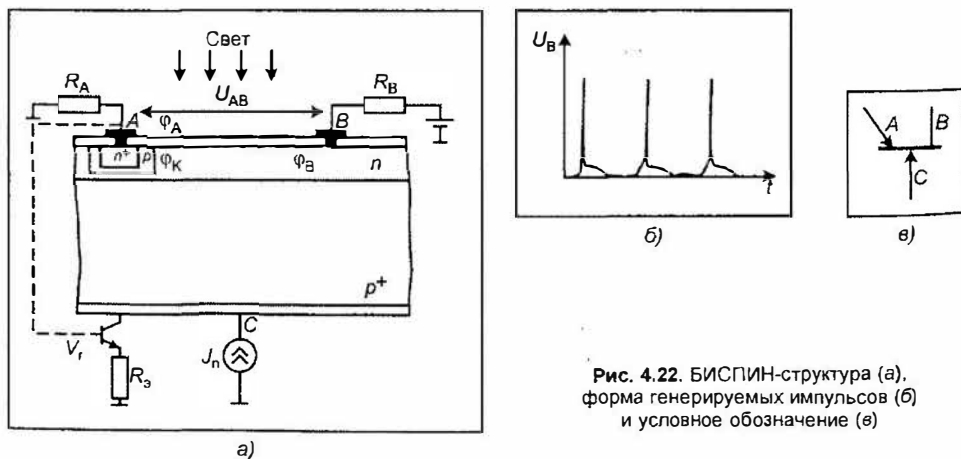


Рис. 4.22. БИСПИН-структура (а), форма генерируемых импульсов (б) и условное обозначение (в)

На поверхности сильно легированной p^+ - или n^+ -подложки формируется омический контакт C . Если к контактам A и B приложить определенную разность потенциалов U_{AB} таким образом, чтобы n^+ -область была включена в прямом направлении, а затем структуру осветить со стороны слаболегированной области, то во внешней цепи появляются импульсы тока определенной частоты следования. При большой скважности импульсы тока имеют

релаксационную форму (рис. 4.22, б). В промежутках между импульсами тока на контакте C имеют место пилообразные импульсы напряжения. Аналогичная генерация проявляется и без подсветки, но при пропускании тока через распределенный $p^+ - n$ -переход. Увеличение интенсивности подсветки или уровня тока в подложку приводит к пропорциональному росту частоты следования импульсов без изменения их амплитуды, величина которой определяется разностью потенциалов U_{AB} на структуре. Таким образом, БИСПИН может рассматриваться как релаксационный генератор, имеющий три независимых канала управления: потенциальный, токовый и гальванически развязанный — световой. Каждый из каналов характеризуется порогом генерации, величина которого зависит от состояния двух других каналов, и с их помощью легко изменяется в большую или меньшую сторону.

В зависимости от напряжения питания, уровня освещенности или тока в подложку БИСПИН может находиться в трех состояниях:

- закрытое, с низкой проводимостью, когда внутреннее сопротивление структуры между контактами A и B определяется сопротивлением обратносмещенного коллекторного перехода локального транзистора и составляет $\sim 10^8$ Ом;
- открытое, с высокой проводимостью, когда сопротивление структуры не превышает единиц кОм. В это состояние структура переходит при больших уровнях токов в подложку или подсветку;
- автоколебательное.

Из сказанного выше следует, что БИСПИН может работать как ключ, если время нарастания подаваемого на его вход сигнала не превышает периода следования импульсов при максимальной частоте повторения.

Физический механизм работы БИСПИНа в автоколебательном режиме состоит в следующем. После включения источника напряжения благодаря тому, что сопротивление обратносмещенного коллекторного перехода локального транзистора существенно превышает сопротивление R_n n -слоя, коллекторный потенциал ϕ_K и потенциал под омическим контактом ϕ_B совпадают с точностью до 10^{-3} В, т. е. $\phi_K \approx \phi_B$. Такого же уровня достигает потенциал подложки ϕ . Следовательно, распределенный $p^+ - n$ -переход находится под нулевым смещением. При включении света за счет внутреннего фотоэффекта происходит зарядка распределенного $p^+ - n$ -перехода и возникает инжекция дырок из подложки в n -слой. При этом часть дырок попадает в коллекторную область встроенной локальной $n^+ - p - n$ -структуры. Для дырок в n -слое коллекторный переход является экстрактором, благодаря чему они поступают в p -базу локального транзистора и скапливаются в ней. Последнее обстоятельство приводит к понижению потенциального барьера эмиттерного $n^+ - p$ -перехода и инжекции электронов из n -эмиттера в узкую p -базу, после пролета которой они оказываются в n -слое структуры. Возникает электронный ток I_n , протекающий по n -слою вдоль распределенного $p^+ - n$ -перехода к омическому контакту. Следует иметь в виду, что этот электронный ток во много раз больше тока дырок, поступающих в p -базу локального транзистора.

Падение напряжения $I_n R_n$ на распределенном сопротивлении n -слоя ($R_n \approx 3 - 5$ кОм) приводит к снижению потенциала ϕ_K вблизи коллекторного перехода

$$\phi_K = U_{AB} - I_n R_n.$$

Поскольку потенциал подложки ϕ вследствие ее высокого уровня легирования можно считать независимым от координаты, то смещение на распределенном $p^+ - n$ -переходе вдоль него изменяется. Прямое смещение этого перехода под коллектором локального

транзистора увеличивается, а в остальной, большей его части, сначала уменьшается, а затем становится отрицательным. Инжекция дырок из p^+ -области в p -базу транзистора при этом растет, что в свою очередь вызывает еще больший перекося в смещении p^+ — n -перехода и т. д. Процесс развивается лавинообразно и приводит транзистор в режим насыщения.

Вольтамперная характеристика структуры имеет S -образный вид, и в этом случае следует ожидать токовую неустойчивость.

При небольших уровнях фототока основной поток дырок в p -область локального транзистора поступает из p^+ -подложки за счет двух процессов: перераспределения фототока и перезарядки барьерной емкости p^+ — n -перехода. В начале процесса дырки, образующиеся вследствие внутреннего фотоэффекта, захватываются полем объемного заряда p^+ — n -перехода и перебрасываются в p^+ -подложку, благодаря чему он весь равномерно переходит в состояние прямого смещения (до 0,3 В). Вследствие этого дырки также равномерно инжектируются в n -слой по всей площади распределенного перехода, и в базу транзисторной структуры поступает лишь небольшая их часть, определяемая величиной площади этого перехода, расположенной под коллектором локальной структуры. Вследствие развивающейся описанной выше петли положительной обратной связи (включающей ток вдоль n -слоя, падение потенциала Φ_k и рост прямого смещения на p^+ — n -переходе) распределенный переход оказывается бисмещенным, происходит перераспределение инжекционного тока дырок, в результате чего все фотодырки, собранные основной обратносмещенной частью p^+ — n -перехода, поступают через прямосмещенную часть в базу локальной транзисторной структуры. К этому току добавляется ток перезарядки емкости p^+ — n -перехода через открытую вертикальную n^+ — p — p^+ -структуру. Оценки показывают, что начальный всплеск тока в импульсе носит емкостный характер, и его амплитуда линейно зависит от разности потенциалов на структуре.

После прекращения перезарядки p^+ — n -перехода количество дырок, поступающих в базу локальной транзисторной структуры, резко сокращается и остается только поток "фотодырок". Концентрация их в базе из-за рекомбинации резко сокращается и все большая часть коллекторного перехода выходит из смещения, сопротивление растекания коллектора увеличивается, коллекторный ток снижается и площадь прямосмещенной части p^+ — n -перехода уменьшается.

Если фототок настолько велик, что ток вдоль p^+ — n -перехода обеспечивает сохранение на нем прямосмещенного участка, то структура остается открытой. В противном случае она переходит в закрытое состояние, и весь процесс повторяется (автоколебательный режим).

Таким образом, в процессе развития импульса p^+ — n -переход проходит четыре состояния: с нулевым смещением, прямосмещенное, обратносмещенное и бисмещенное. Последнее обстоятельство послужило основанием для названия структуры и прибора.

Возможности практического применения БИСПИНов определяются их многофункциональностью, высокой чувствительностью по входу; большой амплитудой выходного сигнала; малыми темновыми токами; широким диапазоном перестройки; большой крутизной преобразования "аналоговый сигнал—частота следования импульсов"; наличием на выходе сигналов двух типов — релаксационного токового и пилообразного (напряжения); широким диапазоном напряжений питания; большим интервалом сопротивления нагрузки.

Особого внимания заслуживает вопрос о применении БИСПИНов в устройствах контроля. Здесь возможны два направления:

- использование БИСПИНа как датчика, например, светового потока, температуры, тока, напряжения и т. д.

- использование БИСПИНа в качестве преобразования "аналоговый сигнал—частота" для любых стандартных и выпускаемых промышленностью датчиков.

На рис. 4.23 приведены фотошаблоны для изготовления БИСПИН-прибора на основе пластин легированного бором кремния.

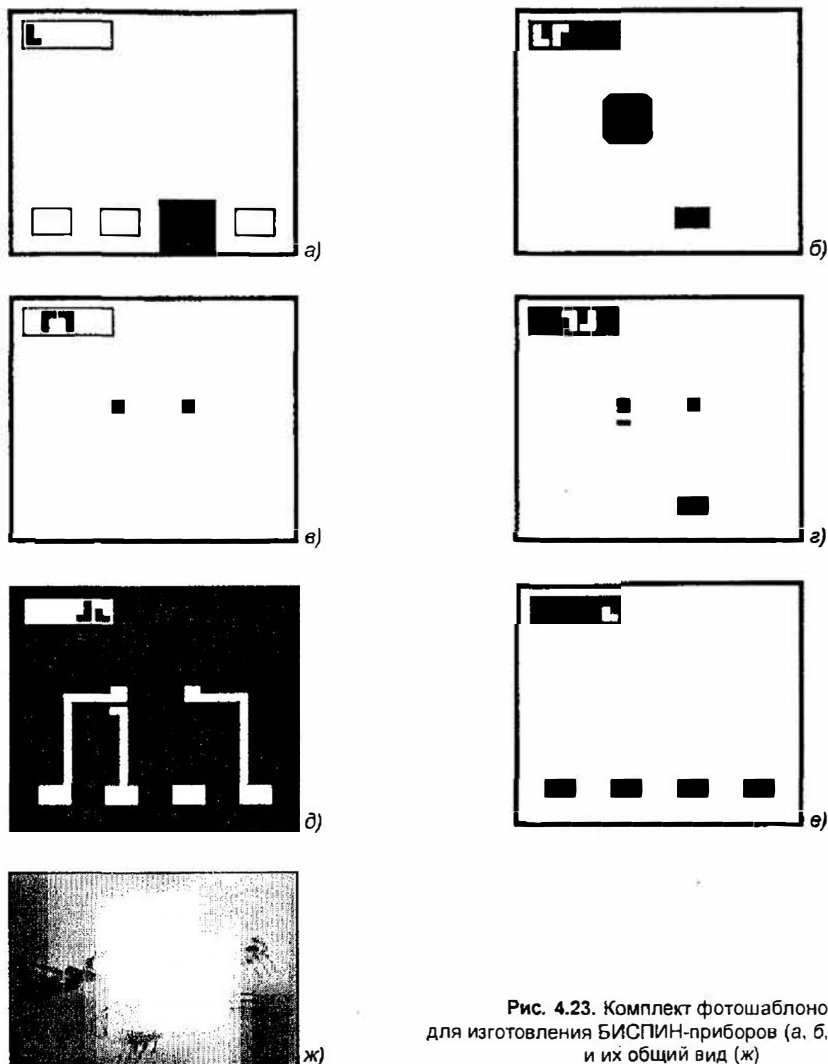


Рис. 4.23. Комплект фотошаблонов для изготовления БИСПИН-приборов (а, б, в, г, д, е) и их общий вид (ж)

После процессов окисления, нанесения фоторезиста и сушки производится первая фотолитография областей разделения (рис. 4.23, а). Затем следуют процессы травления окисла, снятия фоторезиста, химическая обработка и разделительная высокотемпературная диффузия бора. Новый технологический цикл процессов снятия окисла, химической обработки пластин, окисления и второй фотолитографии (рис. 4.23, б). Циклы травления окисла,

снятие фоторезиста, химическая обработка позволяют подготовить пластину к проведению базовой диффузии бора.

База формируется размером 350×350 мкм, глубиной залегания 2,4 мкм и концентрацией примеси бора 10^{18} см^{-3} .

Третья фотолитография проводится с целью формирования эмиттеров в виде линейного и нелинейного контактов (рис. 4.23, в). С этой целью проводятся операции травления окисла, снятия фоторезиста, химической обработки, а затем эмиттерная двухстадийная диффузия фосфора. Размер эмиттерных областей составляет 90×90 мкм, глубина залегания 1,1 мкм и концентрация примесей 10^{20} см^{-3} .

Затем пластина готовится к процессам формирования контактных окон и с этой целью проводится четвертая фотолитография (рис. 4.23, г). В очередном цикле технологических процессов предусмотрено травление окисла, снятие фоторезиста, химическая обработка перед процессом металлизации алюминием.

Пятая фотолитография проводится с целью формирования металлических контактов к линейному, нелинейному и базовому контактам (рис. 4.23, д).

Процессы травления алюминия, снятие фоторезиста, химическая обработка предваряют процесс вжигания алюминиевых токопроводящих дорожек. Контакты имеют размеры 50×100 мкм, их удельное сопротивление составляет 2 Ом/см .

После этого следуют циклы химической обработки и, наконец, последняя шестая фотолитография (рис. 4.23, е). Очередные циклы травления, снятия фоторезиста, химической обработки и термообработка при 300°C . После проверки ВАХ проводится термокомпрессионная обработка. Пластины отмываются, скрайбируются и развариваются в корпуса. Тестирование приборов проводится при напряжении $U_{кб} > 50 \text{ В}$, $U_{из} \cong 20 \text{ В}$.

На рис. 4.23, ж представлены корпусированные БИСПИН-приборы, в центре которых просматриваются кристаллы.

На их основе созданы различные типы датчиков с частотным выходом: оптоэлектронные преобразователи, преобразователь типа "аналоговый сигнал—количество импульсов", измеритель световых, магнитных потоков. Значительный интерес представляют датчики различных физических параметров, реализованные на БИСПИН-приборах. В этом случае резко повышается точность измерений, появляется возможность передачи данных по радиоканалу. На основе БИСПИН-приборов разработаны датчики для дистанционного контроля в системах экологического мониторинга.

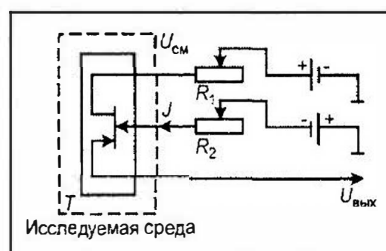
На рис. 4.24 приведены температурно-частотные зависимости датчиков температуры на основе БИСПИН-приборов при различных значениях тока в подложку I и напряжении смещения 6,3 В. Наблюдается линейность в диапазоне температур от -60 до $+80$ градусов. Оценки показывают, что в этом случае возможно получение точности измерения температуры порядка 10^{-3} градуса.

Частотный выход датчиков на основе БИСПИН-приборов позволил создать на их основе комплект интеллектуальных датчиков для экологического мониторинга. Частотный выход БИСПИН-структур позволяет на их основе реализовать передачу телеинформации по радиоканалу. На этом принципе реализованы датчик магнитного поля, измеритель мощности световых потоков, датчик температуры с использованием термистора и др.

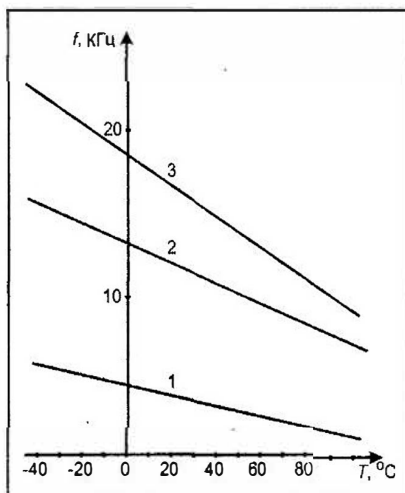
Следует заметить, что вся эта схема может быть реализована в виде одной интегральной схемы. Одновременно можно разместить несколько датчиков для контроля нескольких

физических величин. Реально встроить схемы автоматического переключения каналов, адаптации к изменению эксплуатационных условий, схем промежуточной обработки информации. Такие конструкции получили название *интеллектуальные датчики* (сенсоры).

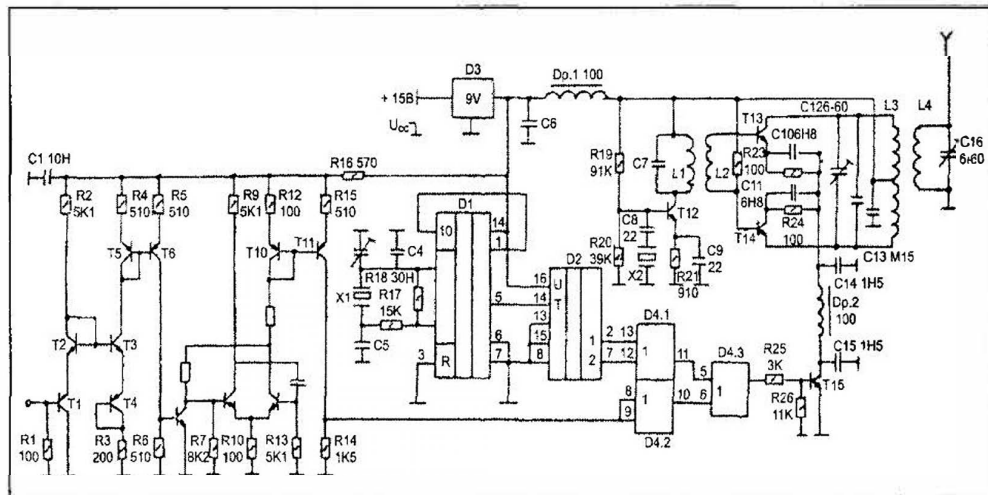
Рис. 4.24. Схема датчика температуры на БИСПИН-приборе (а) и его температурно-частотные зависимости (б) и схема телеканала передачи данных (в)



а)



б)



в)

4.2.5. Приборы на волнах пространственного заряда

В континуальных средах, представляющих собой тонкие слои полупроводников и обладающих отрицательной объемной дифференциальной проводимостью (ОДП), при определенных условиях возбуждаются волны пространственного заряда (ВПЗ).

Такие волны по существу являются динамическими неоднородностями, обладающими уникальными свойствами:

- скорость волн близка к скорости дрейфа электронов ($\sim 10^5$ м/с);
- величина и направление скорости ВПЗ легко управляемы за счет изменения величины и направления статического электрического поля, приложенного в плоскости полупроводникового слоя;
- ВПЗ могут эффективно усиливаться, слабая дисперсия скорости позволяет организовать когерентное многоволновое взаимодействие со средой.

Это явление положено в основу нового класса приборов. Следует особо подчеркнуть, что технология создания континуальной среды, а также приборов на основе ВПЗ вполне совместима с технологией микроэлектроники.

Конструкция одного из процессоров, выполняющего функции усилителя и фильтра, приведена на рис. 4.25. Континуальной средой такого устройства является тонкий слой GaAs n-типа проводимости.

Возбуждение ВПЗ осуществляется многоэлементным преобразователем, выполненным в виде встречно-штыревого преобразователя (ВШП). Динамические неоднородности в виде ВПЗ, генерируемые каждым из электродов ВШП, распространяются к детектирующему ВШП и наводят в нем электрический сигнал. Этот сигнал пропорционален суммарному значению переменных составляющих концентрации электронов. Другими словами, электрический сигнал с детектора равен сумме всех парциальных ВПЗ, возбужденных каждым из электродов.

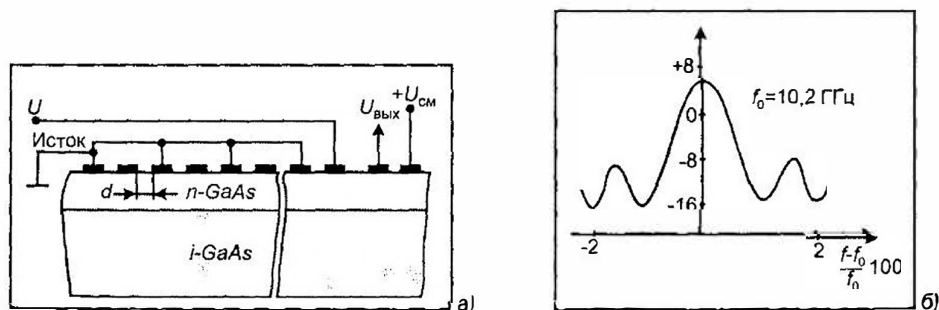


Рис. 4.25. Схема процессора на ВПЗ для усиления и фильтрации сигналов (а) и его частотная характеристика (б)

Максимальный выходной сигнал будет наблюдаться в том случае, если все парциальные ВПЗ придут на выходной электрод синфазно. Геометрия ВШП такова, что расстояние между соседними штырями должно быть равно половине длины ВПЗ λ ; тогда частота определяется соотношением

$$\omega_0 = \pi V_0 / d,$$

где V_0 — скорость волны.

Такой процессор одновременно выполняет функцию фильтра с амплитудно-частотной характеристикой вида $\frac{\sin x}{x}$, причем

$$\frac{\Delta f}{f_0} \approx \frac{1}{N^3},$$

где N — число штырей ВШП (рис. 4.25, б). Подавление побочных максимумов можно осуществить путем аподизации ВШП подобно тому, как это рассматривалось в акусто-электронных устройствах.

Другим примером процессора на ВПЗ является конвольвер сигналов СВЧ-диапазона. В этом процессоре используется нелинейное взаимодействие ВПЗ. Конвольвер представляет собой многослойную сэндвич-структуру (рис. 4.26). Две континуальные среды, отличающиеся степенью легирования и скоростью дрейфа волн пространственного заряда, расположены одна над другой и разделены тонким слоем диэлектрика.

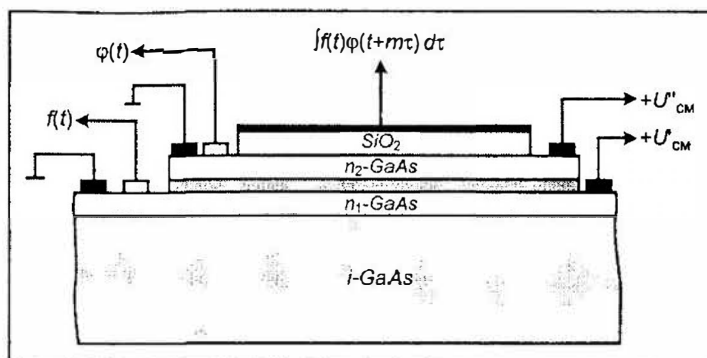


Рис. 4.26. Монолитный конвольвер на ВПЗ

На входные электроды прибора подаются сигналы в СВЧ-диапазоне. Генерируемые электродами динамические неоднородности в виде ВПЗ распространяются в смежных континуальных средах. Суммарная толщина полупроводниковых слоев и разделяющих их пленок диэлектрика меньше характерной длины ВПЗ и поэтому волны, распространяющиеся в континуальных средах, эффективно взаимодействуют. Результатом их взаимодействия является результирующий сигнал, описываемый соотношением

$$U(t) = A \int f(t) \varphi(t + m\tau) d\tau,$$

где A — размерный коэффициент; m — множитель, определяемый относительной разностью скоростей дрейфа в двух континуальных средах. Если две ВПЗ движутся в противоположных направлениях, то можно получить классическое преобразование — операцию свертки сигналов:

$$U(t) = A \int f(t) \varphi(t - \tau) d\tau.$$

Интервалы интегрирования определяются временем прохождения волны в приборе.

Помимо процессоров на ВПЗ можно реализовать устройства памяти аналоговых сигналов. На рис. 4.27 представлено устройство хранения аналоговых сигналов в СВЧ-диапазоне.

Область хранения сигналов представляет собой эквидистантную систему хранящих электродов, связанных с шиной хранения через емкости. В режиме записи напряжение на шину хранения не подается. Хранящие затворы находятся под плавающим потенциалом и мало влияют на условия распространения ВПЗ в пролетной области. В этом случае реализуется режим широкополосного усилителя бегущей волны.

В режиме хранения на шину подается импульс хранения отрицательной полярности такой амплитуды, чтобы обедненные области под "хранящими" электродами полностью сомкнулись. В этом случае в пленке образуются изолированные области — *карманы*, содержащие динамические неоднородности в виде зарядовых пакетов. Наличие или отсутствие зарядовых пакетов свидетельствует о наличии битов хранимой информации; ВПЗ в этом случае как бы замораживаются в среде. Положительный потенциал с выхода устройства может быть снят.

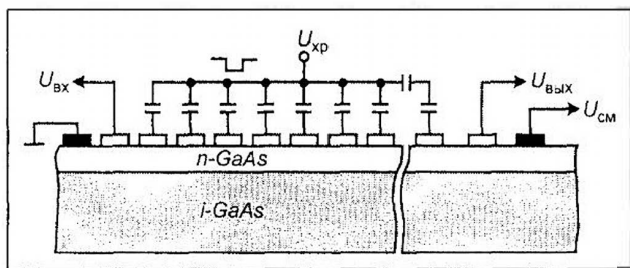


Рис. 4.27. Запоминающее устройство аналоговых сигналов

В режиме воспроизведения сигнала на сток подается положительное тянущее напряжение, снимается отрицательный потенциал с шины хранения. Происходит размораживание зарядовых пакетов, а затем их детектирование на электроде. Помимо одномерного рассмотрения процесса распространения ВПЗ исследованы двумерные явления. Управляя ВПЗ в заданной плоскости, созданы различные конструкции процессоров сигналов. Например, фазированные антенные решетки, линии задержки с дискретным или непрерывным управлением, селекторы СВЧ-импульсов с линейной частотой модуляции и т. п.

4.2.6. Ганновские приборы

В полупроводниковой континуальной среде с *N*-образной объемной вольтамперной характеристикой возникает генерация высокочастотных колебаний электрического тока, представляющая собой *эффект Ганна*.

Генерируется динамическая неоднородность в виде электрического домена, которая называется *доменом Ганна* (рис. 4.28).

Если к торцам кристалла GaAs *n*-типа длиной *l*, обладающего *N*-образной ВАХ, приложить напряжение *U* такое, что $E_{кр} < U/l < E_{нпр}$, то возникают локальные флуктуации плотности заряда (рис. 4.28, а). Эта флуктуация величиной X_1, X_2 расположена на падающем участке ВАХ в области отрицательной дифференциальной проводимости (рис. 4.28, б, в).

Напряженность поля вследствие флуктуации возрастает на величину ΔE , а плотность тока *j* в области $X_1 < X < X_2$ окажется ниже, чем вне области. Электроны, движущиеся против сил поля, начинают скапливаться вблизи X_1 и образовывать отрицательный заряд. На координате X_2 остается нескомпенсированный положительный заряд (рис. 4.28, в). Образуются электростатический домен, обедненный свободными электронами.

Под воздействием электрического поля домен перемещается в континуальной среде от катода к аноду со скоростью дрейфа электронов $V_{др}$ (10^5 — 10^7 м/с). На аноде происходит рекомбинация электронов или их детектирование. Динамическая неоднородность распа-

дается, вызывая импульс тока во внешней цепи. Форма импульса тока приведена на рис. 4.31, г. Одновременно у катода зарождается новая динамическая неоднородность в виде домена и процесс повторяется. Размер домена составляет $\Delta x = 10 \div 20$ мкм. Специфичность эффекта Ганна состоит в том, что преобразование мощности постоянного тока происходит во всем объеме среды, а не в узкой области, например, р—п-переходе. Поэтому генераторы на основе эффекта Ганна имеют значительную мощность. В качестве среды используются InP, CdTe, ZnS, InSb, InAs n-типа, а также в Ge р-типа с характерным размером — 1 мм.

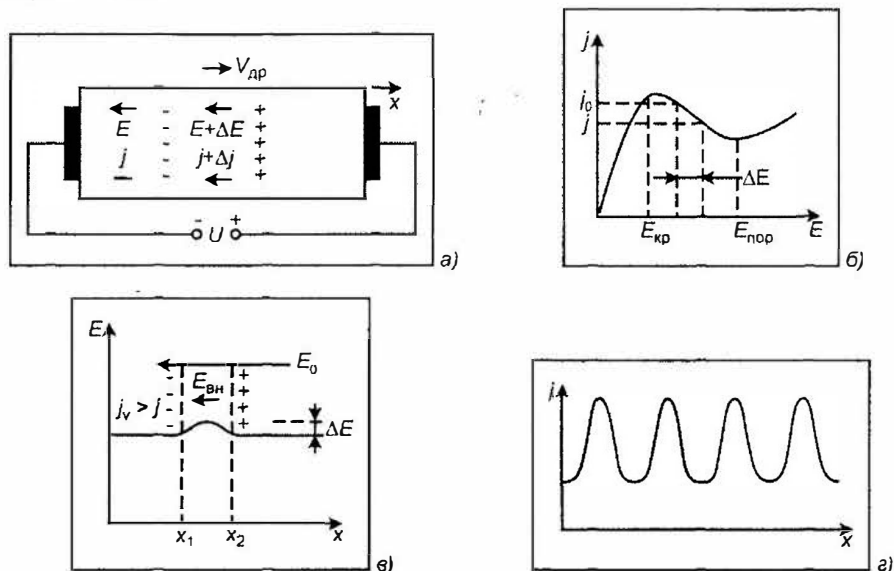


Рис. 4.28. Устройство на доменах Ганна (а), ВАХ (б), распределение поля в полупроводнике (в) и форма импульсов генерации (г)

На основе эффекта Ганна разработаны конструкции процессоров и памяти. Рассмотрим некоторые из них.

Управление импульсами тока может производиться функцией, являющейся произведением профиля легирования среды на площадь поперечного сечения образца. По существу эта функция определяет заряд электростатического домена. В этом случае легко генерировать колебания тока сложной формы, придавая соответствующую форму образцу (рис. 4.29, а).

Форму колебаний можно задавать также с помощью профиля металлического контакта на поверхности образца, описываемого заданной функцией $f(x)$. В этом случае процессор произведет преобразование функции координат $f(x)$ во временную функцию $f(t)$. Металлический электрод изолируется от образца диэлектрической пленкой (рис. 4.29, б).

На основе прибора Ганна, неудачно названного диодом, можно реализовать устройства логики, используя два его состояния: омическое (без домена) и с доменом сильного поля. В диод Ганна вводится устройство управления в виде затворов Шоттки, которые расположены перпендикулярно направлению тока. В таком приборе можно стимулировать или подавлять ганновский домен, а значит формировать нужную логическую функцию.

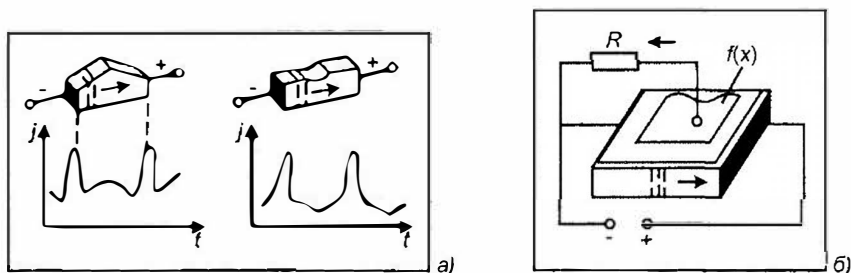


Рис. 4.29. Общий вид образцов и форма генерируемых импульсов тока (а); прибор с профилированным электродом (б)

Диоды Ганна могут быть использованы в устройствах памяти высокого быстродействия ($\sim 10^{-10}$ с). Как правило, как элементы памяти, так и ЗУ реализуются на диодах Ганна, объединенных в устройство на основе различных схемотехнических решений, и особого интереса для практической электроники не представляют.

Контрольные вопросы

1. Что такое функциональная полупроводниковая электроника?
2. Какие динамические неоднородности полупроводниковой природы вы знаете?
3. Как устроена линия задержки на ПЗС-структурах?
4. Что представляет собой трансверсальный фильтр, реализованный на ПЗС-структуре?
5. Что такое линейный ФПЗС и как он работает?
6. Что такое матричный ФПЗС и как он работает?
7. Что такое цифровые процессоры на ПЗС-структурах?
8. Какие принципы лежат в основе конструкции запоминающего устройства на ПЗС-структурах?
9. Как использовать в качестве фильтра сигналов линию задержки, реализованную на ПЗС-структурах?
10. Как реализовать операции суммирования и вычитания на ПЗС-структурах?
11. Как реализовать функции И, ИЛИ на ПЗС-структурах?
12. Что такое БИСПИН-прибор и какими свойствами он обладает?
13. Что такое волны пространственного заряда и какими свойствами они обладают?
14. Опишите конструкцию монолитного конвольвера на ВПЗ.
15. Что такое домен Ганна, и при каких условиях он возникает?
16. Опишите конструкцию процессора на доменах Ганна.

Рекомендуемая литература

1. Булгаков С. С., Кнаб О. Д., Лысенко А. П., Мома Ю. А. БИСПИН — новый прибор микроэлектроники: Обзоры по электронной технике. Сер. 2. — М.: ЦНИИ "Электроника", вып. 6(1563), 71 с., 1990.
2. Бужанева Е. В. Микроструктуры интегральной электроники. — М.: Радио и связь, 1990.

3. Кузнецов Ю. А., Шилин В. А. Микросхемотехника БИС на приборах с зарядовой связью. — М.: Радио и связь, 1988.
4. Носов Ю. Р., Шилин В. А. Основы физики приборов с зарядовой связью. — М.: Наука, 1986.
5. Рычина Т. А., Зеленский А. В. Устройства функциональной электроники и электрорадиоэлементы. Учебник для вузов. — М.: Радио и связь, 1989.
6. Степин В. Я. Применение микросхем с зарядовой связью. — М.: Радио и связь, 1989.
7. Щука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

5. Функциональная магнитоэлектроника

5.1. Физические основы

Функциональная магнитоэлектроника представляет собой направление в функциональной электронике, в котором изучаются магнитоэлектронные эффекты и явления в магнитоупорядоченных континуальных средах, а также возможность создания приборов и устройств обработки и хранения информации с использованием динамических неоднородностей магнитоэлектронной природы.

Магнитное упорядочение заключается в существовании определенной закономерности расположения элементарных магнитных моментов атомов, ионов, электронов. Простейшие типы магнитного упорядочивания наблюдаются в ферри- и ферромагнетиках. *Ферромагнетик* представляет собой вещество, в котором ниже определенной температуры (точки Кюри) большинство атомных магнитных моментов параллельно друг другу и вещество обладает самопроизвольной (спонтанной) намагниченностью. Под *ферримагнетиками* понимают магнетики с несколькими магнитными подрешетками и с отличным от нуля суммарным магнитным моментом. Магнитное упорядочивание в любых классах магнетиков исчезает при температуре выше точки Кюри для ферромагнетиков и точки Нееля для ферри- и антиферромагнетиков.

5.1.1. Динамические неоднородности в магнитоэлектронике

В континуальных магнитоупорядоченных средах существует несколько видов физических носителей информационного сигнала — динамических неоднородностей.

Цилиндрические магнитные домены (ЦМД) представляют собой изолированные однородно намагниченные подвижные области в ферро- или ферримагнетиках. ЦМД, которые имеют форму круговых цилиндров и направление намагниченности, противоположное направлению намагниченности магнетика.

ЦМД возникают при определенных условиях в тонких монокристаллических пластинках или пленках ферритов, обладающих сильной одноосной перпендикулярной анизотропией. Единственная ось легкого намагничивания направлена перпендикулярно поверхности континуальной среды, в которой формируются динамические неоднородности в виде ЦМД (рис. 5.1). Устойчивое равновесие ЦМД поддерживается под действием трех сил: сжимающей силы со стороны внешнего магнитного поля, магнитостатической силы растягивания домена за счет "магнитных зарядов", сжимающей силы поверхностного натяжения домена. Изолированный домен существует в определенном интервале значений внешнего поля

$$H_{\min} < H_{\text{оп}} < H_{\max},$$

где H_{\min} — поле, при котором ЦМД растягивается в полосовой домен, H_{\max} — поле, при котором ЦМД схлопывается — коллапсирует. Значение этих полей зависит от отношения толщины шелки к ее характеристической длине.

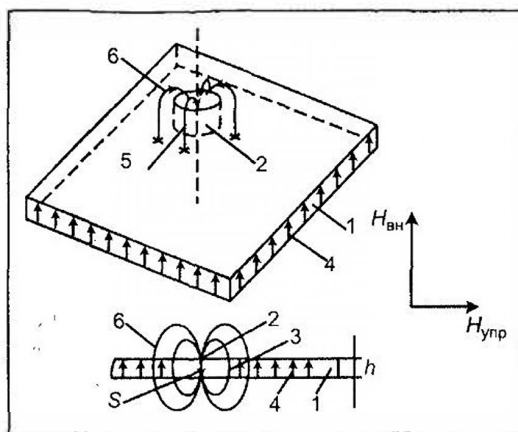


Рис. 5.1. Изолированный ЦМД: 1 — в пленке феррита-граната; 2 — ЦМД; 3 — стенка Блоха; 4 — вектор намагниченности соседнего домена; 5 — вектор намагниченности домена; 6 — поле рассеяния домена

ЦМД можно перемещать в пленке воздействием на него полем $H_{\text{вн}}$. Скорость перемещения определяется соотношением

$$V = \frac{\mu}{2} \left(\Delta H_{\text{вн}} - \frac{8}{\pi} H_c \right) \sim 1 \div 10 \text{ м/с}$$

где $\Delta H_{\text{вн}}$ — изменение напряженности внешнего поля на расстоянии, равном диаметру домена, μ_g — подвижность домена, H_c — коэрцитивная сила материала.

Переход вектора намагниченности из одного положения в другое происходит на границе раздела между двумя доменами. Эта граница или переходной слой называется *доменной границей*, и различают следующие их типы.

Неелевская граница существует в очень тонких магнитных пленках (менее 30 нм). Вектор намагниченности ориентирован по оси легкой намагниченности (ОНЛ), поворот атомных магнитных моментов в границе происходит только в плоскости пленки.

Блоховская граница существует в пленках толщиной более 100 нм и во всех массивных ферромагнетиках. Разворот магнитных моментов возможен как по часовой (правовинтовая граница), так и против часовой стрелки (левовинтовая граница). Особенностью блоховской доменной границы является отсутствие на ее краях магнитных зарядов.

Промежуточным типом является граница с поперечными связями, существующая в пленках толщиной от 30 до 100 нм. В настоящее время значительный интерес представляют вертикальные блоховские линии (ВБЛ) в полосовых доменах. ВБЛ являются одним из типов динамических неоднородностей и на их основе созданы приборы для обработки и хранения информации.

В двоичном исчислении логические "1" и "0" образуются наличием или отсутствием ВБЛ. Расстояние между соседними ВБЛ достаточно мало, поэтому в стенке одного такого полоскового домена можно хранить множество битов информации. В полосковом домене размером 0,5 мкм можно хранить до 100 бит информации. Динамические неоднородности в виде ВБЛ, переносящие или хранящие биты информации, считываются на одном конце полоскового домена и генерируются (записываются) на другом (рис. 5.2).

В оптических дисках со стиранием и перезаписью информации используются домены в дисках из тербий-железо-кобальтового сплава с точкой Кюри 240 °С. Домены первоначально ориентированы в одном направлении (рис. 5.3, а). При воздействии лазерного из-

лучения происходит локальный разогрев до точки Кюри (рис. 5.3, б). Если в этом месте действовать магнитным полем, то можно обратить поляризацию домена (рис. 5.3, в). Заметим, однако, что такие динамические неоднородности превращаются на определенное время в статические.

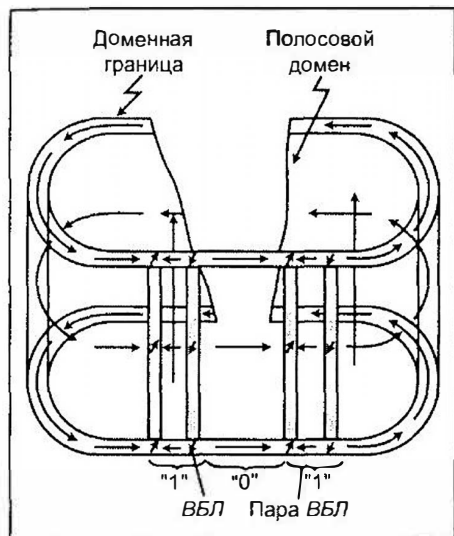


Рис. 5.2. Фрагмент полосового домена с тремя парами ВБЛ, представляющие собой логические единицы

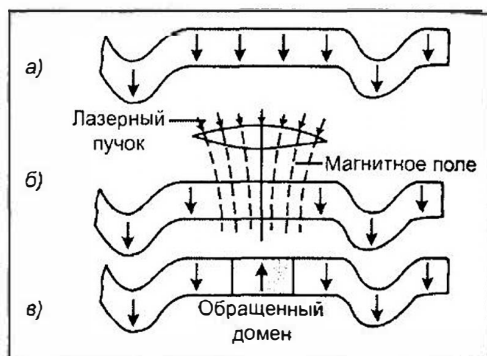


Рис. 5.3. Динамика обращения домена в оптическом диске

В функциональной магнитоэлектронике в качестве динамических неоднородностей часто используются резонансы и волны.

В квантовых системах под *резонансом* будем понимать резкое возрастание квантовых переходов при равенстве частот внешнего излучения $h\nu$ и квантового перехода $\epsilon_i - \epsilon_j$ или

$$h\nu = \epsilon_i - \epsilon_j.$$

Явление резонанса проявляется в увеличении интенсивности обмена энергией в процессах поглощения и излучения.

Механический момент каждого магнитного иона в ферромагнетике (спин) совершает прецессию под воздействием поля $\vec{H}_{\text{упр}}$, перпендикулярного внешнему $\vec{H}_{\text{вн}}$. Величина прецессии определяется *ларморовской частотой*.

В реальной ситуации магнитные моменты ионов в ферромагнетике взаимодействуют между собой благодаря дипольным полям, обмену энергией, взаимодействию с полем кристаллической структуры.

Возникают объемные магнитостатические волны (ОМСВ) и поверхностные магнитостатические волны (ПМСВ).

Магнитостатические волны представляют собой электромагнитные волны, сопровождающие колебания спинов в магнитоупорядоченных средах.

Фазовая скорость магнитостатических волн много меньше скорости света. Магнитостатические волны возбуждаются неоднородным магнитным полем, создаваемым током, протекающим по проводнику на поверхности пленки. На рис. 5.4 представлены два состояния ионов во внешнем магнитном поле $\vec{H}_{вн}$: в отсутствие управляющего поля $\vec{H}_{упр}$ спины выстраиваются вдоль направления постоянного поля (рис. 5.4, а); при наложении управляющего поля $\vec{H}_{упр}$, например, за счет протекания тока по проводнику, создается спиновая волна, распространяющаяся по пленке и являющаяся динамической неоднородностью (рис. 5.4, б). Характер распространения возмущения напоминает движение одиночной волны по струне, причем увеличение натяжения струны аналогично увеличению магнитного поля $\vec{H}_{упр}$. В обоих случаях наблюдается увеличение частоты. Эти волны называют также *спиновыми*.

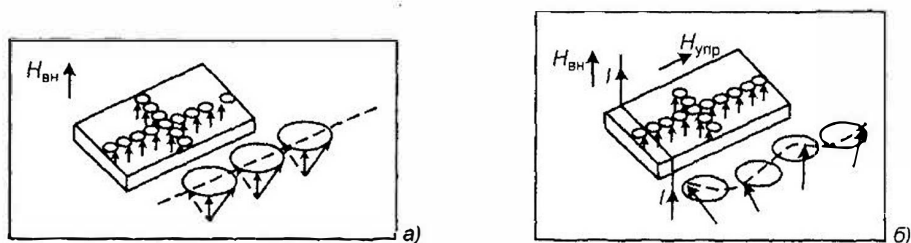


Рис. 5.4. Возникновение динамической неоднородности в виде спиновой волны: а — в отсутствие управляющего поля; б — при наличии управляющего поля

Наряду со спиновыми волнами в ферромагнетиках можно возбуждать упругие волны в ультразвуковом диапазоне, электромагнитные волны. Эти волны можно рассматривать как динамические неоднородности, между которыми могут возникать процессы взаимодействия в случае, если они распространяются одновременно. Взаимодействие упругих и спиновых волн, например, представляет собой магнитострикционную связь, которая позволяет изменить свойства волн вблизи точек синхронизма и образовать гибридную *магнитоупругую волну* (МУВ). Фазовые скорости таких волн выше скорости звука.

Если рассмотреть слоистую структуру ферромагнетик-полупроводник, то спиновые волны в ферромагнитной пленке могут взаимодействовать с электронами проводимости полупроводника. Дело в том, что электрическое поле спиновой волны проникает через границу раздела в полупроводник и создает там переменный ток. Джоулево рассеяние мощности приводит к затуханию волны. В случае приложения поля к полупроводниковой пленке в ней возникает электрический ток, носители приобретают упорядоченное движение. На эти носители действует сила Лоренца, возникающая вследствие проникновения в полупроводник магнитного поля спиновой волны. В случае совпадения скорости упорядоченного движения носителей \vec{V}_H в полупроводнике и фазовой скорости спиновой волны \vec{V}_p , сила Лоренца колеблется в противофазе с силой электрического поля волны. При этом, если $V_H < V_p$, то происходит ослабление спиновой волны; если $V_H \cong V_p$, то не наблюдается поглощение спиновой волны и, наконец, если $V_H > V_p$, то происходит усиление спиновой волны или магнитоакустический резонанс. Это яркий пример взаимодействия динамических неоднородностей различной физической природы в различных континуальных средах. Магнитоакустический резонанс возникает на частотах $1-10^2$ ГГц.

В магнитоупорядоченных средах можно возбудить динамические неоднородности в виде *магнонов* — квазичастиц, представляющих собой квант колебаний спиновых волн.

В твердых телах, находящихся в сильном магнитном поле $\vec{H}_{\text{эл}}$, возникают волны плазмы. Эти волны имеют круговую поляризацию, их называют *спиральными волнами* или *геликонами*. Геликоны в процессе распространения взаимодействуют с фотонами и при совпадении частоты звука с частотой колебаний плазмы возникают смешанные колебания. В этом случае наблюдаются явления возбуждения звуковых колебаний колебаниями плазмы и наоборот.

В достаточно однородных сверхпроводниках II рода могут быть возбуждены магнитные вихри или *вихри Абрикосова*. Магнитный вихрь представляет собой нить нормальной фазы, окруженную экранирующим током. Вдоль оси такой нити проходит квант магнитного потока или флюксон, величина которого определяется как

$$\Phi_0 = \hbar c / 2q = 2,07 \times 10^{-15} \text{ Вб}.$$

Прикладывая к вихрю внешнюю силу (сила Лоренца), можно перемещать вихрь в плоскости пленки. Особый интерес представляет использование флюксонов в пленках высокотемпературных сверхпроводников (ВТСП).

5.1.2. Континуальные среды

При выборе континуальных сред руководствуются условием максимальной эффективности возбуждения заданного типа динамической неоднородности. Первые ЦМД приборы изготавливались на редкоземельных ортоферритах с общей формулой LnFeO_3 . Однако на их основе не могла быть обеспечена высокая плотность записи информации, экономичность приборов.

Наиболее оптимальной средой для создания ЦМД приборов оказались монокристаллические пленки феррит-гранатов (МПФГ), эпитаксиальные пленки феррит-шпинелей. В настоящее время магнитоодноосные пленки ферритов-гранатов, выращенных на намагниченных подложках галлий-гадолиниевого граната (ГГГ), не испытывают заметной конкуренции со стороны других материалов.

Различают гранаты следующих систем: $(\text{LaLuSmCd})_3(\text{FeGa})_5\text{O}_{12}$; $(\text{YLuSm})_3(\text{FeGaSe})_5\text{O}_{12}$; $(\text{YSm})_3(\text{FeAl})_5\text{O}_{12}$. При формировании континуальных сред стараются получить ЦМД субмикронного диаметра, для чего используют, например, гранаты без ионов Sm^{3+} . Для доменопередвигающих структур используются пермаллоевые пленки (79%Ni + 21%Fe) с нулевым коэффициентом магнитострикции.

Для устройств ФЭ, использующих в качестве динамических неоднородностей спиновое эхо, могут быть использованы соединения типа $\text{NiFe}^{57}\text{O}_4$, пленки или порошок Co^{59} , пленки $\text{Co}^{59}\text{Fe}^{57}$.

5.1.3. Генерация, детектирование и управление динамическими неоднородностями

Информация в устройствах на динамических неоднородностях, например, типа доменов, кодируется двумя основными способами, наличием или отсутствием доменов в заданный момент времени в заданном месте, либо состояниями границ доменов.

Процесс генерации доменов осуществляется в основном локализацией электромагнитного поля, например, с помощью аппликации в виде петли, по которой протекает ток /

(рис. 5.5). При подаче импульса тока будет создано размагничивающее поле $H_{\text{разм}}$, превосходящее значение рабочего поля $H_{\text{рел}}$ и имеющее противоположный знак. Под аппликацией пленка локально перемагнитится и образуется домен.

Другим примером процесса генерации доменов может служить деление зародышевого домена. Если домен зародится, то будет записана логическая единица, в противном случае — логический нуль.

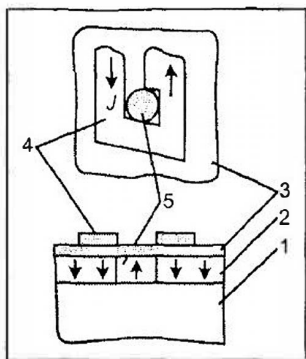


Рис. 5.5. Генерация доменов с помощью петли тока: 1 — подложка ГГГ; 2 — зпитаксиальная ферритовая пленка; 3 — изолирующая пленка; 4 — аппликация; 5 — ЦМД

Если в качестве динамической неоднородности применяется ВБЛ, то используется растянутый ЦМД или *страйп-домен* (рис. 5.6). Операцию записи производят в следующем порядке: вначале отрицательную ВБЛ, находящуюся в вершине страйпа (рис. 5.6, а), переводят в другое положение (рис. 5.6, б, в). Прилагая к вершине страйпа импульс поля смещения, можно получить пару ВБЛ разных знаков (рис. 5.6, г). Положительную ВБЛ переводят в вершину страйпа (рис. 5.6, д), определяют ЦМД (рис. 5.6, е), и уничтожают его путем аннигиляции. Итогом генерации является отрицательная пара ВБЛ в вершине страйпа. Наличие пары отрицательных ВБЛ соответствует логической единице, а ее отсутствие — логическому нулю.

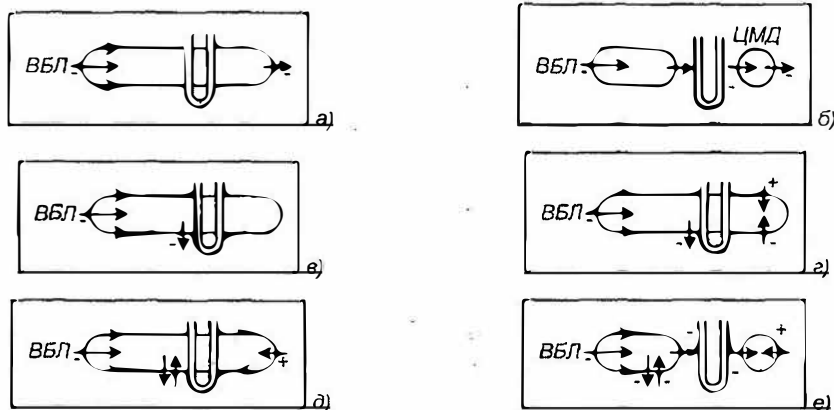


Рис. 5.6. Запись ВБЛ в накопительный регистр (а, в, г, д, е) и ее считывание (б)

Для других устройств функциональной магнитоэлектроники существуют свои методы генерации, в основе которых также лежит принцип возбуждения полем. В устройствах на

спиновых волнах используется проводник с током, токовая петля, а также другие генераторы управляющего магнитного поля. Управление динамическими неоднородностями магнитоэлектронной природы осуществляется различными физическими эффектами и явлениями, связанными с наложением физических полей. По своей природе ЦМД имеют магнитное поле, которое, взаимодействуя с внешним полем, создает силы, перемещающие домены в направлении минимальной интенсивности внешнего поля. Другими словами, домен перемещается в таком направлении, в котором его энергия будет минимальной.

Токовые аппликации, генерирующие магнитное поле, имеют вид петли из металлической пленки (рис. 5.7, а). Направление тока в аппликации таково, что создается поле, противоположное внешнему, и образуется "яма", которая движется в заданном направлении вследствие последовательной подачи импульсов на фазы Φ_1 , Φ_2 , Φ_3 .

Предпочтение отдается трехфазной схеме, чтобы избежать взаимодействия следующих друг за другом доменов, а также возвратного их движения. Тактовые аппликации просты, но ненадежны из-за большого количества проводников.

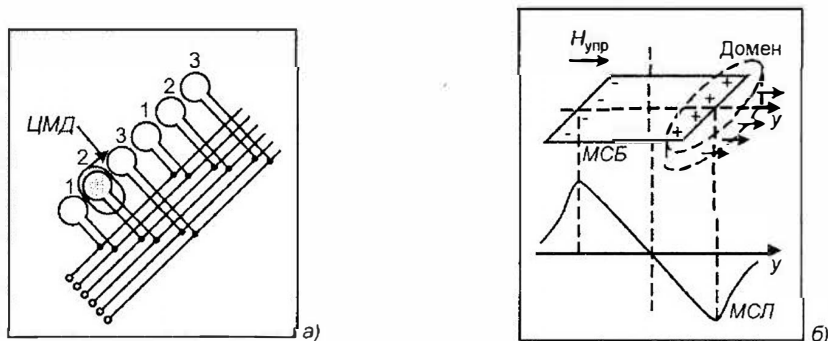


Рис. 5.7. Управление движением ЦМД: а — с помощью токовых аппликаций; б — с помощью пермалловых аппликаций

Более удобны магнитные аппликации, управляющее внешнее поле в которых создает магнитостатические ямы и барьеры (МСБ) (рис. 5.7, б). В таких конструкциях домен продвигается от аппликации к аппликации, форма которых оптимизируется в конкретном случае и представляет собой доменопередвигающую структуру (ДПС). ДПС является эффективным устройством управления динамической неоднородностью в виде домена.

На рис. 5.8 проиллюстрирована идея перемещения ЦМД по классическим π -образным ДПС в различные моменты времени, связанные с направлением вектора управляющего магнитного поля $H_{упр}$. Генерация доменов происходит в генераторе G , а детектирование — в детекторе D .

ДПС могут полностью управлять перемещением и коммутировать ЦМД по любому ранее заданному направлению или каналу. Существуют множество конструкций ДПС, однако, наиболее перспективны ионно-имплантированные ДПС.

Для других видов динамических неоднородностей существуют свои конструкции устройств управления, основанные на магнитоэлектрических эффектах и явлениях.

Считывание информации сводится к выявлению динамических неоднородностей, несущих логические "0" и "1", и осуществляется в устройстве, называемом детектором. При детектировании динамических неоднородностей используются, как правило, физические процессы и явления, обратные процессам генерации. Это эффекты индукции, магниторе-

зистивный, магнитооптический и т. п. Общие принципы построения детекторов включают ряд положений, среди которых: обеспечение отношения сигнал/шум, достаточного для получения заданной вероятности сбоя; технологическая совместимость детектора, устройства управления и генератора; согласованность детектора с входом усилителя считывания; устойчивость к внешним воздействиям.

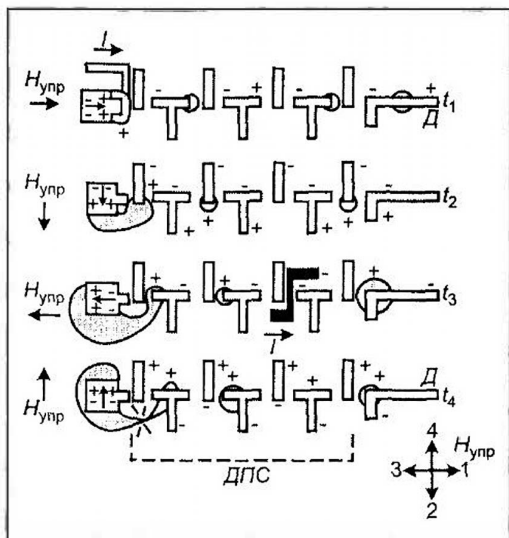


Рис. 5.8. Движение ЦМД по доменопередвигающим структурам. Г — генератор; Д — детектор

Детекторы для считывания ЦМД могут быть различной конструкции в зависимости от используемого физического эффекта. Наиболее распространенной является конструкция с использованием шевронных ферромагнитных или токовых проводниковых расширителей ЦМД.

Увеличение площади считывающего ЦМД с помощью расширителя позволяет получить большую величину выходного сигнала.

Процесс считывания информации в конструкциях на ВБЛ можно понять из рис. 5.7, б. На пару проводников подается импульс тока, и концы страйпов сближаются. Происходит репликация (отделение) ЦМД, которая каналируется в регистр вывода информации и затем детектируется. На конце страйпа сохраняется отрицательная ВБЛ.

Детекторы спиновых волн конструктивно практически не отличаются от генераторов и пришедшая к ним МСВ возбуждает в проводнике импульс тока.

Детектирование флюксонов осуществляется путем их аннигиляции в детекторе.

5.2. Приборы и устройства функциональной магнитоэлектроники

5.2.1. Процессоры сигналов на ЦМД

Уникальные свойства динамических неоднородностей магнитоэлектронной природы позволяют создать весьма эффективные устройства обработки информации. ЦМД как носитель информационного сигнала обладает следующими уникальными свойствами:

- домен в процессах сохраняет форму круглого цилиндра;
- свободно перемещается в двух измерениях;
- управляемо меняет форму, растягиваясь в полосовой домен;
- позволяет перейти к другому физическому носителю информационного сигнала, в том числе многозначному;
- управляемо зарождается или коллапсирует;
- реплицируется, делясь на две части и восстанавливая первоначальную форму и размер;
- визуально и электрически регистрируется;
- дипольно взаимодействует с другими соответствующими ЦМД.

Учитывая эти свойства ЦМД, можно выделить следующие отличительные от интегральных схем свойства приборов функциональной электроники:

- управление процессами обработки информации и ее передача физическими носителями по информационным каналам имеет тотальный и синхронный характер;
- компланарность носителей информационного сигнала имеет две степени свободы его передачи и поэтому время обработки информации уменьшается;
- динамические неоднородности в виде ЦМД при взаимодействии друг с другом неограниченно долго сохраняют свои конечные размеры и количество в системе до, во время и после взаимодействия.

Эти свойства позволяют создать процессоры для обработки дискретной информации.

Анализ показывает, что можно создать логические элементы, реализующие следующие функции: повторение, ИЛИ, И, НЕ, исключающее ИЛИ. Эти элементы и их сочетания позволяют обрабатывать функционально полные системы логических элементов.

На рис. 5.9 показаны реализации логических элементов с ортогональным и антипараллельным расположением входов.

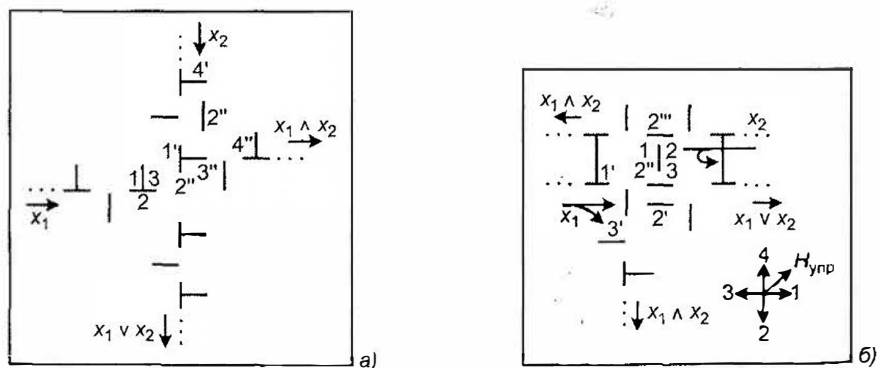


Рис. 5.9. Логические элементы на основе ЦМД:

а — с ортогональным; б — с параллельным расположением входов X_1 и X_2

5.2.2. Процессоры сигналов на МСВ

Перспективными процессами обработки сигналов в гигагерцевом диапазоне являются линии задержки на магнитоэлектрических волнах. Эти устройства напоминают процессоры на ПАВ.

Конструктивно линия задержки на МСВ реализуется на многослойной структуре, состоящей из прочной подложки (Al_2O_3), тонкопленочной континуальной среды из железо-иттриевого граната (ЖИГ) и галлий-гадолиниевого граната (ГГГ) (рис. 5.10).

Такой "сэндвич" является эффективной континуальной средой. Следует отметить, что тонкопленочная технология производства сред для устройств на МСВ совместима с технологией интегральных схем. В зависимости от свойств среды магнитостатические волны распространяются с дисперсией или бездисперсионно, время задержки на всех частотах одинаково. На рис. 5.10, а представлена бездисперсионная линия задержки. На входной детектирующий микрополосковый преобразователь поступает сигнал, который возбуждает МСВ, несущие с задержкой информацию на выходной генерирующий микрополосковый преобразователь. МСВ с частотой линейно не связаны, поэтому в линиях задержки наблюдается зависимость времени задержки от частоты (рис. 5.10, б). Для компенсации такой зависимости прикладывают магнитное поле, перпендикулярно направленное к пленке железо-иттриевого граната (ЖИГ). Используя также две заземленные обкладки, можно изменить параметры распространения МСВ и обеспечить постоянное время задержки в заданном частотном интервале.

В зависимости от направления приложенного поля можно возбудить различные типы волн. При направлении поля по оси z , приложенного перпендикулярно слою ЖИГ, возбуждается прямая объемная волна. Задержка сигнала возрастает с частотой.

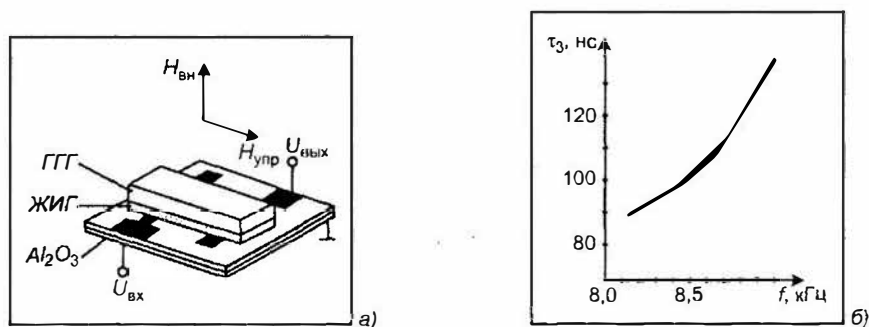


Рис. 5.10. Линия задержки на МСВ (а) и ее характеристика (б)

При приложении поля вдоль оси x генерируется обратная объемная волна с убывающей зависимостью задержки от частоты. И наконец, при направлении магнитного поля вдоль оси y генерируются поверхностные волны, аналогичные ПАВ.

Линия задержки, использующая прямую волну, имеет возрастающую линейно-частотную модуляцию (ЛЧМ), а с обратной волной — убывающую. Если спектр сигнала, поступающего в линию задержки, соответствует полосе пропускания этой линии задержки, но имеет противоположный закон изменения задержки с частотой, например, убывающей, то выходной сигнал будет иметь форму узкого пика. Коэффициент сжатия может составлять несколько порядков.

Линия задержки может быть использована в качестве трансверсального фильтра. Известно, что требуемую характеристику фильтра получают путем суммирования и взвешивания выходных сигналов с ряда слабо связанных отводов. С их помощью снимают сигналы, распространяющиеся в линии задержки. Для приборов на магнитостатических по-

верхностных волнах существует возможность легкого ветвления волны в любых точках тракта распространения.

Заметим, что в зависимости от типа используемых входных и выходных преобразователей, соответственно генераторов и детекторов МСВ, можно получить различные амплитудно-частотные характеристики (АЧХ) преобразования. Преобразователь меандрового типа генерирует АЧХ, описываемую функцией $\frac{\sin x}{x}$. Преобразователь с параллельными

полосками работает в длинноволновой части спектра и генерирует вдвое больше гармоник, чем меандровая структура. Эти преобразователи легли в основу конструкции трансверсальных фильтров на МСВ, полосковых фильтров.

Устройство преобразования спектра СВЧ-сигналов — *серродина* представлен на рис. 5.11. Основой конструкции является линия задержки, в которую встроена катушка подмагничивания для модуляции магнитного поля. В качестве континуальной среды использовалась пленка ЖИГ толщиной 10 мкм на подложке из ГГГ. Серродина работает в диапазоне частот 2—5 ГГц. Для динамической перестройки серродина необходимо изменить магнитное поле на несколько эрстед. Это можно реализовать с помощью малоиндукционных катушек, входящих в устройство управления серродинам. Динамический диапазон по входному сигналу составляет 30—50 дБ, величина управляемого сдвига частот от сотых долей до десятка килогерц при искажении спектра частот ~ 5%. Расширение функциональных возможностей приборов на магнитостатических волнах будет обеспечено благодаря использованию эффектов взаимодействия этого типа динамических неоднородностей с акустическими, оптическими и другими типами волн.

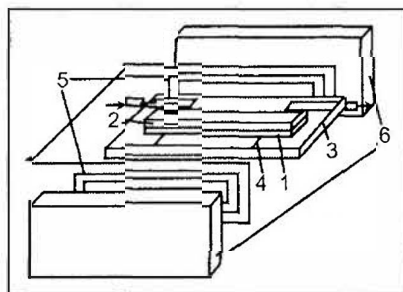


Рис. 5.11. Конструкция серродина на МСВ:
1 — пленка ЖИГ; 2 — подложка из ГГГ;
3 — поликоровая пластина;
4 — микрополосковые преобразователи;
5 — модулирующие катушки;
6 — самарий-кобальтовые постоянные магниты

5.2.3. Запоминающие устройства на ЦМД

Наиболее широкое применение ЦМД нашли в устройствах памяти, позволяющих получить высокую плотность записи информации, энергонезависимость, малую потребляемую мощность, высокое быстродействие, низкую стоимость. Отсутствие движущихся носителей позволяет ЗУ на ЦМД использовать в бортовых системах. Информационные структуры ЗУ на ЦМД могут быть организованы различными способами. Для ЗУ большой информационной емкости характерна организация последовательно-параллельного типа. Рассмотрим структуру с отдельными регистрами ввода-вывода информации и n -накопительными регистрами (рис. 5.12).

Накопительные регистры HP формируются на основе ДПС и информация в них поступает из генератора G через регистр ввода $PВ$ и однонаправленного переклю­чат­еля $П$. Накопи-

тельные регистры связаны с регистром вывода $P_{\text{выв}}$ с помощью репликаторов-переключателей P/P , которые при считывании работают в режимах реплицирования. Копии выбранного блока информации из HP попадают в $P_{\text{выв}}$, а далее в детектор D . В свободные поля через однонаправленные переключатели записываются новые информационные блоки (рис. 5.12, а). На рис. 5.12, б приведены схемы операций обработки информации. Позиция 1 иллюстрирует операцию стирания в P/II , 2 — запись в однонаправленном переключателе и 3 — считывание в P/II . В этом случае информация считывается без ее разрушения.

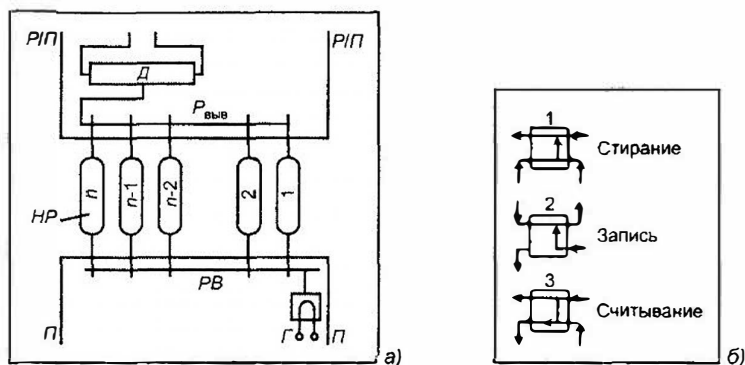


Рис. 5.12. Организация последовательно-параллельной структуры хранения информации: а — с раздельным входом-выходом; б — схема обработки информации

Конструкция ЗУ на ЦМД приведена на рис. 5.13. Микросборка монтируется в корпусе 5 типа DIP. Отдельный чип изготавливается по планарной технологии групповым методом. В последнее время в качестве подложки чипа используется сапфировая подложка. На нее наносится феррит гранатовая пленка, в которой могут образовываться домены. С помощью фотолитографии формируются токовые шины, пермаллоевые (Ni/Fe) доменопередвигающие структуры. ЧИП защищен пленкой нитрида кремния и закрепляется на диэлектрической немагнитной пластине. Два постоянных самарий-кобальтовых магнита 3 создают внешнее поле $\vec{H}_{\text{см}}$, формирующее в феррит-гранатовой пленке ЦМД оптимальных размеров. Между постоянными магнитами помещены две ортогональные катушки 2 и 4, управляемые смещенными на 90° по фазе токами. Это позволяет создать управляющее поле $\vec{H}_{\text{упр}}$, вращающееся по часовой стрелке в плоскости ЧИПа: собранная конструкция помещается в корпус со стандартными штырьками, который экранируется от внешнего магнитного поля.

ЗУ на ЦМД формируется в ЦМД-накопитель, состоящий из нескольких ЦМД-микросборок. Последние выполняют ряд сложных функций, среди которых роль носителей записанного информационного массива, обслуживающих электронных схем и узлов для обеспечения замен, считывания и регенерации информации.

Основным элементом ЦМД-накопителя являются ЦМД-микросборки, содержащие в корпусе одну (однокристалльные) или несколько (многокристалльные) ЦМД-микросхемы, катушки управления полем, постоянные магниты, плату управляющих схем и электромагнитный и электростатический экраны.

Повышение плотности записи информации в устройствах на ЦМД-доменах может быть достигнуто за счет уменьшения диаметра домена. На доменах диаметром ~ 1 мкм созданы

ЗУ с плотностью записи $\sim 10^8$ бит/см². Разрабатываются устройства с субмикронными размерами доменов.

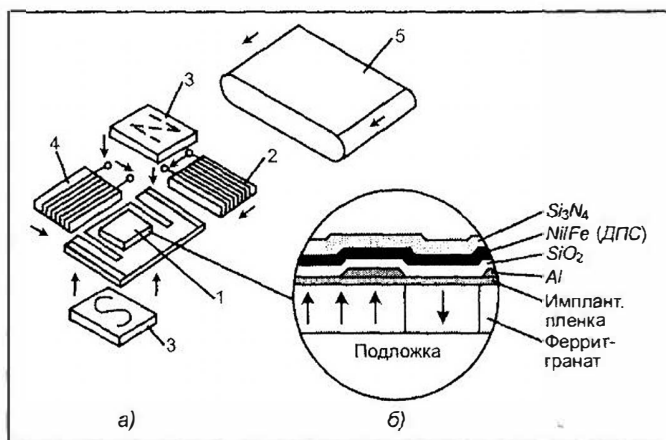


Рис. 5.13. Конструкция микросборки ЗУ на ЦМД (а) и структура слоев отдельного ЧИПа (б)

Уменьшения размеров носителя информации можно получить, перейдя на ЗУ на вертикальных блоховских линиях. Информационный массив в этом случае формируется из страйпов.

Канал ввода информации состоит из генератора ВБЛ, доменопередвигающей структуры с числом позиций равной числу страйпов и токовых шин. Эта система осуществляет преобразования типа ЦМД \rightarrow ВБЛ.

Канал вывода информации имеет в своем составе систему репликаторов (по числу страйпов), осуществляющих преобразование типа ВБЛ \rightarrow ЦМД. Сформированная кодовая последовательность ЦМД направляется в детектор, где происходит считывание информации. Информационная емкость ЗУ на ВБЛ достигает $1,5 \times 10^9$ бит.

Ведутся исследования по созданию "интеллектуальных" ЦМД-ВБЛ систем, в которых на одном кристалле размещены логический процессор и информационный массив. В такой системе можно совместить на одной плате процессы хранения и обработки информационных массивов в реальном масштабе времени со скоростью выше 10^{10} байт/с.

5.2.4. Запоминающие устройства на магнитных вихрях

Принципы организации ЗУ на магнитных вихрях аналогичны ЗУ на ЦМД. Генератор магнитных вихрей (флюксонов) формирует вихри, находящиеся в состоянии безразличного равновесия. Устройство управления представляет собой схему продвижения вихрей по каналу, осуществляемой силой Лоренца, создаваемую транспортным током (рис. 5.14, а). Однонаправленность движения вихрей обеспечивается асимметрией канала продвижения. Считывание магнитных вихрей может осуществляться с использованием квантовых интерферометров, а также одиночных Джозефсоновских контактов.

Возможен вариант формирования сдвигового регистра на магнитных вихрях, либо вариант мажор-минорной организации. В этом случае реализуется накопитель на магнитных

вихрях с параллельно-последовательной организацией записи и считывания информации (рис. 5.14, б). Плотность записи информации в этом случае ожидается $\sim 2 \times 10^8$ бит/см².

Анализ показывает, что скорость обработки информации в этом случае достигает величин $\sim 10^9$ бит/с. Буферная или внешняя память на флюксонных запоминающих устройствах технологически и оперативно совместима с микроэлектронными системами и криогенными процессорами.

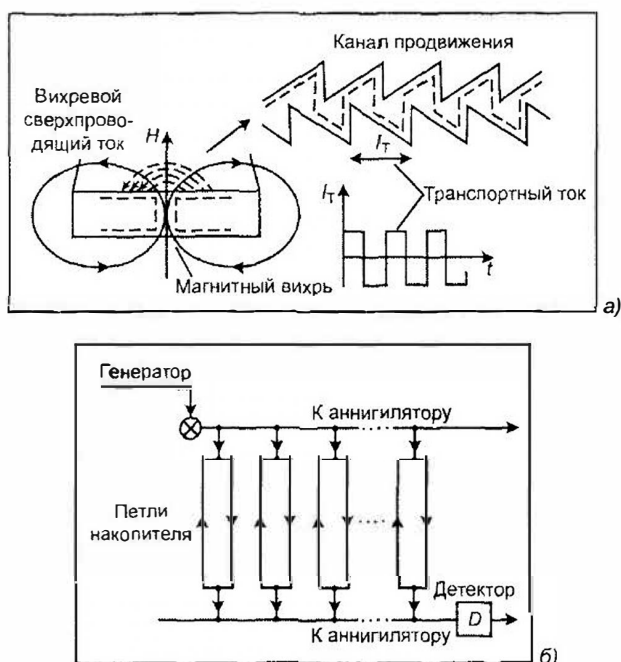


Рис. 5.14. Магнитный вихрь и его канал продвижения (а), мажор-минорная организация ЗУ на флюксонах (б)

Контрольные вопросы

1. Что такое функциональная магнитоэлектроника?
2. Что такое ЦМД и какими свойствами он обладает?
3. Какие динамические неоднородности магнитоэлектронной природы вы знаете?
4. Что такое ВБЛ и какими свойствами он обладает?
5. Что представляют собой магнитостатические волны?
6. Как устроены логические элементы на основе ЦМД-структур?
7. Опишите работу доменопередвигающей структуры.
8. Как происходит генерация и детектирование ЦМД?
9. Опишите конструкцию ЗУ на ЦМД. Какими отличительными свойствами оно обладает?
10. Как устроена линия задержки на МСВ?

Рекомендуемая литература

1. Звездин А. К., Попков А. Ф. Магнитные вихри в сверхпроводниках второго рода. — Электронная промышленность, № 8, 1983.
2. Элементы и устройства на цилиндрических магнитных доменах. Под ред. Н. Н. Евтихьева и Б. Н. Наумова. Справочник. — М.: Радио и связь, 1987.
3. Щука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

6. Функциональная оптоэлектроника

6.1. Физические основы

Функциональная оптоэлектроника представляет собой направление в функциональной электронике, изучающее явления взаимодействия динамических неоднородностей с электромагнитными полями в оптической континуальной среде, в том числе и в активной, а также возможность создания приборов и устройств обработки и хранения информации.

Развитие оптоэлектроники связано с успехами в области квантовой электроники, полупроводниковой электроники, физики твердого тела, оптики. Приборы и устройства оптоэлектроники работают в диапазоне длин волн от 0,2 до 20 мкм. Преимущества при передаче и хранении информации, в случае использования оптического излучения, обусловлены, прежде всего, электрической нейтральностью квантов оптического излучения — фотонов, а также высокой частотой электромагнитных колебаний.

6.1.1. Динамические неоднородности оптической природы

Динамические неоднородности оптической природы представляют собой электромагнитные волны. Это могут быть различные волны, как по форме, так и по спектру. В волне светового диапазона электрические и магнитные поля меняются в каждой точке пространства и во времени. Электрическая и магнитная компоненты связаны законами индукции и изменяются во времени по гармоническому закону. Малая длина волны света позволяет модулировать ее по времени и в пространстве.

Различают плоские волны, амплитуда и фаза которых в любой момент времени постоянны в плоскости распространения; параксиальные волны; сферические волны и др.

Волновой фронт, представляющий собой поверхность, во всех точках которой гармоническая волна имеет в данный момент времени одинаковую фазу, также является динамической неоднородностью оптической природы. Распространение такой динамической неоднородности происходит в направлении нормали к волновому фронту.

В качестве динамической неоднородности может использоваться волновой пакет или распространяющееся волновое поле, занимающее в каждый момент времени ограниченную область пространства. Такой волновой всплеск может быть разложен на сумму плоских монохроматических волн. Понятие волнового пакета связывает оптику с квантовой механикой. В частности, волновой пакет в квантовой механике означает, что вероятность нахождения в нем частицы велика. Вне области, занимаемой волновым пакетом, вероятность нахождения частицы практически равна нулю. При этом скорость распространения волнового пакета частицы совпадает с ее механической скоростью. Волновой пакет является волновой функцией такой частицы.

Весьма перспективным носителем информации является *солитон* (рис. 6.1, з), представляющий собой структурно-устойчивую уединенную волну в нелинейной диспергирующей среде. Распространяясь в континуальной среде с нелинейностями, в которой не вы-

полняется принцип неопределенности, солитоны влияют друг на друга. После взаимодействия солитоны не разрушаются, а расходятся, сохраняя исходные параметры. Солитоны подобно частицам могут образовывать связанные состояния из двух или более импульсов, а также специфическую среду, называемую солитонным газом. Солитонный импульс не предусматривает высокочастотного заполнения.

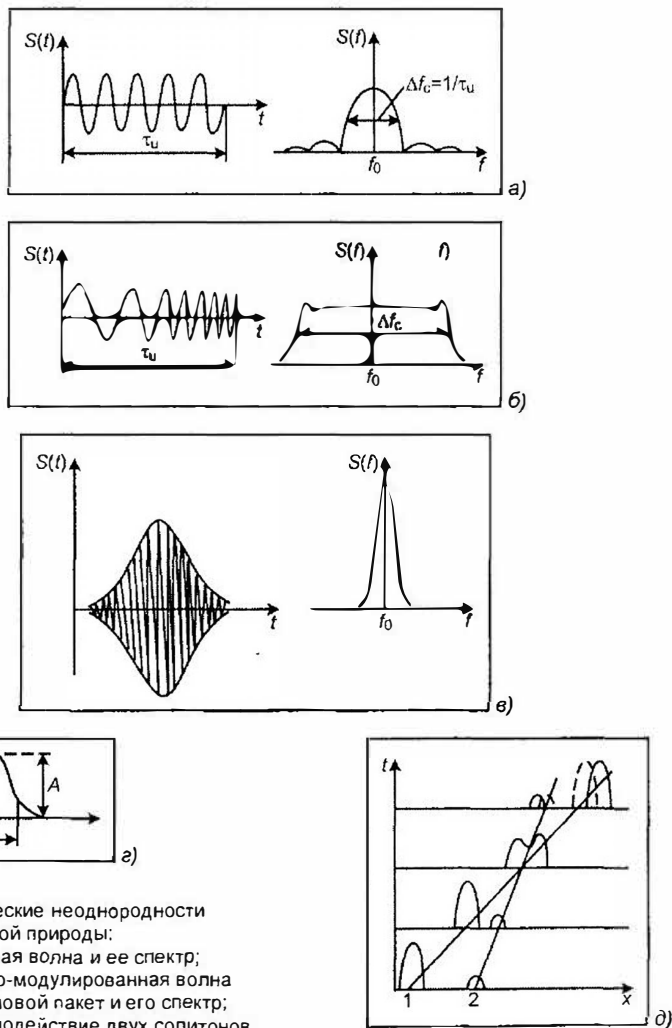


Рис. 6.1. Динамические неоднородности оптической природы:

- а — синусоидальная волна и ее спектр;
 б — линейно частотно-модулированная волна и ее спектр;
 в — волновой пакет и его спектр;
 г — солитон; д — взаимодействие двух солитонов

С помощью оптических солитонов в волоконных линиях связи возможно получение высоких скоростей передачи информации. Возникновение солитонов в световодах обусловлено двумя противоположными эффектами: уширением передаваемого импульса из-за дисперсии и ее компенсацией на нелинейностях волокна. Если через оптическое волокно пропускать короткие импульсы света ($\sim 10^{-10}$ с) с несущей частотой $\sim 10^{15}$ Гц, то передний фронт светового импульса возбуждает среду, и связанные электроны переходят на

верхний уровень. Задний фронт светового импульса возвращает возбужденные электроны на нижний уровень. Возникающее при этом излучение передает энергию световому импульсу.

Как показывают теоретические расчеты и эксперименты, солитоны проходят тысячи километров по световодному волокну без искажения формы. Следует также отметить, что резко возрастает эффективность передачи информации с помощью оптических солитонов в компьютерах с оптической памятью и оптическими линиями связи.

В активной среде существует световое (фотонное) эхо. Под фотонным эхом понимается импульс света, излучаемый системой атомов, предварительно облученных двумя когерентными световыми импульсами. Фотонное эхо наблюдается после второго импульса через время, равное временному интервалу между двумя возбуждающими импульсами. Различают первичное или двухимпульсное эхо и стимулированное (трехимпульсное) эхо (рис. 6.2).

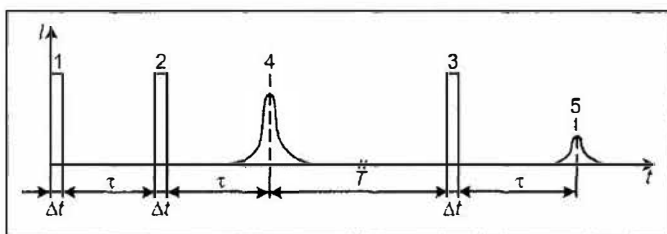


Рис. 6.2. Формирование сигналов светового (фотонного) эха:

1, 2 — записывающие световые импульсы; 3 — считывающий импульс;
4 — двухимпульсное первичное эхо; 5 — трехимпульсное стимулированное эхо

6.1.2. Континуальные среды

Континуальными средами в функциональной оптоэлектронике могут служить как пассивные оптические среды, так и активные. Под активной средой понимаются вещество, распространение частиц (атомов, молекул, ионов) в котором не является равновесным по энергетическим состояниям, а также среда, в которой меняется плоскость поляризации световой волны. В активных средах возможны процессы генерации динамических неоднородностей.

Пассивные оптические среды представляют собой каналы передачи оптического информационного сигнала. Это каналы высокой добротности и оптического качества, например оптические стекла, кварц и т. п.

Для генерации динамических неоднородностей часто используют эффективно люминесцирующие соединения типов $A^{III}B^V$, $A^{IV}B^{IV}$, твердые растворы на их основе: GaAs, InGa, GaP, GaAlAs, GaAsP, ZnS и т. д. В детекторах оптических сигналов используют фоточувствительные соединения типов $A^{III}B^{IV}$, $A^{III}B^V$, $A^{IV}B^{VI}$, например, CdS, CdSe, InAs, PbS, PbSnTe, CdHgTe и т. д. В качестве сред, каналирующих динамические неоднородности и управляющих их перемещением, используются оптические материалы, отличающиеся прозрачностью в заданном спектральном диапазоне, высокой однородностью и строго заданными управляемыми оптическими свойствами. Это могут быть электрооптические материалы ($LiTaO_3$, $LiNbO_3$), акустооптические материалы (TeO_2 , SiO_2 , Ge и др.), магнитооптические материалы (EuO , $MnBi$, $TmFeO_3$ и др.), а также высококачественные кварц, различные полимеры, многокомпонентные стекла и т. п.

В запоминающих устройствах используются светочувствительные материалы, которые позволяют хранить динамические неоднородности, а также преобразовывать их статические неоднородности и хранить таким образом "замороженные" динамические неоднородности. Существуют среды, позволяющие "размораживать" динамические неоднородности и считывать информационный сигнал.

Основными требованиями, предъявляемыми к запоминающим средам, являются высокая энергетическая чувствительность ($\sim 10^{-4}$ Дж/см²), высокая разрешающая способность ($\sim 10^4$ лин/мм), высокая контрастность ($\sim 1 : 100$), большая продолжительность хранения информационного массива (~ 10 лет), малая длительность цикла перезаписи информации для реверсивных материалов. К таким средам можно отнести фотографические материалы, фоторезисторы и фотохромные материалы, магнитооптические материалы, фототермопластики и термооптические материалы, халькогенидные стекла и т. д.

6.1.3. Генераторы и детекторы

Генераторами динамических неоднородностей оптической природы являются источники света различных типов. Выбор того или иного типа генераторов определяется, прежде всего, континуальной средой, в которой должна распространяться динамическая неоднородность. Различают когерентные и некогерентные генераторы оптического излучения. Излучение некогерентных генераторов представляет собой суммарный эффект независимых актов спонтанного испускания фотонов совокупностью возбужденных атомов и молекул. Неодновременность и отсутствие корреляции актов испускания фотонов приводит к хаотичному распределению фаз волн.

Наиболее привлекательным типом генераторов в функциональной оптоэлектронике считают светоизлучающие диоды и лазеры. Их излучение обладает временной и пространственной когерентностью. Именно это обстоятельство позволяет получать динамические неоднородности с высокой воспроизводимостью пространственных и временных параметров.

В качестве детекторов динамических неоднородностей оптической природы, как правило, используют различные типы фотоэлектрических устройств — фотодетекторы. Преобразовав оптический сигнал в электрический, можно легко обрабатывать информацию традиционными методами, включая машинные.

Детекторами могут служить и регистрирующие среды. Регистрирующие среды должны обладать свойством обратимости (реверсивности). Детекторы должны быть такими, чтобы можно было перезаписывать информацию, выделять разностную информацию, осуществлять объемное хранение информации.

6.1.4. Устройства управления

В устройствах управления динамическими неоднородностями оптической природы, как правило, используются электрические и магнитные поля, а также различные модуляторы. С помощью модуляции амплитуды, фазы, интенсивности, поляризации световой волны формируются соответствующие информационные массивы. Управление можно также осуществлять изменением свойств континуальной среды, ее геометрии.

Устройства управления в оптоэлектронике формируются индивидуально к каждому прибору. Поэтому весьма затруднительно выделить общие конструктивные решения, разве что электроды и магнитные элементы.

6.2. Процессоры функциональной оптоэлектроники

Оптоэлектронный процессор представляет собой совокупность оптических и оптоэлектронных элементов, выполняющий операции в соответствии с заданной функцией и алгоритмом обработки информации. Процессоры, реализованные на принципах функциональной оптоэлектроники, способны обрабатывать как аналоговую, так и цифровую информацию.

Традиционные оптроны соответствуют предложенной модели прибора функциональной электроники. Действительно, генератором динамических неоднородностей в виде волновых пакетов служит светоизлучающий диод. Континуальной средой служит либо воздушный промежуток между диодом и фотоприемником, либо световод. Управление можно осуществлять путем подачи соответствующего электрического импульса на светодиод, изменением геометрии световода. В качестве детектора динамических неоднородностей можно использовать различные фотоприемники.

Разработан новый тип переключаемых элементов, которые могут составить конкуренцию транзисторным ключам. Такой оптический аналог электронного транзистора получил название *трансфазор*. Первая часть слова дана от транзистора, вторая — от фазы волны, которая играет решающую роль в физике работы. В основе работы трансфазора лежит свойство материалов изменять свой показатель преломления в зависимости от интенсивности освещения. Изменяя показатель преломления, можно изменить скорость распространения света в среде. В свою очередь, с изменением скорости света меняется длина волны излучения.

Если такой материал поместить в резонатор Фабри-Перо, то меняя интенсивность пучка лазерного излучения, например, добавляя излучение $I_{\text{уп}}$ можно менять оптическую длину резонатора $n \cdot \lambda / 2$ (рис. 6.3, а, б). В этом случае может возникнуть ситуация, когда интенсивность выходного пучка увеличивается и интерферометр Фабри-Перо имеет характеристику бистабильной ячейки.

Зависимость интенсивности выходного пучка от суммы входных позволяет иметь состояния: "открыто — закрыто"; "да — нет"; "1 — 0". Форма петли гистерезиса бистабильного оптического прибора зависит от оптической длины резонатора L , длины волны падающего излучения λ и свойств самого вещества резонатора. Петлю гистерезиса можно сузить, изменить ее форму, однако точки логического нуля и единицы останутся.

Первые эксперименты были проведены с кристаллами антимонида индия InSb, обладающего сильно нелинейной характеристикой в инфракрасной части спектра ($\lambda \sim 5$ мкм). Интерферометр Фабри — Перо был сформирован с помощью специально подготовленных граней образца.

Один и тот же трансфазор может служить как элементом И, так и элементом ИЛИ в зависимости от световых сигналов.

Логические "1" и "0" кодируются высоким и низким уровнем пропускания или интенсивностью выходного светового сигнала. Если интенсивности $I_{\text{вх}} = I_{\text{д}} = I_{\text{уп}} = I_{\text{к}}$, то образуется элемент И, поскольку пучки взаимно когерентны и интерферируют в резонаторе трансфазора.

Если $I_{\text{вх}} = I_{\text{уп}} = I_{\text{к}}$, то формируется элемент ИЛИ, поскольку любой из этих входных световых сигналов способен переключить трансфазор. Если в качестве выходного сигнала использован отраженный пучок, который является инверсией прошедшего пучка, то трансфазор работает как элемент НЕ. В этом случае повышение интенсивности входного пучка до максимального значения уменьшает выходной сигнал до минимума (рис. 6.3, в, штри-

ховая линия). В этом случае также $I_{\text{вх}} = I_{\text{из}} = I_{\text{упр}} = I_{\text{к}}$, однако, зависимость $I_{\text{вых}} = f(I_{\text{к}})$ дополнительна относительно кривых для элементов И и ИЛИ. Обычно во всех схемных решениях до конца не решена проблема межэлементных соединений и подвода световых пучков к трансфазору. Проблему межсоединений в этом случае из плоскости электрических соединений перевели в плоскость оптических соединений. Проблема принципиально не снимается с повестки дня, хотя световые пучки и не взаимодействуют друг с другом. Можно сконструировать и интегральный трансфазор. Это обстоятельство позволяет создать большое число параллельных каналов обработки информации. Преобразование информации в канале осуществляется последовательно от каскада к каскаду. Предполагается, что построенная на основе трансфазоров вычислительная система позволит достичь быстродействия $\sim 10^9$ операций/с.

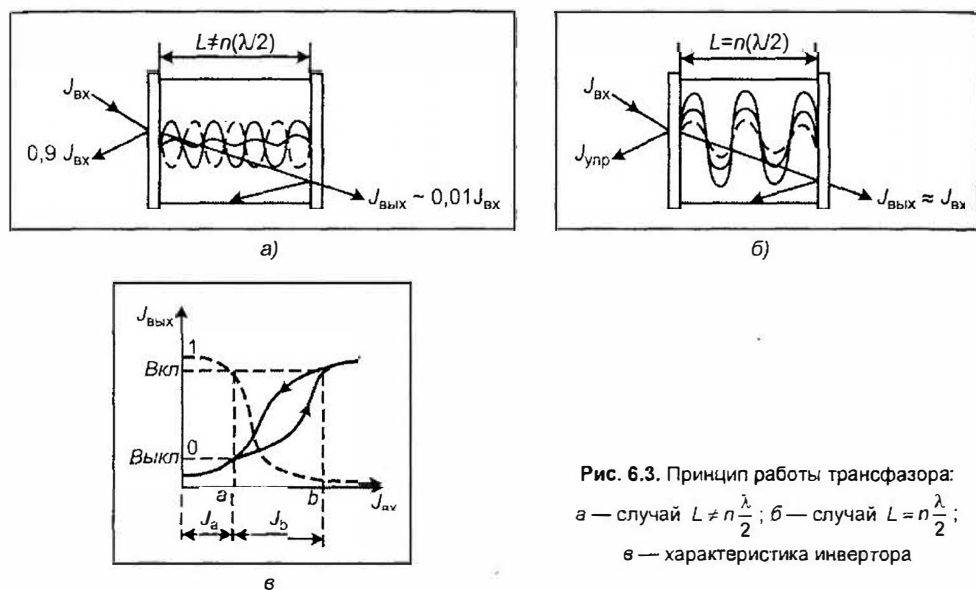


Рис. 6.3. Принцип работы трансфазора:

а — случай $L \neq n \frac{\lambda}{2}$; б — случай $L = n \frac{\lambda}{2}$;

в — характеристика инвертора

Достижения в технологии интегральной оптики позволили создать логические элементы, способные конкурировать с аналогичными электронными приборами. Базовыми элементами волноводной логики являются электрооптические модуляторы и переключатели. В качестве динамической неоднородности в таких устройствах используется оптический волновой пучок с линейной поляризацией. Континуальной средой служит оптическое волокно высокого качества. Управление динамическими неоднородностями осуществляется с помощью электрических полей путем поворота плоскости поляризации.

Волноводный модулятор представляет собой интерференционный прибор, в котором входной световой пучок расщепляется на два канала с определенной оптической длиной. Волноводы каналов симметричны и обладают электрооптическим эффектом (рис. 6.4, а). Вследствие симметричности плеч входной световой сигнал делится на две равные части, а затем на выходе восстанавливаются его пространственно-временные характеристики. Если же к одному из электродов, например к 1, приложить напряжение U_0 , то произойдет сдвиг фаз светового пучка в этом плече интерферометра. Разность фаз может составить $2\pi K$ или $(2\pi K + 1)$ радиан. В первом случае волны на выходе складываются и, следовательно, сигнал восстанавливается. Во втором случае происходит интерференционное вы-

читание (деструктивная интерференция) и формируется нулевой выходной сигнал. На рис. 6.4 приведены логические схемы конъюнктора (б), дизъюнктора (в), инвертора (д). На основе этих схем можно составить различные логические схемы подобно тому, как это делается в традиционной интегральной электронике.

Среди других процессоров, использующих динамические неоднородности оптической природы, следует указать на оптические управляемые транспаранты. С их помощью можно реализовать многоканальную параллельную обработку информационного массива с достаточно высоким быстродействием.

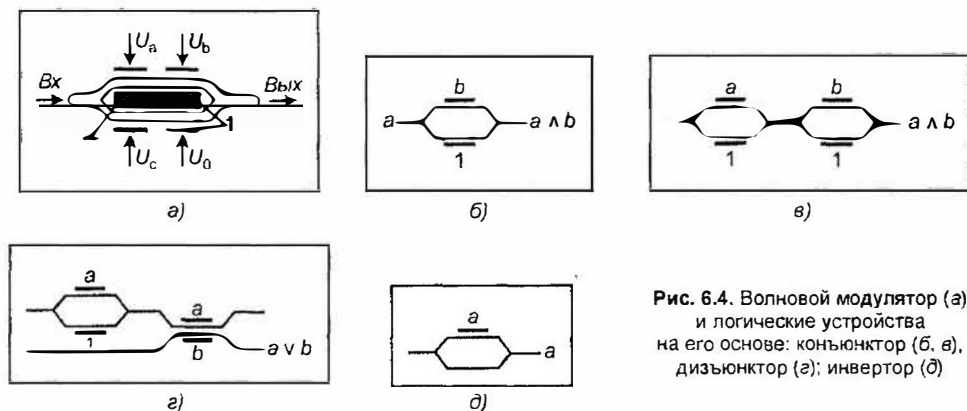


Рис. 6.4. Волновой модулятор (а) и логические устройства на его основе: конъюнктор (б, в), дизъюнктор (г); инвертор (д)

Уместно заметить, что поскольку выполнение всех логических и арифметических операций основано на булевой алгебре в двоичной системе типа "1 — 0", то определяющая роль отводится минимизации логических выражений в процессе вычислений. При построении форм сложных логических функций используют элементарные функции И, ИЛИ, НЕ и их комбинации. Способы реализации этих функций были рассмотрены выше.

Можно обойтись без разложения информационного массива на минимальные формы, т. е. формы, содержащие наименьшее число знаков двоичных переменных и их отрицаний, а также знаков логических операций.

С этой целью необходимо производить обработку информации в аналоговой форме с помощью функций высшего порядка. Примером может служить преобразование Фурье, осуществляемое тонкой линзой.

6.3. Запоминающие устройства функциональной оптоэлектроники

Среди различных типов ЗУ весьма перспективными являются *оптические ЗУ*. С появлением оптических дисков ожидалось рекордные показатели по емкости памяти и скорости выборки. Различают оптические ЗУ (диски) одnorазовой и многократовой записи. В отличие от накопителей (ЗУ) винчестерного типа, оптические диски одnorазовой записи не позволяют стирать информацию. В этом случае носителем информации являются статические неоднородности, формируемые в континуальной среде в виде ямок, пузырьков, областей с измененным фазовым расстоянием. Это направление разработок к функциональной оптоэлектронике относить не будем.

Существуют конструкции накопителей на оптических дисках, использующие реверсивную континуальную среду. К таким средам относятся полимерные красители, подверженные фазовым изменениям под действием лазерного излучения. Наиболее перспективными являются среды, в которых динамическая неоднородность в виде магнитного домена формируется лазерным лучом на принципах магнитооптического эффекта (рис. 6.5).

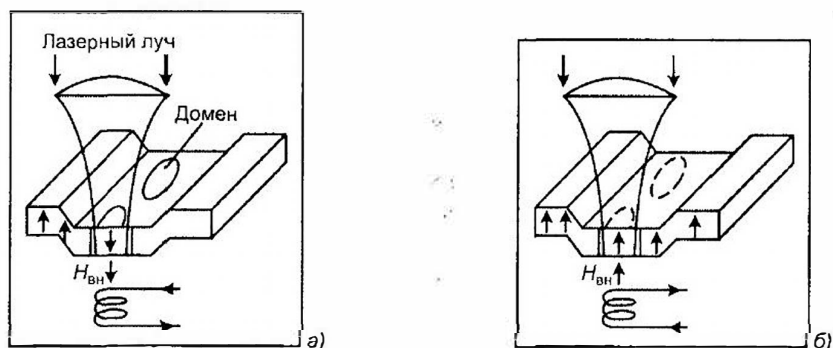


Рис. 6.5. Схема записи (а) и стирания (б) на оптическом диске

Лазерный луч нагревает пятно на поверхности диска. Магнитное поле $\vec{H}_{вн} \uparrow$, ориентированное соответствующим образом, намагничивает домен определенного направления. Стирание производится полем, ориентированным в противоположном направлении $\vec{H}_{вн} \downarrow$. Сформированная магнитным полем динамическая неоднородность в виде магнитного домена легко считывается. Стирание производится на втором или заданном оборотах путем возвращения поляризации домена в исходное состояние.

Магнитооптическая технология обеспечивает $\sim 10^7$ циклов перезаписи, высокое быстродействие. Общий объем записанной информации $\sim 10^{12}$ бит. Это еще один "побитовый" способ организации ЗУ, возможности которого ограничены.

Однако интерес к оптическим ЗУ сохраняется в связи с разработкой перспективных ЗУ на основе использования динамических неоднородностей в виде стимулированного светового (фотоинного) эха. Фотонным эхом является импульс света, спонтанно излучаемый системой атомов, предварительно облученных двумя когерентными световыми импульсами.

На рис. 6.6, а представлена последовательность воздействия импульсов света. Для записи информации используются два импульса (1 и 2) длительностью Δt с интервалом $\tau > \Delta t$. Волновые векторы этих импульсов направлены под углом друг к другу и пересекаются в заданной ячейке памяти. Средой для записи может служить, например, кристалл трехфтористого лантана. Возникает двухимпульсное первичное эхо 4, следующее через время τ .

Считывание информации осуществляется третьим импульсом 3, следующим через время $T \gg \tau$ после первого импульса (рис. 6.6, а). Через время $T + \tau$ после первого импульса возникает трехимпульсное стимулированное эхо 5, которое и несет информацию о проведенной ранее записи. Возникает также сигнал распада возбужденного состояния 6. Схема ЗУ приведена на рис. 6.6, б. Излучение импульсного лазера 7 делителем 8 делится на два импульса, имеющих разные волновые векторы \vec{k}_1 и \vec{k}_2 . Два импульса попадают на кристалл LiF_3 и в нем формируют ячейку памяти 10.

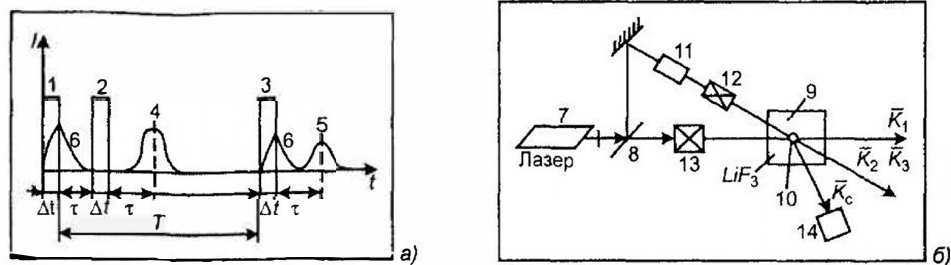


Рис. 6.6. Генерация фотонного эха (а) и ЗУ на его основе (б)

Для управления лучами используется линия задержки 11, оптический затвор 12 и система отклонения луча 13. Волновой вектор сигнала \vec{k}_c детектируется с помощью фотоэлектронного умножителя 14. Время хранения информации в таком ЗУ составляет ~ 5 минут. Поэтому для перезаписи 10^{15} бит информации требуется 10^4 диодных лазеров (решетка 100×100), система отклонения лучей, обеспечивающая их наведение в любую из 10^{15} ячеек памяти. Точность наведения составляла при этом $0,1 \text{ мкм}$, а фокусировка должна осуществляться в объеме 1 мкм^3 .

Особенностями ЗУ на основе фотонного эха являются высокая плотность записи информации ($> 10^{12} \text{ бит/см}^2$), отсутствие межсоединений в элементах памяти, возможность произвольной выборки информации, малое потребление энергии, возможность перезаписи информации без предварительного стирания.

Контрольные вопросы

1. Что такое функциональная оптоэлектроника?
2. Какие динамические неоднородности оптической природы вы знаете?
3. Что такое солитон и как он используется в световолоконных системах связи?
4. Что такое фотонное эхо?
5. Какие ЗУ функциональной электроники вы знаете?
6. Что такое трансфазор и какими свойствами он обладает?
7. Что такое волноводный модулятор?

Рекомендуемая литература

1. Абловиц М., Сигур Х. Солитоны и метод обратной задачи. — М.: Мир, 1987.
2. Кухаркин Е. С. Электрофизика информационных систем. Учебное пособие. — М.: Высшая школа, 2001.
3. Новокшенов В. Ю. Введение в теорию солитонов. — Ижевск: Институт компьютерных исследований, 2002.
4. Слепов Н. Н. Современные технологии цифровых оптоволоконных сетей связи. — М.: Радио и связь, 2000.
5. Щука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

7. Функциональная молекулярная электроника

7.1. Физические основы

Функциональная молекулярная электроника представляет собой направление в функциональной электронике, в котором изучаются процессы переноса и хранения информационного сигнала в молекулярных системах, а также разработка молекулярно-инженерных технологий для создания приборов и устройств обработки и хранения информации.

Идеи молекулярной электроники несколько отличаются от идей микроэлектроники. В процессе создания и микроминиатюризации приборов микроэлектроники создаются технологические процессы, связанные с удалением лишнего материала, изменением проводящих свойств материала с целью создания физических барьеров и переходов.

В молекулярной электронике развиваются методы конструирования и изготовления органических молекул с заданными свойствами, методы агрегации молекул нескольких типов. Первые методы связаны с созданием материалов с заданными электрическими свойствами путем подбора размеров молекул, их формы, взаимного пространственного их расположения, параметров различных функциональных групп молекул. Этот метод получил название *молекулярная инженерия* и с его помощью разработана концепция создания молекулярного электронного прибора.

Молекулярный электронный прибор приобрел статус микроэлектронного прибора после того, как были изготовлены переключатели и инверторы на молекулярном уровне. Проводниками, линиями межсоединений в таких структурах служат одномерные полимеры типа трансполИАцетилена $(CH)_x$ или нитрида хлора $(SN)_x$. Электрический механизм переключения на молекулярном уровне эквивалентен изменению валентности молекул. Валентность же связана с туннелированием электронов через неширокие периодические решетки, образованные молекулами (рис. 7.1).

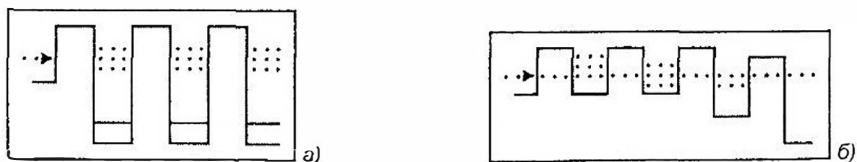


Рис. 7.1. Туннелирование электрона сквозь периодическую молекулярную решетку

Электрон способен преодолеть периодическую молекулярную решетку только в том случае, если его энергия равна или больше энергетического барьера внутри решетки. Электроны не могут преодолеть молекулярную периодическую решетку, если их энергия соответствует величине каждого энергетического барьера внутри решетки (рис. 7.1, а), и свободно проходят через систему потенциальных барьеров при условии равенства или превышения энергии псевдостационарного уровня (рис. 7.1, б).

Идея переключения в подобных структурах будет осуществляться, если высота барьера или глубина ямы регулируются внешними факторами. Среди таких механизмов отметим перемещение положительного или отрицательного заряда внутри молекулярной цепочки, переключение потока туннелирующих электронов путем смещения высоты псевдоэнергетических барьеров. Этот метод развивает схемотехнические принципы обработки информации и ее хранения.

Молекулярный аналог элементарной логической ячейки И — НЕ строится на базе тетрамерной производной с диазосвязями (рис. 7.2).

Периодический потенциал формируется четвертичными атомами азота, входящими в структуру. Две из четырех контрольных групп могут управляться потоками заряда через цепочки $(\text{CH})_x$ путем нейтрализации положительного заряда.

Молекулярная ячейка типа ИЛИ — НЕ представляет собой набор колец фталоцианида галлия, связанных фтором. Соединения типа $\text{Ni} - \text{S}$ обеспечивают заземление и связь с отрицательным потенциалом, а также с выходным выводом $(\text{SN})_n$ (рис. 7.2, б). Эти базовые элементы могут стать составными частями биологических компьютерных систем. Размеры структур логических ячеек могут составлять менее одной сотой размера полупроводниковой логической ячейки. Ожидаемая плотность размещения составит 10^{18} вентилях/см³. При всей привлекательности такой идеи молекулярных схем (даже с точки зрения использования при создании компьютеров) в ней содержится врожденный порок схемотехники. Речь идет о тех же схемотехнических решениях, тех же проводах, хотя это уже не плоскостные тоководущие дорожки, а молекулярные цепочки. Все это не исключает возможность возникновения традиционных для схемотехнических решений недостатков, а также новых, специфика которых состоит в налаживании надежных контактов между отдельными соединениями.

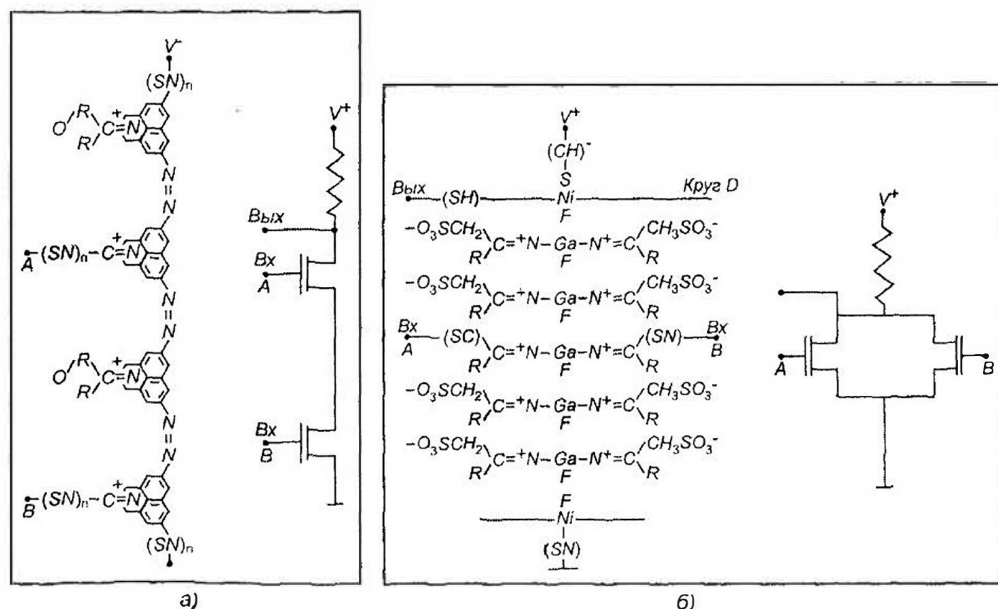


Рис. 7.2. Молекулярные логические вентиля типов И — НЕ (а) и ИЛИ — НЕ (б)

Методы агрегации определенного числа молекул нескольких типов или межмолекулярной самосборки позволяют получать заданные размеры и форму функционального элемента за счет выбора параметров, участвующих в самосборке молекулы, создать серию функциональных элементов без разброса параметров со строгой атомной детализацией.

Для получения сверхтонких пленок используется метод Ленгмюра — Блоджетт. Суть метода сводится к использованию нерастворимых поверхностно-активных веществ. Эти вещества формируют из двух фрагментов. Один фрагмент представляет собой гидрофильное вещество, имеющее сродство к воде, и хорошо в ней растворяющееся. Вторым компонентом — гидрофобный. Он не растворяется в воде и не позволяет молекуле поверхностно-активного вещества погрузиться в воду. Формирующиеся на поверхности воды однородные мономолекулярные пленки могут быть перенесены на твердые структуры подложки с различными типами слоев.

7.1.1. Динамические неоднородности

Передача информации в молекулярной электронике осуществляется ансамблями электронов или солитонов.

В масштабах микромира под солитоном понимают структурное возмущение, способное перемещаться в одном или двух направлениях подобно частице. Перенос солитона связан с возмущением, которое меняет положение молекулярных, так называемых π -орбиталей между атомами углерода. В центре солитона существует движущаяся *межфазная* или *междолинная* граница между эквивалентными фазами *A* и *B* (рис. 7.3, а). Прохождение солитона через сопряженную систему приводит к переходу между фазами *A* и *B* и к обмену одинарных и двойных связей.

Солитоны могут быть интерпретированы как топологические узлы валентной π -электронной системы.

Вследствие врожденности основного состояния электропроводящих полимеров солитоны могут свободно перемещаться вдоль цепи сопряжений подобно волне плотности π -электронов (рис. 7.3). Скорость перемещения солитонов близка к скорости звука.

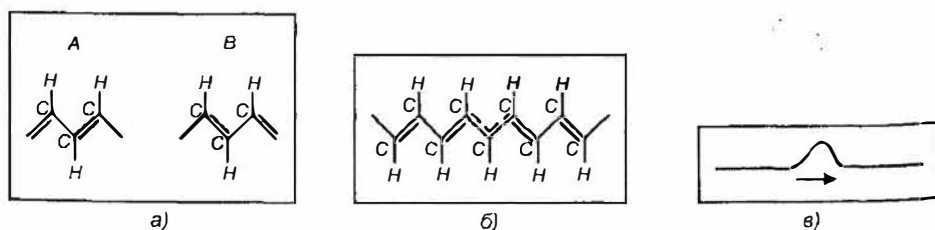


Рис. 7.3. Генерация и распространение солитона в полиацетиленовой цепи (а, б); интерпретация в виде волны возмущения (в)

Солитон обладает свойствами квазичастицы, имеет определенную энергию, форму, момент импульса. На молекулярном уровне также наблюдается реверсирование солитонов.

Различают два режима отражения солитонов. Две трансполиацетиленовые цепочки сопрягаются с трициклическим пентоном. Солитон проходит, отражаясь от пентона словно от стенки (рис. 7.4, а). Другой вариант предусматривает использование гидрированного каротена в качестве реверсивной среды (рис. 7.4, б). В этом случае солитон обходит молекулу по часовой стрелке. Движение валентных π -электронов происходит по про-

нумерованному порядку. Движение электрона со стрелкой под номером 10 может произойти только после того, как завершится движение со стрелкой 1. Солитон проходит по одной цепочке дважды.

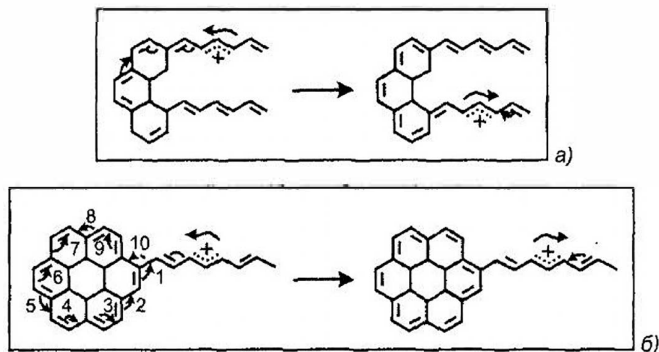


Рис. 7.4. Процесс реверсирования солитонов

Наряду с солитонами в молекулярной системотехнике используются процессы переноса электронов по цепочкам различных циклических молекул, даже разделенных зазором до 1 нм. Электроны туннелируют по молекулам, переходя в соседние молекулы за время $\sim 10^{-10} \div 10^{-12}$ с. Основным эффектом при этом является быстрое резонансное и одновременно однонаправленное туннелирование за счет того, что уровень для лишнего электрона в каждой последующей молекуле на 0,1 ÷ 0,2 эВ ниже, чем в предыдущей. Эффект быстрого резонансного туннелирования проявлялся в континуальных молекулярных средах, у которых молекулы в цепочке подобраны так, что энергия электрона, поляризовавшего одну молекулу, равна энергии электронного уровня другой, не поляризованной молекулы. Другими словами, цепочка молекул должна строиться так, чтобы энергии уровней для лишнего электрона снижались на 0,1—0,2 эВ.

В ленгмюровских пленках с помощью света можно возбудить экситоны, представляющие собой мигрирующее электронное возбуждение, не связанное с переносом электрического заряда и массы.

В качестве динамической неоднородности можно использовать один электрон. Это позволяет достичь энергетического предела, минимизировать энергию на одну информационную операцию. Если между отдельными молекулами создать разность потенциалов V , а энергетический барьер по полю составляет

$$E_{\text{уп}} = E_{\text{ср}},$$

то в обратном направлении этот барьер будет определяться величиной

$$E_{\text{обр}} = E_{\text{ср}} + eV,$$

где e — заряд электрона. Барьер одновременно будет препятствовать движению электрона в обратном направлении. Однако использование только одного электрона накладывают значительные ограничения на "одноэлектронику".

7.1.2. Континуальные среды

Ленгмюровские пленки представляют собой многофункциональную континуальную среду. На основе пленок Ленгмюра — Блоджетт (ЛБ) можно создать молекулярно-гладкую

континуальную среду со свойствами проводника, подзатворного диэлектрика, плазмостойкого резистора, пьезоэлектрика и биосенсора (рис. 7.5).

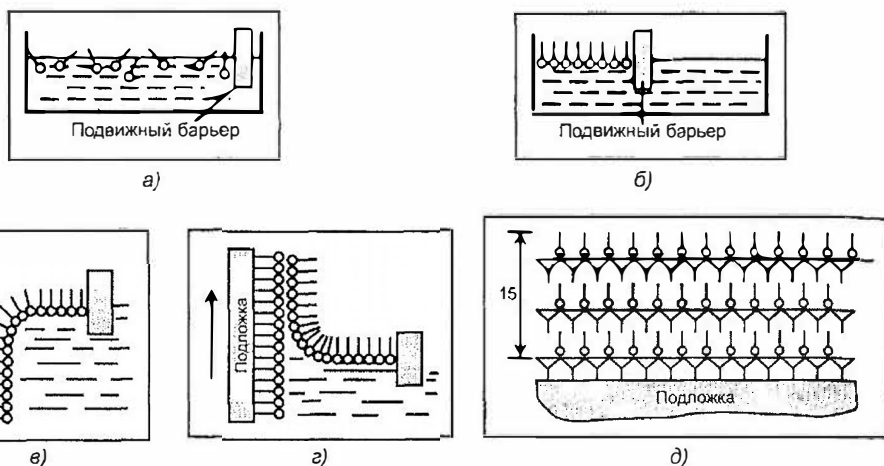


Рис. 7.5. Высаживаемые на различные подложки пленки Ленгмюра — Блоджетт, состоящие из разного типа молекул (а, б, в, г), и создание континуальных сред на их основе (д)

На рис. 7.6 показаны структуры ряда синтезированных органических полимеров.

Весьма перспективной средой для функциональной молекулярной электроники являются органические полимеры. Они состоят из цепочки слабосвязанных молекул с частично заполненными зонами валентных электронов.

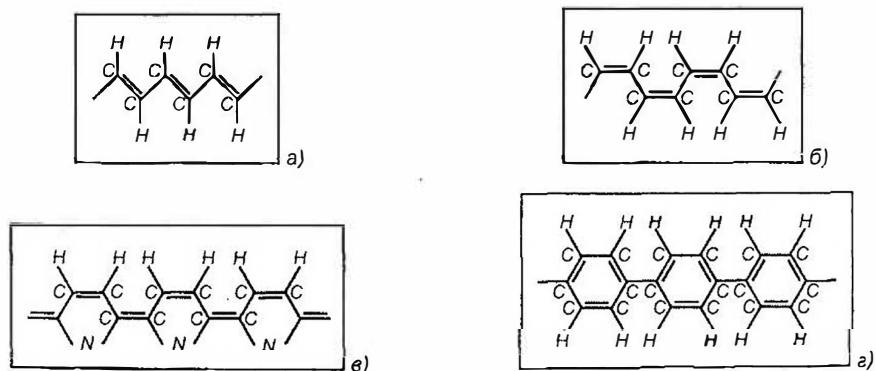


Рис. 7.6. Структуры полимерных континуальных сред: а — полиэтилен (трансформа); б — полиэтилен (цисформа); в — полипропил; г — полифинитен

Квазиодномерный органический полимер полиацетилен может быть легирован донорами (K, Na, Li) и акцепторами (Br, AsF₃, PF₆), и имеет слабое взаимодействие электронов с решеткой.

Цепочки образуют поликристаллические волокна диаметром 20 нм. Пленка полиацетилена представляет собой переплетенные волокна (паутина). Различают две модификации

полиацетилена: цис- и трансизомеры. Полиацетилены могут быть получены в виде молекулярных кристаллов диацетилена (мономера).

Возможно получение пленок из поликристаллов полиацетилена. Полипролл представляет собой пленки с локальным кристаллическим порядком. В них можно осуществить регулярность в расположении легирующих примесей и стехиометричность состава легируемых материалов. Полифинилены существуют в виде порошков и пленок. Это далеко не полный перечень сред. Число известных синтезированных органических веществ превысило пять миллионов.

7.1.3. Другие элементы приборов

Динамическими неоднородностями типа электронов и солитонов можно управлять. В молекулярной системе, в которой двойная связь является частью большой полиацетиленовой цепочкой, под воздействием поляризованной фотоактивации включается процесс переноса электронов (рис. 7.7). При прохождении солитона (см. нижнюю цепочку) фотоактивационный процесс становится невозможным.

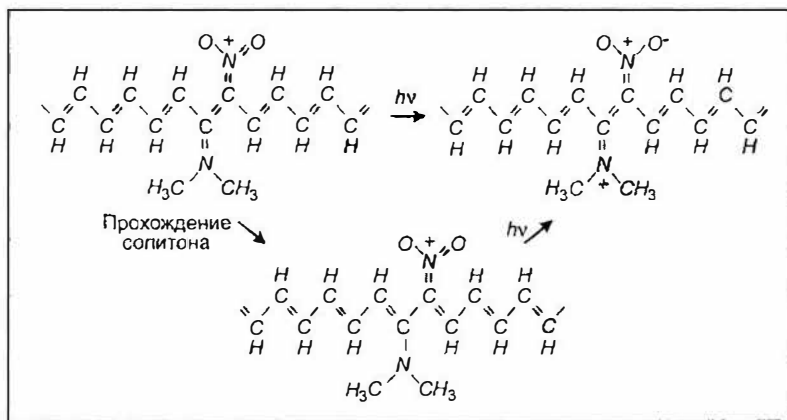


Рис. 7.7. Молекулярный детектор солитонов

Другими словами, солитон включает реакцию внутреннего переноса заряда, а изменения в спектре поглощения олефина может служить детектором, регистрирующим прохождение солитона.

Генерация солитонов осуществляется в процессе протонного туннелирования, например, между вторичным анином и кетонной группой трансполиацетилена в присутствии электрического поля.

Итак, показана возможность генерации, управления и детектирования солитонов. Солитон, являясь динамической неоднородностью, способен переносить информацию; с его помощью можно и хранить информацию.

7.2. Молекулярные устройства

Хранение бита информации в солитонной памяти определяется наличием или отсутствием солитона, а число хранимых битов зависит от скорости распространения солитона и

длины сопряженного полимера (рис. 7.8). Сопряженный полимер связывает генератор солитонов (ГС) и электронный туннельный переключатель (ТП). Показана возможность накопления четырех битов информации. При одновременном прибытии положительного и отрицательного солитонов срабатывает туннельный переключатель. Солитонные устройства памяти не имеют высокого быстродействия, однако, отличаются высокой плотностью хранения информации. Оценки показывают, что если расстояние между солитонами составляет $\sim 200 \text{ \AA}$, а расстояние между центрами цепочек $\sim 50 \text{ \AA}$, то информационная емкость близка к $\sim 10^{18}$ бит/см³.

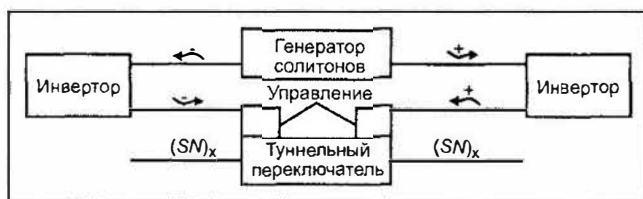


Рис. 7.8. Схема устройства памяти, реализованная на солитонах

В устройствах памяти могут быть использованы и пленки Лентгюра — Блоджетт. Каждая молекулярная система имеет свою частоту поглощения. Запись информации осуществляется избирательно в ансамбль молекул трехмерной молекулярной пленки с помощью лазерного излучения определенной частоты. Считывание информации осуществляется за счет обратных физических явлений. Использование свойств молекулярных структур позволяет создать трехмерные запоминающие устройства высокой информационной емкости.

Существуют различные идеи построения молекулярных компьютеров. Их микросборка, например, может быть осуществлена химическим осаждением из газовой фазы. Последовательность химических превращений при создании молекулярной вычислительной машины рассчитывается по заданному алгоритму.

В итоге к базовой молекуле добавляются все новые и новые звенья, формируются проводники, изоляторы, вентили, компоненты вычислительных устройств. Так, например, размер молекулярного вентиля может быть меньше одной сотой микрометра, а плотность сборки может достичь 10^{18} вентилях/см³.

Процессор и память гипотетического молекулярного компьютера занимают объем $\sim 1 \text{ см}^2$ и монтируются на криостате, который предотвращает нагрев контактов и снижает химическую активность молекул. Для уменьшения количества межсоединений предусматривается использование оптических каналов для ввода-вывода информации. Обработка информационных массивов молекулярными системами имеет ряд особенностей, среди которых важнейшими являются:

- ☐ гигантский параллелизм переработки информации;
- ☐ высокая эффективность преобразования информации;
- ☐ значительная информационная сложность исходных операций;
- ☐ способность к изменчивости и эволюции молекулярных компонентов устройств переработки информации и эволюционному обучению (самообучению) устройств;
- ☐ динамические механизмы переработки информации, основанные на сложных нелинейных процессах.

Другим примером высокой производительности может служить процессор, реализованный на пленках Ленгмюра — Блоджетт. На таких пленках можно создать молекулярные системы с высокой степенью параллелизма обработки информации и большой вычислительной мощностью.

В качестве примера приведем идею сложения двух образов, проецируемых на искусственную мембрану типа ленгмюровской пленки со встроенными молекулами фотоактивируемого фермента (динамические неоднородности). В этом случае выходной сигнал, представляющий собой концентрацию продукта ферменто-активной реакции, будет пропорционален сумме оптических сигналов на входе системы.

Оценки показывают, что при поверхности пленки $\sim 1 \text{ см}^2$ и при разрешающей способности 10 мкм число ячеек составит 10^6 . Каждая ячейка будет содержать $\sim 10^4 + 10^5$ молекул фермента при 1% заполнении поверхности. Если оценить время операции в одной ячейке даже $\sim 10^{-2} \text{ с}$, то, учитывая высокую степень параллелизма обработки информации ($\sim 10^6$), эффективное время одной операции составит $\sim 10^{-8} \text{ с}$. Другими словами, обработка больших информационных массивов методами ФЭ позволяет предельно распараллеливать информацию в процессе ее обработки. Современные компьютеры не имеют такой производительности при обработке больших массивов информации.

7.3. Автоволновая электроника

Особое место в молекулярной электронике занимают перспективные идеи автоволновой электроники.

В автоволновой электронике изучаются процессы, происходящие в неравновесных открытых кинетических системах, которые проявляются в виде разнообразных нелинейных автоволн, стационарных пространственных структур, стратов, доменов и других типов динамических неоднородностей.

Наибольший, пожалуй, интерес представляют собой автоволны — класс нелинейных волн, распространяющихся в активных континуальных средах за счет запасенной энергии. Следует особо подчеркнуть, что автоволны распространяются в средах, в каждом локальном объеме которых осуществляется независимая накачка энергии (равно как и вещества) и имеется диссипативный сток для отработанных "продуктов" накачки (рис. 7.9).

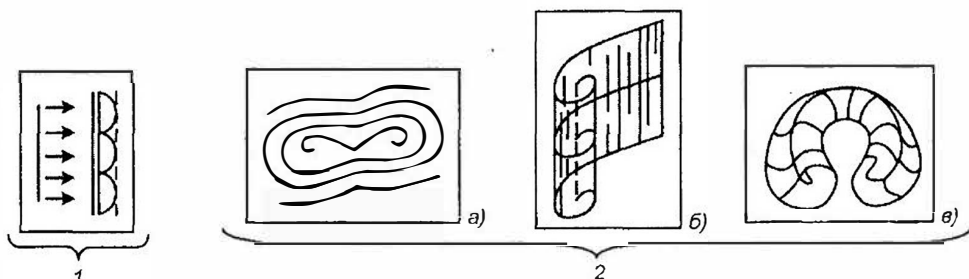


Рис. 7.9. Типы автоволновых процессов: 1 — плоский волновой фронт огибает препятствие; 2 — спиральные волны; а — двумерная среда, б — вращающийся свиток, в — вихревое облако

В качестве континуальных сред в устройствах автоволновой электроники используются активные распределенные среды, в которых возможна реализация большого числа само-

задерживающихся нелинейных пространственно-временных возбуждений. К таким средам можно отнести биологические активные мембраны, мышечные волокна, нейронные сети, распределенные биохимические среды, а также некоторые сообщества живых организмов.

Генератором динамических неоднородностей автоволновой природы являются: импульсы внешней накачки, различные флуктуации.

Автоволны способны аннигилировать при столкновении двух встречных волновых фронтов, не отражаются от препятствий и границ среды, однако дифрагируют на препятствиях в соответствии с принципом Гюйгенса. Различают плоские и спиральные волны (рис. 7.9).

Управление автоволновыми процессами может осуществляться с помощью тепловых полей, локальным энергетическим воздействием, введением дополнительной массы вещества.

Автоволновые среды обладают свойством ассоциативной памяти. Это обусловлено рядом причин, среди которых важной является отсутствие локальной пространственной адресации записываемой информации. Адресация и считывание возможны только по содержанию информации, по характерным признакам. Автоволновые среды способны регистрировать предысторию своего функционирования путем накопления определенных качеств или свойств в материале. Такие среды обладают распределенной памятью и способны к обучению под воздействием внешних условий.

Реализация памяти с использованием автоволновых процессов происходит по следующей схеме. В сосуде 1 находится вещество 2 и его расплав 3 (рис. 7.10). В режиме записи (рис. 7.10, а) опорный источник создает однородный поток тепла. Возникший градиент температуры $\frac{dT}{dy}$ возбуждает в диссипативной автоволновой среде конвективные ячейки.

Если на опорную матрицу воздействовать тепловым сигналом от источника В, то произойдет перестройка потоков, адекватная информационному воздействию потока В. Возникнет структура, описывающая поверхность изотермы плавления $\overline{a \cdot b}$. Эта структура характеризует структуру конвективных потоков, интенсивности тепловых выбросов и т. д. Другими словами, можно изготовить копии рельефа в виде $a \cdot b$. Таким образом реализуется память.

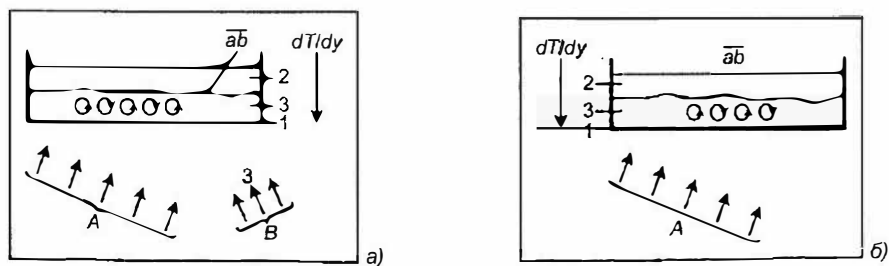


Рис. 7.10. Схема ассоциативной памяти: а — запись; б — воспроизведение структуры

Ассоциативное восстановление можно провести по схеме рис. 7.10, б. Условия неравновесности создаются источником А, а дополнительным организующим фактором является

структура рельефа $\overline{a \cdot b}$. Достижение комплементарного соответствия вложений $\overline{a \cdot b}$ и $a \cdot b$ является актом распознавания образа. Этот метод имеет некоторое сходство с волновой голографией, а именно — реализуется в опорную упорядоченность в условиях самоорганизации.

Можно ожидать, что в дальнейшем исследования будут развиваться с использованием синергетического подхода к поведению автоволновых сред. Под *синергетикой* будем понимать процессы, происходящие в системах, состоящих из многих подсистем самой различной природы, например, электроны, атомы, молекулы, клетки, нейроны, фотоны, животные организмы. Такие структуры возникают в процессах самоорганизации, переходов типов "беспорядок—порядок", "порядок—порядок" и т. п.

Контрольные вопросы

1. Что такое функциональная молекулярная электроника?
2. Что представляют собой пленки Ленгмюра — Блоджетт?
3. Какие динамические неоднородности молекулярной природы вы знаете?
4. Опишите возможность создания ЗУ на солитонах.
5. Что такое автоволны?

Рекомендуемая литература

1. Блинов Л. М. Ленгмюровские пленки. — УФИ, т. 155, вып. 3, 1988.
2. Гуляев А. М. Молекулярная электроника. Обзор. Шумовые и деградиационные процессы в полупроводниковых приборах. — М.: МИЭТ, 2003.
3. Давыдов А. С. Солитоны в молекулярных системах. — Киев: Наукова думка, 1984.
4. Жаботинский А. М., Отмер Х., Филд Р. Колебания и бегущие волны в химических системах. — М.: Мир, 1988.
5. Хакен Г. Синергетика. — М.: Мир, 1985.
6. Щука А. А. Функциональная электроника. Учебник для вузов. — М.: МИРЭА, 1998.

8. Приборы функциональной электроники второго поколения

Ко второму поколению приборов и устройств функциональной электроники отнесем устройства, использующие одновременно динамические неоднородности различной физической природы в различных континуальных средах.

В этих приборах прослеживается тенденция интеграции функций на основе интеграции эффектов и взаимодействия динамических неоднородностей различной физической природы в различных континуальных средах и, соответственно, расширение функциональных возможностей приборов и устройств.

8.1. Приборы с акустическим переносом зарядов

В приборах с акустическим переносом зарядов (ПАПЗ) используется взаимодействие двух типов динамических неоднородностей: поверхностной акустической волны (ПАВ) и зарядовых пакетов. В качестве континуальных сред используются слоистые структуры из пьезоэлектрика и полупроводника, в которых генерируются и взаимодействуют динамические неоднородности соответствующей природы. На рис. 8.1 приведена схема устройства прибора с акустическим переносом зарядов (ПАПЗ) на основе слоистой структуры типа $\text{ZnO}/\text{Al}/\text{SiO}_2/\text{Si}$. Поверхностная акустическая волна (ПАВ) возбуждается с помощью ВШП 1 в пьезоэлектрической среде (ZnO). Электрическое поле, создаваемое ПАВ, захватывает неосновные носители, сформированные в виде зарядовых пакетов, из другой полупроводниковой континуальной среды (Si). Зарядовые пакеты генерируются входным диодом 2 и переносятся к обратно смещенному выходному диоду 3.

Выходной диод является детектором динамических неоднородностей, и он преобразует зарядовый пакет в выходное напряжение, которое индицируется как логическая "1" или "0". Для изоляции канала от основных и неосновных носителей используются кольца смещения 4, которые позволяют изолировать канал переноса от зарядов полупроводникового слоя. Возбуждаемые ПАВ поглощаются в конце структуры поглотителем 5. Инжекция неосновных носителей заряда в канал переноса может осуществляться электрическим путем с помощью диода 2 или оптическим путем через центральный сегмент 6 облучением светодиодами 7. У приборов такого типа характерным недостатком является неэффективность переноса зарядов. Для увеличения эффективности переноса зарядов используются полупроводниковые континуальные среды типа GaAs , а также совмещение канала переноса зарядов с каналом распространения ПАВ.

При заданном значении эффективности переноса зарядов число электронов в зарядовом пакете чувствительно к изменению потенциала ПАВ.

Приборам с акустическим переносом присущи те же недостатки, что и ПЗС-приборам, например, неэффективность переноса заряда, небольшой динамический диапазон. Для увеличения динамического диапазона предложено использовать GaAs . При этом удалось совместить канал переноса ПАВ с каналом переноса зарядов.

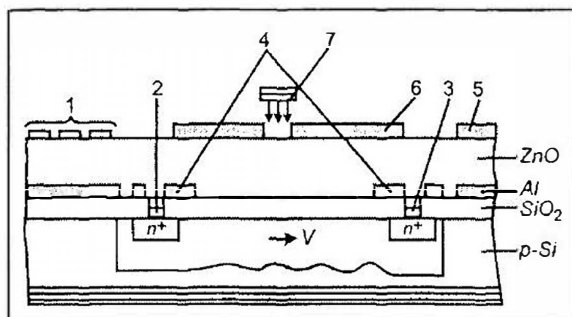


Рис. 8.1. ПАПЗ на основе слоистой структуры ZnO/Al/SiO₂/Si: 1 — входной ВШП; 2 — входной диод; 3 — выходной диод; 4 — кольца смещения; 5 — поглотитель ПАВ; 6 — центральный сегмент; 7 — светодиод

Перенос зарядов электрическим полем ПАВ имеет место, если напряженность поля

$$E_{3B} \geq \frac{V_{3B}}{\mu},$$

где V — скорость звука, μ — подвижность носителей.

При оптическом вводе информации кванты света генерируют электронно-дырочные пары. Поле ПАВ разделяет их в канале переноса, который находится в состоянии глубокого обеднения. Электроны захватываются электрическим полем и переносятся к выходному устройству. Ток выходного преобразования линейно зависит от оптической мощности сигнала. Для увеличения эффективности переноса канал необходимо изолировать, например, с помощью протонной бомбардировки. В ПАПЗ возможен процесс неразрушающего считывания зарядов.

Это обстоятельство позволяет создать на базе ПАПЗ процессоры сигналов, устройства памяти. Простейшие устройства линии задержки имеют полосу пропускания ~ 150 МГц, время задержки $\sim 10^6$ с.

Линии задержки являются базовыми конструкциями для трансверсальных фильтров и корреляторов.

Основу корреляторов на ПАПЗ составляют две линии задержки, акустические волны в которых распространяются в противоположных направлениях (рис. 8.2). Область нелинейного взаимодействия формируется диодами Шоттки. Катоды диодов соединены с выходными электродами, позволяющими производить пространственное интегрирование сигналов каждого из электродов. Эти электроды представляют собой металлизированные пластины (~ 1 мкм), расположенные перпендикулярно распространению ПАВ, генерируемых ВШП и отражателями на частоте 360 МГц.

Анализируя перспективу развития ПАПЗ следует отметить, что приборы этого класса сочетают в себе преимущества ПАВ-приборов в части высоких рабочих частот и ПЗС-приборов в отношении высокого динамического диапазона, гибкого управления. В этой связи интересны аналоговые процессоры для параллельной обработки информации.

Широкое применение ПАПЗ могут найти в вычислительных интерфейсах для связи аналоговых и цифровых устройств обработки радиосигналов. Применение ПАПЗ позволяет существенно уменьшить габариты, потребляемую мощность, стоимость устройств.

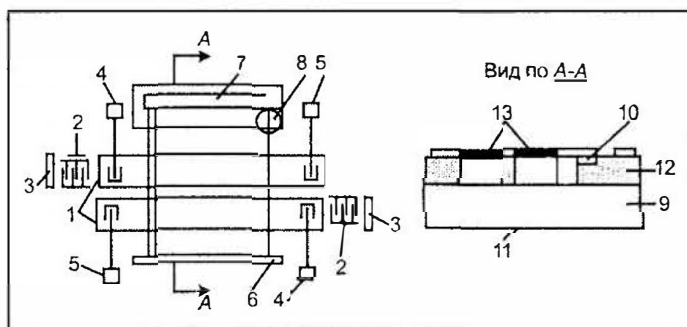


Рис. 8.2. Коррелятор на ПАЭП: 1 — линии задержки; 2 — ВШП; 3 — поглотители ПАВ; 4 — входные диоды; 5 — выходные диоды; 6 — резистор; 7 — выходной затвор; 8 — диоды области нелинейного взаимодействия; 9 — полупроводниковая подложка; 10 — граница области обеднения от затвора Шоттки; 11 — затвор; 12 — область переноса заряда; 13 — электроды неразрушающего считывания

8.2. Приборы акустооптики

В приборах и устройствах функциональной акустооптики используется взаимодействие двух типов динамических неоднородностей: акустической и оптической природы. Физической основой взаимодействия электромагнитных и упругих волн являются электрооптический и уругооптический эффекты.

Акустическая волна модулируется коэффициентом преломления континуальной среды, создавая в ней бегущую фазовую решетку. При выполнении брэгговских условий падения света на фазовую решетку, образованную звуковой волной, происходит рассеяние света. Угол рассеяния θ может быть определен из соотношения

$$\theta \cong \lambda / \Lambda = \lambda \cdot f_3 / V_3,$$

где λ и Λ — длины волн света и звука; V_3 и f_3 — скорость и частота звука.

К приборам акустооптики относятся дефлекторы и модуляторы.

Акустооптический дефлектор предназначен для сканирования, модуляции и переключения света в электронно-оптических системах обработки информации. Акустооптический дефлектор состоит из оптического волновода, сформированного в подложке LiNbO_3 путем имплантации Ti , а также призматических элементов ввода-вывода излучения 3. Волна возбуждается ВШП (рис. 8.3). Оптическое излучение, введенное в континуальную среду с фотоупругими свойствами, испытывает дифракцию на "замороженной" акустической фазовой решетке. В первый дифракционный максимум перекачивается значительная доля падающей световой энергии.

Акустооптические дефлекторы широко используются при обработке информации, т. к. имеют заданное число разрешимых фиксированных позиций от 10 до 125, диапазон частот управляющего сигнала лежит в пределах 130—825 МГц при мощности несколько ватт.

Аналогично устройство *акустооптического модулятора*, который позволяет управлять амплитудой, фазой, частотой и поляризацией световой волны. Такие модуляторы позволяют вносить новую информацию в световой пучок. Мощность модулируемого излучения составляет ~ 1 мВт, ширина полосы составляет 200 МГц.

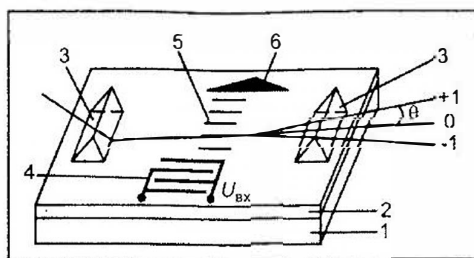


Рис. 8.3. Акустооптический дефлектор:

1 — подложка LiNbO_3 ;
 2 — волновод Ti:LiNbO_3 ; 3 — призма;
 4 — ВШП; 5 — фазовая решетка;
 6 — отражатель; 0 — основной пучок
 излучения; ± 1 — дифракционные
 максимумы первого порядка

Определенный интерес представляют *акустооптические процессоры*, которые эффективно используются в задачах параллельной обработки информации в реальном масштабе времени. Такие процессоры позволяют обрабатывать сигналы с частотой 2—3 ГГц при ширине полосы 0,5—1,0 ГГц в динамическом диапазоне 40—60 дБ. Различают корреляторы с пространственным и временным интегрированием, которые обеспечивают свертку или взаимную корреляционную функцию сигналов. Схема коррелятора с пространственным интегрированием приведена на рис. 8.4.

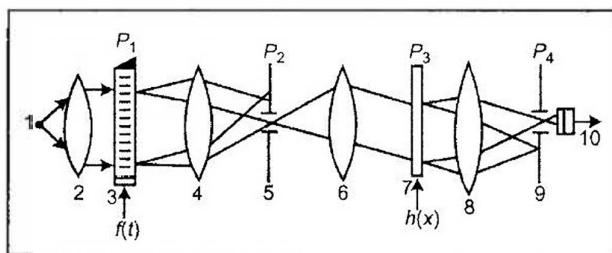


Рис. 8.4. Коррелятор с пространственным интегрированием

Сигнал $f(t)$ подается на пьезопреобразователь акустооптического модулятора 3. Модулятор представляет блок, способный создать акустооптическую копию сигнала $f(t)$ в виде зон с измененной оптической плотностью. Коэффициент преломления в зонах сжатия увеличивается, а в зонах разряжения уменьшается. При этом пучок света от источника 1, пройдя коллимирующую линзу 2, дифрагирует, а амплитуда дифрагирующей составляющей модулируется функцией

$$f\left(t - \frac{x}{V_s}\right),$$

где V_s — скорость звука в модуляторе. Оптическая система из линз 4, 6 и диафрагмы 5 формирует изображение в плоскости P_2 , в которой расположен транспарант 7 с амплитудным пропусканием, заданным функцией $h(t)$. Линза 8 осуществляет пространственное преобразование Фурье над волновым полем в плоскости P_3 . В результате имеем

$$g\left(\frac{2 \cdot \pi \cdot x}{\lambda \cdot F}, t\right) = f\left(t - \frac{x}{V_s}\right) \cdot h(x) \exp\left(\frac{j \cdot \pi \cdot x}{\lambda \cdot F}\right) dx,$$

где F — фокусное расстояние линзы 8, λ — длина световой волны, x — пространственная координата в выходной плоскости P_4 . Точечная диафрагма выделяет результат преобразования, световое распределение регистрируется фотодетектором 10 в виде:

$$I(t) \sim \left| \int \left(t - \frac{x}{V} \right) h(x) dx \right|^2.$$

Другими словами, этот сигнал содержит информацию о квадрате модуля взаимной корреляции функций $f(t)$ и $h(x)$. В корреляторе с временным интегрированием функция $h(t)$ вводится с помощью модулятора.

В задачах обработки больших массивов информации определенные надежды связывают с акустооптическими линейно алгебраическими процессорами, предназначенными для перемножения матрицы на вектор, матрицы на матрицу, тройного перемножения матриц и т. п.

В таких процессорах используются свойства акустооптических ячеек: возможность модуляции интенсивности дифрагированного пучка и его отклонения на угол, пропорциональный частоте акустических колебаний в ячейке (рис. 8.5).

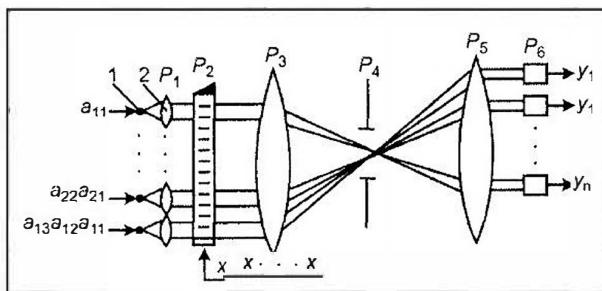


Рис. 8.5. Схема акустооптического алгебраического процессора

Элементы матрицы, соответствующим образом представленные сигналами a_{ij} , подаются на линейку источника света 1 в плоскости P_1 , их коллимированное излучение — на модулятор 3. Одновременно на модулятор подаются импульсы x_1, x_2, \dots, x_N , пропорциональные N -мерному вектору N . Результатом перемножения является вектор $y = Ax$, компоненты которого вычисляются по правилу:

$$y = \sum_{j=1}^N a_{ij} \cdot x_j \quad (i, j = 1 \dots N).$$

Оптическая система из линз 4, 6 и диафрагмы 5 проецирует результат на линейку фотодетекторов 7. Работа устройства синхронизирована так, что в момент времени входа в апертуру ячейки первого акустического импульса первый источник света (фотодиод) излучает световой импульс с заданной интенсивностью a_{11} , а на выходе первого детектора появляется импульс тока, пропорциональный $a_{11} \cdot x_1$. Следующий такт вычисления произойдет, когда первый импульс сместится и будет напротив второго светодиода, излучающего со световой интенсивностью, пропорциональной a_{21} .

В результате суммарный выходной сигнал первого фотодетектора пропорционален $a_{11} \cdot x_1 + a_{12} \cdot x_2$, а выходной сигнал второго фотодетектора $a_{21} \cdot x_1$. Через N таких тактов все элементы вектора x заполняют амплитуду ячейки, и величина сигнала нагрузки на i -ом фотодетекторе будет пропорциональна $\sum_{j=1}^{N-i+1} a_{ij} x_j$. Предел скорости обработки оценивается

величиной $(5-10)^{10}$ операций умножения/сложения в секунду. Аналогично строится процессор для перемножения матрицы на матрицу.

Акустооптические процессоры достаточно полно теоретически разработаны, известны алгоритмы обработки различных видов сигналов. Что касается технологических аспектов производства акустооптических процессоров, то не решены проблемы эффективных материалов для звукопроводов, не доведены до совершенства технологические процессы изготовления этого типа приборов.

В случае решения этих задач появится возможность разработать эффективную элементную базу для создания оптических интегральных схем.

Сравнительные оценки эффективности использования цифровых и оптических процессоров для корреляционного анализа радиолокационных сигналов показывают, что экономически выгоднее использовать акустооптические процессоры. Выяснилось также, что скорость обработки сигналов с полосой частот выше 300 МГц недоступна цифровым процессорам, и вполне эффективно обрабатывается акустооптическими процессорами.

В этой связи становится объяснимым внедрение в радиоэлектронные комплексы глобальной системы ориентации (Global Position System), системы электронного противодействия (Electronic Combat) США именно акустоэлектронных процессоров.

Контрольные вопросы

1. Что такое приборы функциональной электроники второго поколения?
2. Опишите работу прибора с акустическим переносом зарядов.
3. Какие приборы акустооптики вы знаете? Какие физические принципы лежат в основе их работы?
4. Как устроен акустический алгебраический процессор?

Рекомендуемая литература

1. Балацкий В. И., Парыгин В. Н., Чирков Л. Е. Физические основы акустооптики. — М.: Радио и связь, 1985.
2. Егоров Ю. В., Наумов К. П., Ушаков В. Н. Акустооптические процессоры. — М.: Радио и связь, 1991.
3. Кухаркин Е. С. Электрофизика информационных систем. Учебник для вузов. — М.: Высшая школа, 2001.
4. Щука А. А. Функциональная электроника. — М.: МИРЭЛ, 1998.



ЗАКЛЮЧЕНИЕ

Когда все уже написано, к автору приходят свежие мысли:

- ☐ Что получилось?
- ☐ Чего не хватает?
- ☐ Что не так сказано?
- ☐ Что неясно выражено?
- ☐ Названы ли вещи своими именами? И т. д. и т. п.

Видимо это неизбежно. Пока шла работа над книгой, тысячи исследователей и инженеров плодотворно работали в электронике.

Вспомним. Вначале был электрон. Свободный электрон в пустоте. Его приучили работать. Он стал работать в "пустотном реле" — электронной лампе. Родилась *вакуумная электроника*, ныне здравствующая.

Выяснилось, что электроны формируют ионы. Их тоже привлекли на службу. Возникла *плазменная электроника*, которая сегодня обретает новое звучание.

Свободные электроны обнаружили и в твердом теле. И не только электроны, но и покинутые ими места — дырки. Они стали "работать" на *твердотельную электронику*. С появлением "его величества" транзистора развилась *полупроводниковая электроника*, а затем ярко расцвела *микроэлектроника*. Сегодня микроэлектроника важнейшая из всех направлений электроники. В перспективе ожидается развитие *нанoeлектроники* на базе возникшей в последние годы нанотехнологии.

Стало ясно, что и ансамбли связанных электронов обладают уникальными способностями. Родилась *квантовая электроника*. Ныне квантовая электроника процветает. Каких только лазеров нет, где только их не встретишь!

А может быть электроника без электронов?

И да, и нет!

Нет! Потому, что электроны зримо и незримо присутствуют во всех процессах, связанных с динамическими неоднородностями, которые являются носителями информационного сигнала в *функциональной электронике*.

Да! Потому, что наступает эра *солитонов*. Солитон стал первой широко известной динамической неоднородностью и лик его ныне многообразен.

До шестидесятых годов прошлого столетия солитоном называли уединенную волну — волновой пакет неизменной формы, распространяющуюся с постоянной скоростью в жидкости. Математики многих поколений всесторонне описали это явление природы. Полученные решения показали на существование в природе различных видов солитонов в различных средах.

Физики частично эти решения овестьили. В каждой среде у солитона свое лицо:

- ☐ в кристаллической решетке — это дислокации;
- ☐ в плазме — это ионно-звуковые солитоны;

- в ферромагнетиках — это доменные стенки;
- в оптических волокнах — это оптические солитоны;
- в органических полупроводниках — это волны зарядовой плотности;
- в полимерах — это топологические солитоны;
- в сверхпроводниках — это вихри магнитного потока и т. д.

Солитоны существуют не только в электронике в ее нынешних границах.

Созданы солитонные модели элементарных частиц. В четырехмерном пространстве времени найдены динамические неоднородности солитонного происхождения — *инстан-тоны*...

Именно поэтому *функциональная электроника* имеет право на жизнь, на дальнейшее развитие.

Может быть со временем это направление в электронике получит название *солитонная электроника*.

А может быть просто — *солитоника*... Электроника всегда на марше!

ПРИЛОЖЕНИЕ

Вехи развития отечественной электроники

Российский период

Начало электроники в России по существу было положено в конце XIX века работами профессора Столетова А. Г. по фотоэффекту и профессора Попова А. С. в области радиосвязи.

1888

Профессор Столетов А. Г. открыл законы внешнего фотоэлектрического эффекта и изобрел прибор, преобразующий свет в электричество. Ныне этот прибор называется фотоэлементом (рис. П.1).



Рис. П.1. Александр Григорьевич Столетов —
русский физик,
профессор Московского университета



Рис. П.2. Александр Степанович Попов —
изобретатель радио, русский физик
и электротехник, профессор Петербургского
электротехнического института

1895

25 апреля (7 мая) Александр Степанович Попов выступил на заседании физического отделения Русского физико-химического общества с докладом и демонстрацией системы передачи и приема радиосигналов (рис. П.2). Устройство позволяло передавать информацию в системе азбуки Морзе на расстояние более 60 метров.

1896

В журнале "Журнал РФХО", который имел международную рассылку, увидела свет публикация профессора Попова А. С. "Прибор для обнаружения и регистрации электрических колебаний". Это была первая в мире публикация по передаче информации с помощью электромагнитных колебаний.



Рис. П.3. Петр Николаевич Лебедев — русский физик-экспериментатор, профессор Московского университета, создатель первой физической школы в России

физик Лебедев П. Н. предложил способ генерирования электромагнитных колебаний длинами волн 6—10 мм с помощью искрового генератора (рис. П.3). Ученый доказал, что именно электрический метод генерирования волн даст выгодное соотношение между мощностью излучения и его частотой.

1897

Профессор Попов А. С. достиг дальности радиосвязи в несколько километров. Он и его ассистенты при проведении экспериментов на транспорте "Европа" и крейсере "Африка" впервые обнаружили эффект отражения и интерференции радиоволн от крейсера "Лейтенант Ильин", проходящего между двумя судами.

1898

Фирма Дюкре (Франция) приступила к серийному производству радиостанций по системе профессора Попова А. С.

1899

Рыбкин Н. Д. и Троицкий Д. С. под руководством Попова А. С. осуществили прием радиосигналов на головные телефоны — "телефонный приемник депеш". Началась эра радиоэлектронной связи.

1900

Профессор Попов А. С. демонстрирует на Всемирной выставке в Париже свою связную аппаратуру и получает золотую медаль и диплом.

В одном из корпусов электромеханического завода Кронштадтского военного порта при участии профессора Попова А. С. организована мастерская для ремонта и производства приборов беспроводной связи. Так возникла российская радиопромышленность. Кронштадтская мастерская переросла в радиотелеграфное депо (1910 г.), преобразованная в Радиотелеграфный завод Морского ведомства (1915 г.)

Профессор Попов А. С. достиг дальности радиосвязи в 47 километров при проведении спасательных работ на Балтийском море.

1901

В Кронштадтской мастерской изготовлено 12 корабельных радиостанций системы профессора Попова А. С.

1903

Профессор Попов А. С. и его аспирант Лифшиц С. Я. передали звуки голоса по радио, открыв эру радиотелефонии.

1904

Организовано "Отделение для беспроволочной телеграфии по системе профессора Полова А. С. на электротехническом заводе "Сименс и Гальске" в Петербурге.

1907

Физик Розинг Б. Л. предложил электронную развертку изображения на основе осциллографической трубки и безинерционного фотоэлемента в передающем устройстве (рис. П.4). Он создал более 120 схем и систем телевизионных устройств.



Рис. П.4. Борис Львович Розинг — русский, советский физик и изобретатель, основоположник электронного телевидения, профессор Ленинградского технологического института и Архангельского лесотехнического института

1908

В Санкт-Петербурге организовано "Общество беспроволочных телеграфов и телефонов", которое с 1910 года стало называться "Русское общество беспроволочных телеграфов и телефонов" или РОБТиТ.

1909

Электротехник Коваленков В. И. сконструировал "микрофонно-телефонно-ионный усилитель", в котором использовал положительную обратную связь. Это было в Петербургском электротехническом институте.

1910

Создано первое научно-исследовательское учреждение в области радиоэлектроники "Поверочное отделение радиотелеграфных мастерских морского ведомства". Возглавил его Коринфинский Е. Л.

Электротехник Коваленков В. И. впервые разработал трехэлектродную лампу.

1911

Кронштадтская мастерская была переведена в Петербург, существенно расширилась и получила название "Радиотелеграфное депо морского ведомства". В ней была организована исследовательская лаборатория, которой последовательно руководили Петровский А. А., Исаков Л. Д., Шулейкин М. В.

Электротехник Коваленков В. И. изготовил двухсточную лампу с высокой крутизной характеристики для схем телефонной трансляции.

Изобретатель Розинг Б. Л. впервые в мире осуществил передачу изображения с помощью электронно-лучевой трубки.

1913

Инженер Федорицкий Н. А. основал в Петрограде мастерскую рентгеновских трубок, которая вскоре переросла в "Первый русский завод рентгеновских трубок".

Физик Иоффе А. Ф. провел прямые экспериментальные доказательства квантовой природы фотоэффекта (рис. П.5). Работа была отмечена премией Петербургской Академии наук.



Рис. П.5. Абрам Федорович Иоффе — русский, советский физик, профессор Петербургского политехнического института, основатель Ленинградского физико-технического института АН СССР, создатель советской школы физиков, академик АН СССР



Рис. П.6. Николай Дмитриевич Папалекси — русский, советский физик, профессор Одесского политехнического института, один из создателей Центральной радиолaborатории, академик АН СССР

1914

Ренгартен И. И. построил первый отечественный радиопеленгатор.

Профессор Папалекси Н. Д. на заводе Федорицкого Н. А. создал первые отечественные генераторные лампы мощностью до 100 Вт (рис. П.6).

1915

"Радиотелеграфное депо" преобразовано в завод, представляющий собой первое радиоэлектронное предприятие России с численностью около 300 человек. Завод специализировался на выпуске радиостанций мощностью 10 и 25 кВт, выпуске радиоприемников и радиопеленгаторов.

Первая в России радиотелефонная связь была осуществлена между Петербургом и Царским Селом. В устройствах передачи и приема информации использовались лампы профессора Папалекси Н. Д.

Началось производство электронных ламп на Радиозаводе морского ведомства под руководством Волынкина В. И.

Военные радиоспециалисты Бонч-Бруевич М. А. и Лещинский В. М. наладили производство отечественных газонаполненных приемоусилительных радиолamp в вакуумной мастерской на Тверской военной радиостанции (рис. П.7).

1916

Поручик Бонч-Бруевич М. А. организовал выпуск вакуумных усилительных ламп в Тверской радиостанции.

Профессор Папалекси Н. Д. впервые применил высокочастотный нагрев металлов в вакууме для их обезгаживания.



Рис. П.7. Михаил Александрович Бонч-Бруевич — один из пионеров радиотехники, создатель производства первых отечественных радиоламп, руководитель Нижегородской радиолaborатории, профессор Нижегородского университета, МВТУ, Ленинградского института инженеров связи (ныне носит его имя), член-корреспондент АН СССР

Радиоинженер Шулейкин М. В. на линкоре "Андрей Первозванный" установил радиотелеграфную связь между Петроградом и Гельсингфорсом на незатухающих высокочастотных колебаниях (20 кГц). В качестве генератора колебаний использовалась машина Волгодина В. П.

1917

Военные радиостанции передали историческое обращение "К гражданам России".

Советский период

Бурное развитие электроники произошло в годы становления и расцвета советской науки и техники. Уровень развития исследований и разработок был, как правило, на уровне или выше мирового по всему фронту разработок.

1918

19 июля Ленин В. И. подписал первый декрет о радио "О централизации радиотехнического дела Советской республики", положивший начало советской радиопромышленности. Создана Нижегородская радиолaborатория под руководством Бонч-Бруевича М. А.

1919

Радиоинженер Бонч-Бруевич М. А. впервые в мировой практике применил платиновые и красномедные аноды с водяным охлаждением и освоил выпуск генераторных ламп мощностью до 100 кВт.

Инженер Дикарев А. В. в Казанской радиолaborатории с помощью 9-лампового усилителя собственной конструкции осуществил радиоприем из Москвы.

1920

Состоявшийся в декабре VIII Всероссийский съезд Советов одобрил план ГОЭРЛО, в который входили вопросы радиофикации России.

Создана секция "Электросвязь", куда вошли петроградские заводы "Сименс и Гальске", "Эриксон", "Гейслер" и завод пустотных аппаратов (бывший завод Федорицкого), московские "Объединенный завод РАДИО", завод "Морзе", мастерская по ремонту приборов слабого тока, а также нижегородский телефонный завод "Сименс". Секция переехала из Петрограда в Москву.

Инженер Термен Л. С. изобрел электронный музыкальный инструмент "Терменвокс".

1921

Инженер Зилитинкевич С. И. впервые получил собственные колебания электронов в тормозящем поле анода и генерировал короткие монохроматические радиоволны длиной волны 30 см.

Инженер Дикарев А. В. сконструировал 12-ламповый усилитель с громкоговорителем, позволяющим воспринимать речь на расстоянии до 500 метров. По указанию Ленина громкоговорители установили на шести площадях Москвы.

1922

9 марта создан Государственный электротехнический трест заводов слабого тока (ГЭТЗСТ), в который вошли 11 заводов Москвы, Петрограда, Нижнего Новгорода, а также два строящихся радиообъекта на Шаболовке в Москве и Детском Селе под Петроградом.

Инженер Рчеулов изобрел магнитную запись видеосигналов, а также электровакуумную передающую трубку — оптический диссектор.

Профессор Богословский М. М. наладил в Петроградском политехническом институте серийный выпуск "ламп Богословского" с вольфрамовым катодом. Приемоусилительные лампы маркировались литерами R и M, а генераторные — G.

В том же институте в лаборатории Чернышева А. А. также выпускались генераторные и усилительные лампы (рис. П.8).

Рис. П.8. Александр Алексеевич Чернышев — русский, советский электротехник и радиотехник, один из создателей электронной промышленности, академик АН СССР



Под руководством Бонч-Бруевича М. А. вступила в строй первая радиовещательная станция имени Коминтерна мощностью 12 кВт. Станция была первой в Европе и самой мощной в мире.

Лосев О. В., изучая свойства кристаллического детектора, обнаружил у кристалла падающий участок вольтамперной характеристики. Он впервые построил генерирующий детектор, т. е. детекторный приемник, способный усиливать электромагнитные колебания. В своем приборе Лосев использовал контактную пару "металлическое острие—кристалл цинкита". На эту контактную пару подавалось небольшое напряжение. Прибор Лосева вошел в историю полупроводниковой электроники как *кристадин*.

923

Налажен выпуск приемоусилительных ламп с вольфрамовым катодом в Одессе. Производство возглавляли Папалекси Н. Д. и Мандельштам Л. И. (рис. П.9), работали Тамм И. Е., Романюк К. Б., Стохарский К. В., Щеголев Е. Я.

В лаборатории Вологодина В. П. разработаны и производились первые ртутные мощные выпрямители, предназначенные для питания генераторных ламп радиотелефонных передатчиков.



Рис. П.9. Леонид Исаакович Мандельштам — русский, советский физик, профессор Московского университета, академик АН СССР



Рис. П.10. Сергей Аркадьевич Векшинский — советский ученый в области электровакуумной техники. Один из основателей отечественной электронной промышленности, академик АН СССР

1924

Бонч-Бруевичем М. А. разработана и выпущена первая в мире генераторная лампа с водяным охлаждением мощностью 100 кВт.

На Электровакуумном заводе налажен выпуск серии из шести типов генераторных ламп с мощностью до 1000 Вт, пяти типов выпрямительных ламп, ртутных выпрямителей, электронных рентгеновских трубок. Техническим директором завода был Богословский М. М., а главным инженером — Векшинский С. А. (рис. П.10).

Профессора Харьковского университета Слуцкий А. А. и Штейнберг Д. С. предложили магнетронный способ генерации электромагнитных колебаний.

1925

На территории Электровакуумного завода была организована Центральная радиолaborатория (ЦРЛ), в которую вошли лучшие специалисты тех времен (Вологдин В. П., Рожанский Д. А., Мандельштам Л. И., Папалекси Н. Д., Шорин А. Ф., Бонч-Бруевич М. А., Щеголев Е. Я. и др.). Значительный вклад в развитие вакуумной электроники внес советский физик Дмитрий Апполинариевич Рожанский (рис. П.11). В состав ЦРЛ была включена и вакуумная лаборатория, которой руководил Шапошников А. А., и где работали Волюнкин В. И., Зусмановский С. А., Иванов А. А., Астафьев В. А.

1926

Шапошниковым А. А. изготовлены первые советские лампы с оксидным катодом.

Инженер Романюк К. Б. изготовил рентгеновскую трубку с автокатодом, разработал карбидированный вольфрамовый катод.

Инженер Оболенский С. А. предложил катодорезный метод покрытия оксидных катодов. Физик Френкель Я. И. высказал гипотезу о том, что дефекты кристаллической структуры представляют собой "пустое место" (или "дырку"), которое способно перемещаться по кристаллу.



Рис. П.11. Дмитрий Апполинариевич Рожанский — русский, советский физик, профессор Ленинградского политехнического института. Автор работ и изобретений в области вакуумной электроники, радиофизики и радиолокации



Рис. П.12. Петр Иванович Лукирский — русский, советский физик-экспериментатор, один из основоположников отечественной школы физической электроники, создатель отечественной школы физиков-экспериментаторов, профессор Ленинградского политехнического института, академик АН СССР

1927

На Электровакуумном заводе (завод бывшего Русского общества беспроводного телеграфа и телефона — РОБТиТ) под руководством Векшинского С. А. был налажен выпуск осциллографических трубок с горячим катодом и с экраном трех цветов (синим, желтым и зеленым).

Инженер Хараджа Ф. Н. на том же заводе начал серийный выпуск рентгеновских трубок. Лосев О. В. открыл явление свечения кристаллов карборунда при прохождении тока через точечный контакт (свечение Лосева). Он дал правильное физическое истолкование открытому им явлению свечения кристаллов, доказав экспериментально существование некоторого "активного слоя" в детектирующем контакте. В терминах сегодняшней электроники это явление называют рекомбинационным излучением $p-n$ -переходов.

1928

Электровакуумный завод объединился с заводом "Светлана", который до этого производил электрические лампы накаливания. В рамках объединения главный инженер Векшинский С. А. создал исследовательскую лабораторию на базе лаборатории Шапошников-ва А. А. В ней работала группа физиков (Шальников А. А., Ансельм А. И., Рыжанов С. Г., Лукирский П. И. (рис. П.12), Гринберг Г. А. и др.), инженеры электровакуумной промышленности (Романюк К. Б., Александров А. Г., Мошкович С. М. и др.), а также молодые инженеры (Бабат Г. И., Болдырь Ю. Д., Красилов А. В., Подгурский Е. Л., Полевой К. П., Юноша Ю. А. и др.).

На Московском электрозаводе была создана лаборатория электронных ламп, которая освоила выпуск преумножительных ламп типов Р-5, ПТ-2, использующих торированный катод. Выпускались также генераторные лампы Г-6 мощностью 50 Вт, Б-250 (250 Вт), БТ-500 (500 Вт).

1929

Под руководством Минца А. Л. была построена и введена в строй крупнейшая радиовещательная станция имени ВЦСПС мощностью 100 кВт (рис. П.13). На этой радиостанции

впервые использовался многокаскадный передатчик на 18 параллельно соединенных 15 кВт лампах с кварцевой стабилизацией частоты.



Рис П.13. Александр Львович Минц — советский ученый в области радиоэлектронной техники, основатель Радиотехнического института АН СССР, академик АН СССР

1930

На заводе "Светлана" был создан первый тетрод типа СО-44 с оксидным катодом, а затем разработаны тетроды с торированным катодом СТ-80 и катодом косвенного накала СО-95 с внутренней экранировкой от внешних электростатических полей.

Инженер Кубецкий Л. А. создал первый фотоумножитель с фокусировкой электронов с помощью магнитного поля.

Физик Константинов А. П. разработал электронно-лучевую трубку для передачи изображения, в которой использовался принцип накопления зарядов.

1931

На заводе "Светлана" Мошковым С. М. была создана лампа-варимю СО-148 и выходной пентод СО-122 на основе технологии изготовления барированных катодов.

Инженер Катаев С. И. разработал устройство передачи движущихся изображений, которое, по сути, было передающей электронно-лучевой трубкой с накоплением заряда — аналог иконоскопа.

Введена в строй первая передающая телевизионная станция.

Инженер Катаев С. И. изготовил передающую телевизионную трубку — иконоскоп на основе одностороннего мозаичного фотокатода.

Физики Кикоин И. К. и Носков М. М. открыли фотомагнетозлектрический эффект. Эффект заключался в возникновении разности потенциалов на гранях освещенной пластинки закали меди, находящейся в магнитном поле. Эффект Кикоина — Носкова нашел широкое применение при изучении свойств полупроводников.

Физики Иоффе А. Ф. и Френкель Я. И. предложили туннельную теорию в полупроводниках.

Инженеры Шмаков П. В. и Тимофеев П. В. изобрели передающую телевизионную трубку — супериконоскоп.

1932

Инженер Остряков П. А. создал первые мощные лампы с воздушным охлаждением.

Профессор Рожанский Д. А. открыл принцип скоростной модуляции группирования электронов.

На базе бывшей шелкоткацкой фабрики в Щелковском районе Московской области был создан завод "Радиолампа", на котором было развернуто производство радиоусилительных ламп с бариевыми и оксидными катодами, а также генераторных ламп с карбидными катодами.

1933

Инженерами Минцем А. Л. и Огановым Н. И. были созданы первые разборные мощные генераторные лампы.

Инженер Полевой К. П. разработал первые приемные телевизионные трубки — кинескопы.

Профессора Тимофеев П. В. и Шмаков П. В. разработали передающую телевизионную трубку, в которой осуществлялся перенос изображения со сплошного фотокатода на сплошную диэлектрическую мишень. Она называлась "суперэмитрон" и представляла собой современный супериконоскоп.

Вступила в строй спроектированная Минцем А. Л. 500 кВт радиостанция им. Коминтерна.

Физик Ландау Л. Д. (рис. П.14) впервые указал на существование в инерционной поляризуемой среде особого квантового стационарного состояния электрона — полярона.



Рис. П.14. Лев Давыдович Ландау — советский физик-теоретик, академик АН СССР

1934

Заводская лаборатория "Светланы" была реорганизована в Отраслевую вакуумную лабораторию (ОВЛ), ставшую одним из центров исследований советской электроники. ОВЛ до 1937 года руководил Векшинский С. А., а затем до 1941 года — Зусмановский С. А.

Инженер Коровин Ю. К. в ЦРЛ в Ленинграде экспериментально доказал возможность радиообнаружения самолета. На установке, работающей на магнетроне с длиной волны излучения 20 см и мощностью несколько ватт, удалось детектировать самолет на расстоянии до 3 км.

Инженер Шембель Б. К. разработал радиолокационную станцию непрерывного излучения.

Инженер Мошкович С. М. создал "суперную серию" ламп, включающую смесительные и комбинированные лампы (СО-182, СО-184, СО-185, СО-187, СО-193).

Инженер Кубецкий Л. А. разработал фотоумножитель, известный под названием "трубка Кубецкого".

1935

В лаборатории профессора Чернышева А. А. в Ленинградском электрофизическом институте создана радиолокационная установка на магнетроне с длиной волны излучения 21 см и мощностью 20 Вт, которая позволяла обнаруживать самолет на расстоянии до 8 км.

ЛЭФИ преобразуется в НИИ-9 и ориентируется на разработку оборонной тематики. Институт возглавил Бонч-Бруевич А. М. Была разработана радиолокационная станция "Буря" с дальностью обнаружения самолетов до 11 км.

Инженер Ощепков П. К. в ЛФТИ разработал метод импульсной радиолокации с электронно-лучевой индикацией целей.

Круссер Б. В. и Романова Н. М. создали иконоскоп, с помощью которого Рыфтиным Я. А. была продемонстрирована система электронного телевидения с разрешением на 180 строк.

1936

На заводе "Радиолампа" налажен выпуск металлических прямоусилительных ламп 6С5, 6Ф5, 6К7, 5Ц4 и др.

1937

Инженеры Алксеев Н. Ф. и Маляров Д. Е. разработали многорезонаторные магнетроны на длине волны излучения 9 см, мощностью до 300 Вт и КПД 20%.

Физики Тамм И. Е. (рис. П.15) и Франк И. М. построили теорию излучения электрона, движущегося в среде со скоростью, превышающей фазовую скорость в этой среде — эффект Вавилова — Черенкова. За эту работу в 1958 году им была присуждена Нобелевская премия.

В ЛФТИ Кобзарев Ю. Б., Погорелко П. А., Чернецов Н. Я. создали радиолокационную установку "Модель-2" с дальностью обнаружения самолетов свыше 100 км.



Рис. П.15. Игорь Евгеньевич Тамм — советский физик-теоретик, академик АН СССР, лауреат Нобелевской премии

1938

Инженер Круссер Б. В. с группой сотрудников на заводе "Светлана" выпустил первые суперортиконы с разрешением до 900 линий.

Группа инженера Девяткова Н. Д. разработала триоды и триодные генераторы на дециметровый диапазон длин волн.

Инженер Брауде Г. В. создал высокочувствительный ортикон.

В Ленинграде начались регулярные телепередачи.

ЛФТИ совместно с НИИИСКА разработали подвижную станцию дальнего радиообнаружения "Редут", в которой использовались мощные генераторные лампы ИГ-8 на длине

одны излучения 4 м и 50 кВт в импульсе. Максимальная дальность обнаружения самолетов составила 95 км на высоте 7500 м.

Физик Давыдов Б. И. теоретически доказал, что выпрямление тока в контакте двух полупроводников связано с наличием электронной и дырочной проводимости по разные стороны контакта.

939

Инженерами Девятковым Н. Д., Слиозбергом М. Л., Данильцевым Е. Н. разработаны генераторные и усилительные двухконтурные клистроны.

Ученые Векшинский С. А., Авдеев В. Н., Вильдгрубе Г. С. создали фотоумножитель жабузийного типа.

В Москве в январе начались регулярные телепередачи.

ИФТИ совместно с НИИИСКА разработали подвижную станцию дальнего обнаружения "Редут" на основе отечественных ламп ИГ-4 с мощностью 50 кВт в импульсе, которая в следующем году была принята на вооружение армии.

Туннельный эффект, в соответствии с которым частицы с заданным потенциалом способны с определенной вероятностью проникать через высокопотенциальный барьер, был впервые предсказан физиком Гамовым Г. Г. Диод на его основе был создан в 1958 году Эсахи Л.

Инженер Девятков Н. Д. (рис. П.16) создал первый прямопролетный клистрон мощностью 100 Вт на длине волны излучения 15 см.

Инженеры Девятков Н. Д. и Пискунов И. В. и независимо от них Коваленко В. Ф. предложили конструкции отражательных клистронов.

Инженеры Алексеев Н. Ф. и Маляров Д. Е. создали многокамерный магнетрон сантиметрового диапазона длин волн.



Рис. П.16. Николай Дмитриевич Девятков — советский ученый в области электроники, медицинской электроники, академик АН СССР



Рис. П.17. Валентин Александрович Фабрикант — советский физик

1940

Физик Фабрикант В. А. (рис. П.17) впервые указал на возможность использования индуцированного излучения для наблюдения отрицательного поглощения (усиления) света. Именно на основе этого явления были созданы лазеры.

1941

Инженеры Мошкович С. М. и Александров А. Г. создали первые импульсные генераторные лампы.

Физик Лашкарев В. Е. экспериментально подтвердил, что в медно-закисном купроксном выпрямителе осуществлялся контакт полупроводников различного типа.

1942

Разработана и принята на вооружение радиолокационная станция орудийной наводки (СОН). С этой целью был разработан новый комплект электровакуумных приборов.

1943

Разработана самолетная радиолокационная станция "Гнейс-2", работавшая на длине волны 1,5 метра и излучаемой мощностью в импульсе до 10 кВт. Станция была принята на вооружение Советской армии.

Постановлением Государственного Комитета Обороны № ГОКО-3683сс создан Совет по радиолокации при ГКО. Научно-техническую часть Совета возглавили известные специалисты Берг А. И., Шокин А. И., Калмыков В. Д., Кобзарев Ю. Б., Угер Г. А., Щукин А. Н.

1944

Физик Вул Б. М. открыл и исследовал сегнетоэлектрические свойства титаната бария и других активных диэлектриков.

1945

Постановлением Правительства СССР 7 мая учрежден ежегодный День радио.

Физик Завойский Е. К. открыл явление парамагнитного резонанса или электронный спиновой резонанс.

1946

Созданы модуляторные тетроды ГМИ-83.

Физик Лашкарев В. Е. открыл биполярную диффузию неравновесных носителей тока в полупроводниках.

1947

Совет по радиолокации преобразован в Комитет по радиолокации.

Разработан первый отечественный кинескоп 18ЛК15, серийное производство которого началось на Московском электроламповом заводе.

Профессором Котельниковым В. А. создана теория потенциальной помехоустойчивости.

Профессором Кабановым Н. И. сделано открытие эффекта дальнего рассеянного отражения от Земли дециметровых радиоволн.

1948

Инженером Стельмахом М. Ф. созданы первые образцы ламп обратной волны (ЛОВ).

Промышленностью стали выпускаться кинескопы 30ЛК1Б с экраном диаметром 3000 мм и углом отклонения 60 градусов.

Инженер Красилов А. В. и его группа разработали германиевые диоды для радиолокационных станций.

Физик Лашкарев В. Е. построил общую теорию фото-ЭДС в полупроводниках.

1949

В подмосковном Фрязино в НИИ-160 (НИИ "Исток") Красиловым А. В. (рис. П.18) и Мадояном С. Г. впервые наблюдался транзисторный эффект. Они опубликовали первую в СССР статью о транзисторах под названием "Кристаллический триод".



Рис. П.18. Александр Викторович Красилов — советский ученый, создатель первого отечественного транзистора



Рис. П.19. Сергей Алексеевич Лебедев — советский ученый в области электротехники и вычислительной техники, создатель первых отечественных ЭВМ на электронных лампах, директор Института точной механики и вычислительной техники, академик АН СССР

1950

Разработана серия виброустойчивых импульсных ламп (ГМИ-5, ГМИ-6, ГМИ-7) на анодных напряжениях от 3 до 40 кВ и импульсные токи до 150 А.

Под руководством Лебедева С. А. разработана первая электронная цифровая вычислительная машина на отечественных электронных лампах (рис. П.19).

Академик Иоффе А. Ф. создал теорию термоэлектрических преобразователей и предложил использовать термоэлектрические свойства полупроводников для создания охлаждающих устройств. Его сотрудник Стильбанс Л. С. провел необходимые расчеты, разработал конструкцию и создал первый в мире полупроводниковый холодильник, основанный на использовании эффекта Пельтье.

Созданы макеты видиконов.

Лабораторные образцы германиевых транзисторов были разработаны Вулом Б. М. (рис. П.20), Жановым А. В., Вавиловым В. С. (ФИАН), Тучкевичем В. М., Наследовым Д. Н. (ЛФТИ), Жалашниковым С. Г., Лениным Н. А. (ИРЭ АН СССР) и др.

Физики Гинзбург В. Л. и Ландау Л. Д. построили общую теорию сверхпроводников.

1951

Физики Фабрикант В. А., Вудынский М. М., Бутаева Ф. А. открыли способ квантового усиления электромагнитного излучения при помощи индуцированного излучения, который лег в основу работы всех типов лазеров. Заявка была зарегистрирована 18 июня 1951 года, опубликована в 1959 году, при этом диплом на открытие был выдан только в 1964 году.



Рис. П.20. Бенцион Моисеевич Вул — советский физик, академик АН СССР

1952

Под руководством Понаморева Е. М. разработан суперорбитон ЛИ-17 на основе висмутово-серебряного фотокатода для передвижных телевизионных станций.

В НИИ-108 Пениным Н. А., Кубецким Г. А., Якуниной К. В., Пантелеймоновой Е. А. создан германиевый плоскостной сплавной транзистор.

Создатель отечественного транзистора Красильов А. В. и инженер Щиголь Ф. А. разработали промышленные образцы точечных транзисторов типа С1 и С2.

Физик Абрикосов А. А. выдвинул идею о существовании сверхпроводников второго рода и ввел представление вихревого состояния токов — "вихрей Абрикосова".

1953

Физик Иоффе А. Ф. на основе исследований термоэлектрических свойств полупроводников создал серию термоэлектродгенераторов.

В ЦНИИ-35 изготовлены плоскостные транзисторы типов П1, П2, П3.

Одним из первых, кто поставил проблему преобразования световой энергии в электрическую с помощью полупроводников, стал Иоффе А. Ф. Основоположающими в области фотоэлектрических свойств полупроводников стали работы советских ученых Давыдова Б. И., Курчатова И. В., Кушнера Ю. М., Ландау Л. Д., Лашкарева В. Е., Тучкевича В. М., Алферова Ж. И. и др. Был создан германиевый фотоэлемент, работавший в диодном режиме и управлявшийся светом по обратному току.

1954

Физики Прохоров А. М. (рис. П.21) и Басов Н. Г. (рис. П.22) создали микроволновой квантовый генератор — лазер.

Разработаны первые промышленные образцы германиевых транзисторов.

Слушатель академии им. Жуковского капитан Федотов Я. А. впервые применил один из плоскостных транзисторов в макете супергетеродинного радиоприемника.

1955

В лаборатории профессора Калашникова С. Г. был создан германиевый транзистор, работающий в частотном диапазоне 1,0—1,5 МГц.

1956

Разработан трехлучевой металлоглазанный кинескоп 53ЛК41Ц с трехцветным экраном и теневой маской.



Рис. П.21. Александр Михайлович Прохоров — основоположник квантовой электроники, директор созданного им ИОФАНа, академик АН СССР



Рис. П.22. Николай Геннадьевич Басов — основоположник квантовой электроники, директор ФАН АН СССР, академик АН СССР

Инженером Галинским Н. Д. разработан высокочувствительный суперортикон ЛИ-201, который широко использовался в студийной аппаратуре и получил гран-при на выставке в Брюсселе.

Инженер Щиголь Ф. А. разработал первые кремниевые сплавные транзисторы типа П501-П503.

1957

На орбиту выведен первый в истории человечества искусственный спутник Земли на основе отечественной электронной элементной базы.

Физик Келдыш Л. В. построил систематическую теорию туннельных явлений в полупроводниках.

В НИИ "Пульсар" были разработаны кремниевые сплавные транзисторы типов П101-П106.

В СССР выпущено 2,7 миллиона транзисторов.

1958

В СССР начато серийное производство фотодиодов.

Вышло Постановление СМ СССР о строительстве под Москвой города-спутника (ныне Зеленоград).

Гамм И. Е. совместно с Франком И. М. и Черенковым П. А. удостоен Нобелевской премии по физике за открытие и объяснение "эффекта Черенкова" — эффекта сверхсветового люминесценции.

Выпущены первые германиевые транзисторы П605-П609, работающие на частотах до 100 МГц с мощностью до 10 Вт, изготовленные по конверсионной технологии, предложенной Стружинским В. А.

Физик Прохоров А. М. впервые предложил использовать резонатор открытого типа в квантовых усилителях и генераторах.

1959

Физикам Басову Н. Г. и Прохорову А. М. присуждена Ленинская премия за разработку нового принципа генерации и усиления электромагнитных колебаний, создание молекулярных генераторов и усилителей.

Физики Басов Н. Г., Вул Б. М. и Попов Ю. М. предложили использовать полупроводники для создания лазеров.

Физики Тагер А. С., Мельников А. И., Цебиев А. М., Кобельков Г. П. сделали открытие явления генерации и усиления СВЧ-колебаний при лавинном пробое полупроводниковых диодов.

Идея академика Вула Б. М. о применении электронно-дырочного перехода в качестве переменной емкости позволила создать варикап типа Д901, а также умножительные диоды.

1960

Созданы образцы микромодульной техники на основе миниатюрных элементов.

Главным конструктором Носовым Ю. Р. (рис. П.23) в НИИ "Сапфир" разработаны импульсные диоды для ЭВМ.



Рис. П.23. Юрий Романович Носов — советский, русский ученый, инженер



Рис. П.24. Владислав Иванович Пустовойт — советский, русский физик, член корреспондент АН СССР

1961

Запущен в космический полет Гагарин Ю. А. Вся электроника для систем связи и управления изготовлена на основе отечественной элементной базы.

Шокин А. И. назначен Председателем Государственного Комитета Совета Министров СССР по электронной технике.

Физики Пустовойт В. И. (рис. П.24) и Герценштейн М. Е. открыли эффект усиления ультразвука в полупроводниковых пьезопроводящих кристаллах.

Физики Раутиан С. Г. и Собельман И. И. предложили использовать органические комплексные соединения в качестве генерирующих сред для лазеров.

1962

Под руководством академика Капицы П. Л. создан оригинальный тип СВЧ-генератора магнетронного типа — нигатрон (рис. П.25).

Под руководством Горфинкеля Б. И. в Саратове разработана серия ламп для новых моделей унифицированных телевизоров (1Ц21П, 6Д20Г1, 6П36С, 6Ф5П, 6К13П, 6П124П) со сроком долговечности 10 000 часов.

Вышло Постановление ЦК КПСС и СМ СССР о создании Научного Центра микроэлектроники в Зеленограде.

Физик Денисюк Ю. Н. впервые предложил и осуществил запись голограмм в трехмерных средах.

Физик Келдыш Л. В. предложил использовать пространственно-периодические поля — сверхрешетки для управления электронным спектром и электронными свойствами кристалла.

Физики Хохлов Р. В. (рис. П.26) и Ахманов С. А. предложили новый тип генератора когерентных колебаний — параметрический генератор.

Создана первая отечественная интегральная схема на кремнии, в которой использовались кремниевые сплавные транзисторы.



Рис. П.25. Петр Леонидович Калица — советский физик, автор работ в различных областях физики, в том числе создал общую теорию электронных приборов магнетронного типа, создатель Института физических проблем АН СССР, академик АН СССР



Рис. П.26. Рем Викторович Хохлов — советский физик, академик АН СССР

1963

Создан Центр микроэлектроники в Зеленограде, первым директором которого стал Лукин Ф. В.

Инженер Щиголь Ф. А. разработал планарный транзистор 2Т312 и его бескорпусной аналог 2Т319, ставший основным активным элементом гибридных схем.

Физики Басов Н. Г. и Ораевский А. Н. предложили использовать быстропротекающие химические реакции для создания уровней инверсной заселенности и создания на их основе химических лазеров.

Физики Басов Н. Г., Вул Б. М., Попов Ю. М. с сотрудниками создали первый полупроводниковый лазер.

1964

Физикам, основателям квантовой электроники Басову Н. Г. и Прохорову А. М. была присуждена Нобелевская премия по физике за основополагающие работы по квантовой электронике.

На заводе "Ангстрем" при НИИ точной технологии созданы первые интегральные схемы "Тропа" с 20-ю элементами на кристалле, выполняющей функцию транзисторной логики с резистивными связями.

Физики Гуляев Ю. В. и Пустовойт В. И. открыли эффект акустоэлектронного взаимодействия поверхностных акустических волн и электронов в прилегающем полупроводнике и предложили использовать это явление в акустоэлектронных устройствах.

Физик Гуляев Ю. В (рис. П.27) открыл и исследовал акустомагнетоэлектрический эффект.



Рис. П.27. Юрий Васильевич Гуляев — советский, русский физик, академик АН СССР



Рис. П.28. Александр Иванович Шокин — крупный государственный деятель, первый министр электронной промышленности СССР и России

1965

Государственный Комитет преобразован в Министерство электронной промышленности СССР во главе с Министром Шокиным А. И. (рис. П.28).

В НИИМО в Зеленограде создана технология и начат выпуск первых планарных транзисторов "Плоскость".

Под руководством Малина Б. В. в НИИ-35 (ныне НИИ "Пульсар") была создана первая серия кремниевых интегральных схем ТС-100 (степень интеграции — 37 элементов на кристалле).

1966

Физики Прохоров А. М и Конюхов В. К. предложили идею создания газодинамического лазера на смеси углекислого газа и азота.

Учеными Тагером А. С. и Вальд-Перловым В. М. созданы первые лавинно-пролетные диоды.

В НИИ "Пульсар" начал работать первый экспериментальный цех по производству планарных интегральных схем.

В НИИМЭ под руководством доктора наук Валиева К. А. начат выпуск логических и линейных интегральных схем.

Физик Месяц Г. А. (рис. П.29) открыл явление взрывной электронной эмиссии и провел его фундаментальное исследование.

1967

Созданы первые транзисторы в пластмассовых корпусах типа КТ-315, которые стали поставляться в 27 стран мира.

Физик Гуляев Ю. В. с сотрудниками открыл акустомагнетоэлектрический эффект. Диплом на открытие был получен в 1974 году.



Рис. П.29. Геннадий Андреевич Месяц — советский, русский физик, академик АН СССР



Рис. П.30. Жорес Иванович Алферов — советский физик, академик АН СССР, лауреат Нобелевской премии

1968

Физик Алферов Ж. И. (рис. П.30) с сотрудниками предложил конструкции лазеров на основе гетеропереходов.

НИИ "Пульсар" выпустил партию первых гибридных тонкопленочных ИС с планарными бескорпусными транзисторами типов КД910, КД911, КТ318, предназначенных для телевидения, радиовещания и связи.

В НИИ МО разработаны цифровые и линейные ИС массового применения (серия 155).

1969

В НИИ "Пульсар" под руководством Бачурина В. В. созданы мощные высокочастотные МДП-транзисторы СВЧ-диапазона.

Физик Алферов Ж. И. сформулировал и практически реализовал свои идеи управления электронными и световыми потоками в классических гетероструктурах на основе системы GaAs-AlAs.

Физик Гуляев Ю. В. открыл акустоэлектронные волны "волны Гуляева — Блюстейна", не обладающие дисперсией и нашедшие широкое применение в современной акустоэлектронике.

1970

В стране было выпущено 3,6 миллионов интегральных схем 69-ти серий, из которых 7 серий по МОП-технологии, 32 серии по биполярной технологии.

В НИИМО разработаны функциональные приборы на полевых транзисторах с затвором Шоттки и приборами Ганна на одном кристалле.

1971

Введен в эксплуатацию завод цветных кинескопов "Хроматрон".

1972

Созданы генераторные лампы с магнитной фокусировкой электронного потока (ГК-12А, ГК-13А) и полезной мощностью до 100 кВт на частоте 3 МГц.

Разработаны генераторные тетроды (ГУ74Б, ГУ77Б, ГУ78Б) мощностью до 2,5 кВт.

Физики Алферов Ж. И., Андреев В. М., Гарбузов Д. З., Корольков В. И. и Трегьяков Д. Н. удостоены Ленинской премии за работы в области физики гетероструктур.

В НИИФП в Зеленограде изготовлены линейные ПЗС-приборы с поверхностным каналом.

Физик Абрикосов А. А. выдвинул идею о высокотемпературной сверхпроводимости "металлического экситония".

1973

В НИИ "Пульсар" Кузнецовым Ю. А., Вето А. В., Гольдшером А. И., Шилиным В. А., Скрылевым А. С. и др. создана серия приборов с зарядовой связью.

Созданы интегральные схемы для наручных часов со степенью интеграции 1500 транзисторов на кристалл размером 2×2 мм².

Под руководством Иванова Э. Е. на заводе "Ангстрем" за пять месяцев был разработан и выпущен калькулятор на основе собственных БИС.

1974

В Научном Центре на заводе "Ангстрем" созданы первые отечественные микропроцессоры.

Под руководством Гуськова Г. Я. создана ЦВМ "Салют-31", предназначенная для управления аппаратурой земных станций спутниковой связи "Сургут", размещаемой на подвижных объектах и обеспечивающих дуплексную телефонно-телеграфную связь.

1975

Организован промышленный выпуск цифровых ИС серий 100 и 500 с быстродействием 2 нс для супер-ЭВМ "Эльбрус-2".

На заводе "Ангстрем" создана БИС ЗУ динамического типа емкостью 4 Кбит.

1976

В НИИ "Пульсар" созданы КМОП БИС ЗУ с информационной емкостью 1 Кбит.

Электронная промышленность выпустила 300 миллионов интегральных схем, из них 85% монолитных (полупроводниковых).

На заводе "Ангстрем" была достигнута степень интеграции 20 000 транзисторов на кристалл.

В НПО "Зенит" под руководством Жильцова В. И. было разработано свыше 20 моделей лазеров на красителях с перестраиваемой частотой с точностью до 0,0001%.

1977

В НИИМЭ разработан комплект микропроцессорных ТТЛШ серии 589.

1978

В НИИ "Пульсар" разработаны мощные МДП-транзисторы (КП905, КП907, КП908), позволившие достичь верхней границы дециметрового диапазона частот.

Разработаны мощные полевые СВЧ-транзисторы на арсениде галлия с барьером Шоттки.

В НИИМЭ разработана ППЗУ типа ТТЛ-556Р15 для космических систем и ЭВМ "Эльбрус".

1979

НИИМЭ осуществил поставки ИС для обеспечения программы "Марс-Венера".

На заводе "Ангстрем" создана первая однокристалльная микро-ЭВМ, эквивалентная мини-ЭВМ.

1980

Заводом "Микрон" изготовлена 10 000 000 интегральная схема.

1981

Министр Шокин А. И., выступая на XXVI съезде КПСС, подчеркнул высокий уровень развития отечественной электроники.

1983

В НИИМЭ организован промышленный выпуск базовых матричных кристаллов БМК И-200 и БМК И-300 для отечественных ЭВМ.

1984

В НИИТТ был разработан первый персональный компьютер ДВК-1, а на заводе "Ангстрем" было налажено его серийное производство.

1985

В НИИМЭ получены тестовые образцы кристаллов ИС с топологической нормой 0,5 мкм с использованием электронно-лучевой литографии.

1986

Физик из МГУ Лихарев К. К. предсказал процесс одноэлектронного туннелирования и предложил приборы обработки информации на основе этого явления.

В НИИМЭ выпущены первые ИС с программируемыми логическими матрицами серий 556 и 1556.

1987

Вошел в строй Центр с использованием синхротронного излучения в Зеленограде.

НИИ ФП в Зеленограде совместно с ЦНИИС создали и ввели в эксплуатацию первый в стране участок волоконно-оптической связи.

1988

На заводе "Ангстрем" создан первый 32-разрядный микропроцессор.

1989

На заводе "Ангстрем" налажен выпуск СБИС памяти емкостью 1 Мбит.

1990

НПО "Научный центр" поставил для школ 300 000 вычислительных машин УК-НЦ. 200 000 ДВК были поставлены на предприятия.

В электронной промышленности насчитывалось 816 предприятий, организаций и филиалов, из них 232 научные организации с филиалами. Удельный вес отрасли в общем объеме промышленной продукции составлял 2,4%.

Новейший российский период

На постсоветском этапе развития России заметных успехов в области вакуумной электроники не наблюдается. Проводимые реформы деформировали электронную промышленность, ряд ведущих предприятий вакуумной электроники на грани закрытия, другие после так называемого акционирования утратили производственный профиль деятельности. Удельный вес отрасли снизился в 15 раз, эффективно работающие предприятия составляют всего несколько процентов.

1991

Обвал электронной промышленности СССР, обвал всех рынков сбыта предприятий "оборонки". Предприятия остались без финансового потенциала для конкурентной борьбы.

1992

НИИМП-ЭЛАС совместно с НПО имени Лавочкина разрабатывают системы спутниковой связи, предназначенной для оперативного обмена информацией в структурах банковской и финансовых систем.

ОАО "Ангстрем" развернул свои разработки в интересах Юго-Восточной Азии.

1993

В Зеленограде учреждено Российско-Китайское предприятие "Корона-Семикондактор" для производства СБИС с топологической нормой 0,8 мкм.

В НИИМЭ разработаны ИС ЭСЛ на 10 000 вентилях с быстродействием 0,15 нс и поставлены в Китай.

1994

Правительством РФ принята "Российская государственная программа развития электронной техники" на период 1995—2000 годы. Реальное финансирование составило 10% от предусмотренных бюджетных ассигнований.

1995

Российская электроника имела годовые объемы вложений 150 млн долларов, а мировой рынок оценивается в 210 млрд долларов.

В России только на заводах "Ангстрем" и "Микрон" в Зеленограде можно производить СБИС с топологической нормой 1,2 мкм.

1996

Разработана и принята целевая программа "Электронная промышленность Москвы — России XXI века".

Правительством создана холдинговая компания "Российская электроника", в которую вошли 32 предприятия и научно-исследовательских институтов бывшей электронной промышленности. Компанию возглавил Генеральный директор завода "Ангстрем" Дшхунян В. Л. (рис. П.31), заместителем стал Андреев А. С.

1997

На заводе "Микрон" введена производственная линия по выпуску СБИС с проектными нормами 0,8 мкм на пластинах 150 мм.



Рис. П.31. Валерий Леонидович Джхунян — конструктор изделий микроэлектроники, генеральный директор ОАО "Российская электроника"

3 НИИМО разработана элементная база БикМОП ИС на основе самосовмещенной технологии.

4 разработана и принята целевая программа "Электронная промышленность Москвы — России XXI века".

998

5 на СП "Корона" начато промышленное производство СБИС на пластинах кремния диаметром 150 мм с топологическими нормами 0,8 мкм.

6 указом Президента РФ и Постановлением Правительства создана холдинговая компания "Российская электроника", в которую вошли 32 предприятия и научно-исследовательских института бывшей электронной промышленности.

999

7 создано Российское агентство по системам управления, призванное координировать научную и производственную деятельность в сфере радиоэлектронике.

8 удалось увеличить производство изделий электронной техники на 147% по сравнению предыдущим годом.

9 объем привлеченных инвестиций в три раза превосходит объем государственных ассигнований.

0 сотрудники НПП "Исток" создали мощные многолучевые клистроны в дециметровом диапазоне для телевизионных передатчиков.

000

1 академик Алферов Ж. И. удостоен Нобелевской премии за основополагающие работы области информационных и коммуникационных технологий, в частности за открытие явления суперинжекции в гетероструктурах, открытия идеальных гетероструктур AlAs-GaAs, создание полупроводниковых лазеров на двойных гетероструктурах, создания первых биполярных гетеротранзисторов, солнечных батарей на гетероструктурах.

001

2 первые в мире в ФТИ РАН и АОЗТ "Светлана" разработан лавинно-пролетный диод на основе карбида кремния.

002

3 академик Месяц Г. А. первым получил премию по энергетике за выдающиеся работы области взрывной автоэлектронной эмиссии.

2003

Российские физики академики Гинзбург В. Л. (рис. П.32) и Абрикосов А. А. получили Нобелевскую премию по физике за работы в области сверхпроводимости и сверхтекучести.

Рис. П.32. Виталий Лазаревич Гинзбург — советский академик АН СССР, лауреат Нобелевской премии



ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

А

Автоволна 755
Акустоэлектроника 618
Амплитрон 125
Аналоговый интегратор 453
Анод 110
Аподизация преобразователя 647
Арифметическое логическое устройство 438
Аркотрон 164

Б

Базовый матричный кристалл 371, 460
Барьер Шоттки 230
БИС ПИН-прибор 712
Болометр 205

В

Вакуум 27
◊ высокий 34
◊ низкий 33
◊ средний 34
Вакуумметр:
◊ магнитный 45
◊ тепловой 39
◊ электронно-ионизационный 42
Варистор 205
Вентиль ртутный 164
Видикон 129
Вихри Абрикосова 728
Волна:
◊ Гуляева-Блюштейна 628
◊ Лява 628
◊ магнитостатическая 726
◊ магнито-упругая 727
◊ пространственного заряда 687
◊ Рэлея 628
◊ спиральная 728
◊ Стоунли 628

Волновод 559
Волоконный световод 565
Вольтамперная характеристика
p—n-перехода 225
Вторично-электронный умножитель 136
Вычитатель 454

Г

Газоразрядная индикаторная панель 166
Газотрон 164
Гексод 115
Геликон 687, 694, 728
Генератор 447
Гептод 115
Герметизация 338
Гетеропереход 205, 227
◊ анизотипный 227
◊ изотипный 227
Гетероструктура 227, 254
◊ двойная 541
Голография 599
Графekon 133

Д

Давление газа 28
Двумерный электронный газ 228, 255
Дебаевский радиус экранирования 151
Декатрон 165
Декод 116
Дематрон 125
Демультимплексор 382
Детектор 448
Дефект:
◊ примесный 210
◊ точечный 210
Дефлектор 569
◊ акустооптический 760
Дешифратор 382, 433
Диафрагма 83

Динамическая неоднородность 621

Диод 110

◊ IMPATT 203

◊ p—i—n 576

◊ TRAPPAT 203

◊ Ганна 204, 619

◊ лавинно-пролетный 203

◊ туннельный 202

◊ Шоттки 290

Дислокация 210

Дисплей 127

Диссекторы 131

Дифференциальный усилитель 450

Дифференциатор 453

Диффузионная длина 219

Диффузия:

◊ в замкнутой системе 321

◊ в открытой системе 321

◊ внедрения 320

◊ высокотемпературная 322

◊ замещения 320

◊ радиационно-стимулированная 326

Диэлектрик 188

Добротность 296, 632

◊ резонатора 526

Домен:

◊ Ганна 685, 720

◊ сегнетоэлектрический 671

◊ цилиндрический магнитный 724

Доменная граница 725

Дырка 188, 211

Е

Емкость:

◊ диффузионная 225

◊ барьерная 225

З

Закон:

◊ Больцмана 515

◊ Мура 475

◊ Пашена 159

◊ Снелля 83

◊ Столетова 61

◊ Фаулера 62

◊ Фаулера-Нордгейма 65

◊ Фика 320

◊ Эйнштейна 61

Замедляющая система 91

Запоминающее устройство 406, 441, 678

◊ акустическое 662

◊ голографическое 602

◊ на магнитных вихрях 736

◊ на ПЗС 709

◊ на цилиндрических магнитных доменах 734

◊ оперативное 406

◊ оптическое 745

◊ перепрограммируемое 406

◊ постоянное 406

Зарядовый пакет 684

Затвор 195, 248

◊ плавающий 413

Затравка 303

Зона:

◊ валентная 188, 215

◊ запрещенная 188, 215

◊ проводимости 188, 215

И

Игнитрон 164

Излучатель:

◊ когерентный 593

◊ некогерентный 591

Излучение:

◊ корпускулярное 154

◊ синхротронное 153

◊ тормозное 153

◊ циклотронное 153

Изокон 131

Иконоскоп 128

Инжектрон 126

Инжекция 223, 259

Интегральная оптика 583

◊ элементная база 584

Интегральная схема 197, 365

◊ аналоговая 371, 447

◊ большая 370

◊ гибридная 370

◊ заказная 461

◊ комбинированная 370

◊ малая 370

◊ монолитная арсенид-галлиевая 473

◊ полузаказная 460

◊ полупроводниковая 368

◊ программируемая логическая 464

◊ сверхбольшая 200, 370

◊ ультрабольшая 370

◊ цифровая 370

гегрально-оптическая схема 587
 гибридная 587
 монолитная 589
 гегральный:
 диод 288
 конденсатор 293
 резистор 291
 стабилизатор 288
 гегрференционная картина 599
 щная имплантация 323
 щная литография 314
 шаблон 314
 ок 195, 248

К

дмикон 129
 гал 248
 встроенный 250
 индуцированный 249
 чаллирование 323
 чтилевер 176
 эматрон 126
 осинотрон 122
 год 110
 годоллюминесценция 96
 антовая:
 нить 481
 точка 481, 483, 491
 лгтовое ограничение 480
 антовый:
 интерферометр 492
 каскадный лазер 492
 логический элемент 491
 колодец 228
 переход 511
 нескоп 126
 истрон 89, 118
 герентность 598
 света 529
 манда 438
 мпаратор 382, 455
 мплементарная структура 251
 мпрессия 67
 нвольвер 659, 719
 нтинуальная среда 631
 нтинуум 27
 нтрасткон 131
 ррелятор 702
 рфффициент:
 диффузии 320
 инъекции 259

◇ металлизации 633
 ◇ Эйнштейна:
 ▫ для вынужденного испускания 513
 ▫ для поглощения 512
 ▫ для спонтанного испускания 512
 ◇ электрохимической связи 631
 Кремний 302
 Кремнийкон 130
 Кристаллы:
 ◇ ионный 209
 ◇ ковалентный 210
 ◇ металлический 210
 ◇ молекулярный 210
 Критерий Кнудсена 33
 Кроссовер 67

Л

Лазер 203, 525
 ◇ газовый 544
 ◇ газодинамический 550
 ◇ гетеро- 541, 593
 ◇ жидкостный 542
 ◇ инжекционный 540
 ◇ ионный 546
 ◇ молекулярный 547
 ◇ на парах металла 552
 ◇ на свободных электронах 553
 ◇ полупроводниковый 538
 ◇ с распределенной обратной связью 543, 593
 ◇ с распределенным брегговским отражателем 593
 ◇ твердотельный 536
 ◇ химический 551
 ◇ эксимерный 549
 Лампа:
 ◇ бегущей волны 121
 ◇ генераторная 116
 ◇ модуляторная 118
 ◇ обратной волны 121
 ◇ приемно-усилительная 117
 ◇ прямой волны 121
 ◇ электронная 110
 Легирование 319
 Линза:
 ◇ геодезическая 584
 ◇ голограммная 602
 ◇ иммерсионная 84
 ◇ короткая магнитная 88
 ◇ Ломеберга 584
 ◇ одиночная 85

- ◊ Френеля 584
- ◊ электронная 83
- Линия задержки 640, 701, 733
- ◊ дисперсионная 644
- ◊ многоотводная 643
- Литография 304
- Литокон 133
- Логика:
 - ◊ диодно-транзисторная 389
 - ◊ интегральная инжекционная 387
 - ◊ на полевых транзисторах 397
 - ◊ резисторно-емкостная 387
 - ◊ резисторно-транзисторная 386
 - ◊ транзисторная с непосредственными связями 385
 - ◊ транзисторно-транзисторная 391
 - ◊ с диодами Шоттки 391
 - ◊ эмиттерно-связанная 391
- Логическая схема:
 - ◊ комбинационная 382
 - ◊ последовательностная 382
- Логический элемент 380
 - ◊ динамический 382
 - ◊ инвертирующий 383
 - ◊ неинвертирующий 383
 - ◊ статический 382

М

- Магнетрон 123
- Магنون 728
- Манометр 39
- Межсоединения 468
- Металлизация:
 - ◊ взрывной метод 337
 - ◊ катодное распыление 336
 - ◊ термический способ 335
 - ◊ термоионное испарение 337
- Метод:
 - ◊ боковой изоляции диэлектрической изоляции V-канавками 287
 - ◊ гетеродинирования 579
 - ◊ двойной экспозиции 607
 - ◊ диффузии 320
 - ◊ диффузионный 194
 - ◊ зонной плавки 302
 - ◊ зонной чистки 194
 - ◊ изоляции диэлектриком 286
 - ◊ изоляции комбинированный 287
 - ◊ изоляции обратносмещенным р—п-переходом 286
 - ◊ интерферометрический 607

- ◊ ионного легирования 202
- ◊ Ленгмюра-Блонджетт 750
- ◊ молекулярной инженерии 748
- ◊ пространственной фильтрации 611
- ◊ процессорного ядра 464
- ◊ спекл-интерферометрии 607
- ◊ течения:
 - ◊ манометрический 49
 - ◊ масс-спектрометрический 51
- ◊ травления 202
- ◊ Чохральского 303
- ◊ электрохимический 194
- ◊ эпитаксиальный 197
- Микроконтроллер 442
- Микрополосные линии 469
- Микропроцессор 371, 437
- Микроскоп:
 - ◊ ионный 171
 - ◊ сканирующий зондовый 175
- Микро-ЭВМ 442
- Микроэлектроника 198
- Минизона 480
- Митрон 125
- Многополосковый ответвитель 636
- Мода 527
 - ◊ волноводная 560
- Модулятор 448, 569
 - ◊ акустооптический 572, 760
 - ◊ волноводный 744
 - ◊ магнитооптический 571
 - ◊ частотный 571
- Монохроматическое колебание 529
- Мультиплексор 382, 434, 708

Н

- Накачка 515
- Накопитель 406
- Нанотехнология 476
- Нанотрубка 497
- Наноэлектроника 476
- Напряжение:
 - ◊ насыщения 250
 - ◊ отсечки 252
 - ◊ пороговое 249
 - ◊ пробоя 225
- Насос:
 - ◊ бустерный 36
 - ◊ двухроторный 35
 - ◊ диффузионный 36
 - ◊ ионно-геттерный 37
- Продолжение рубрики см. на с. 796*

ос (прод.):
 риогенный 38
 агниторазрядный 38
 механический с масляным уплотнением
 5
 рбитронный 37
 орбционный 37
 трийный 36
 урбомолекулярный 35
 ители заряда 211
 диффузия 217
 рейф 218
 еосновные 213
 сновные 213
 викон 130

оэлектроника 493
 од 115
 ческий контакт 229
 улятор 553
 рационный усилитель 449
 нвертирующий 452
 югарифмический 455
 еннвертирующий 452
 роника 591
 оэлектроника 591
 рон 595
 икон 130
 лоняющая система:
 лектромагнитная 82
 лектростатическая 80

од 114
 ванс 67
 иферия 406
 септрон 596
 аGRID 115
 оокоэффициент 674
 ооэлектрик 674
 изма 146
 идеальная 151
 квазинейтральная 150
 степен ь ионизации 150
 измон 687
 омбикон 129
 движность носителей заряда 218

Позистор 205
 Положительная обратная связь 526
 Полупроводник 188, 209
 ◇ п-типа 212
 ◇ р-типа 212
 ◇ собственный 211
 Поляритон 688
 Полярон 202, 687
 Пресобразователь 448
 ◇ аналогово-цифровой 457
 ◇ встречно-штыревой 633, 699
 ◇ цифроаналоговый 457
 ◇ электронно-оптический 133
 Прибор с зарядовой связью 619
 Приемник излучения 574
 Примесь 213
 Проводимость:
 ◇ дырочная 189, 212
 ◇ собственная 211
 ◇ электронная 189
 Процессор 702
 ◇ акустооптический 761
 ◇ оптоэлектронный 743
 Пушка:
 ◇ Пирса 67
 ◇ электронная 67
 Пьезокерамическая матрица 680
 Пьезокерамический двигатель 177
 Пьезоэффект 633

Р

Работа выхода 219
 Разрешающая способность 304
 Распределение:
 ◇ Максвелла 30, 149
 ◇ Ферми-Дирака 216
 Расходимость 530
 Ребикон 130
 Регистр 382, 438
 ◇ адреса 438
 ◇ команд 438
 ◇ общего назначения 438
 ◇ операндов 438
 ◇ сдвига 430
 ◇ состояния 438
 ◇ флажков 438
 Резонатор 526
 Рентгенолитография 311
 Рентгенорезист 312

С

Сатикон 129
Сверхрешетка 689
Сегнетомагнетик 676
Сегнетополупроводник 676
Сегнетоэлектрик 674
Секон 130
Селектор импульсов 448
Сенсор 176
Серродин 734
Сетка 112
◊ антидинаatronная 114
◊ защитная 114
◊ экранная 113
Сила Ван-дер-Ваальса 174
Синергетика 757
Сканирующая туннельная микроскопия 171
Скрайбирование 338
Слоистые структуры 677
Соленоид 87
Солитон 49, 618, 739, 750
Средняя длина пробега 31
Стабилитрон 165
Сток 195, 248
Страта 149
Сумматор 382, 430, 453
Супериконоскоп 128
Суперкомпьютер 444
Суперортикон 131
Схема вакуумная интегральная 117
Счетчик 383, 429
◊ асинхронный 430
◊ синхронный 430

Т

Таймер 439
Текнетрон 195
Тензорезистор 205
Термистор 205
Термовизор 204
Термогенерация 212
Термокомпрессия 338
Тетрод 113
Тиратрон 164
Тиристор 202
Токовый шнур 685
Травление 329
◊ полное 330
◊ ионно-химическое 331

◊ лазерно-стимулированное 333
◊ плазмохимическое 332
◊ сухое 330
◊ химическое 329
Транзистор 192, 247
◊ Terahertz 267
◊ V-МДП 257
◊ арсенид-галлиевый 270
◊ БнМОП 195
◊ биполярный 194, 247, 257
◊ вертикальный 201
◊ гетеробиполярный 205
◊ интегральный 197
◊ МДП 248
◊ меза- 194
◊ МЕР 253
◊ многоколлекторный 264
◊ многоэмиттерный 264
◊ на гетеропереходах 256
◊ на горячих электронах с резонансным туннелированием 489
◊ одноэлектронный 493
◊ плоскостной 194
◊ полевой 195
◊ с вертикальным затвором 269
◊ с высокой подвижностью электронов 472
◊ с двумя затворами 268
◊ с диодом Шоттки 265
◊ с управляемой проводимостью канала 268
◊ с управляющим р-п-переходом 252
◊ со встроенным инжектором 265
◊ составной 577
◊ униполярный 247, 248
◊ эпитаксиально-планарный 258
Транзисторный ключ 381
Трансьютер 443
Трансфазор 743
Триггер 382, 421
◊ D- 424
◊ J- 426
◊ RS- 423
◊ RST- 424
◊ T- 425
◊ Шмитта 422
Триод 112
Трубка осциллографическая 128
Туннелирование 481

У

Ударная ионизация 98
Указатель стека 439

равнение:

- Лапласа 83
- Леннарда-Джонса 175
- Менделеева-Клайперона 29
- Рамо 94
- Ричардсона-Дешмана 59
- Эйнштейна 218
- уровень Ферми:
 - в металлах 62
 - в полупроводниках 62, 216
- силитель 447, 654
- стройство управления 438

Ф

- Фазон 671
- Ферримагнетик 724
- Ферромагнетик 724
- Фильтр 448
 - голограммный 602
 - дисперсионный 651
 - на ПЗС 701
 - полосовой 645
 - трансверсальный 647
 - трансверсальный на ПЗС 702
- Флуоктон 672
- Фонон 211, 687
- Фотодиод 574
 - лавинный 575
 - на гетероструктуре 576
- Фототки 576
- Фотолитография 305
- Фотоника 591
- Фотонное эхо 741
- Фотоповторитель 307
- Фоторезист 305
- Фоторезистор 205, 578
- Фототиристор 577
- Фототранзистор 577
- Фотумножитель 135
- Фотшаблон 305
- Фотэлемент 134
 - вентильный 204
- Фотозффект 97
- Фуллерен 497
- Функциональная электроника:
 - акустическая 627
 - диэлектрическая 671
 - магнитная 724
 - молекулярная 748
 - оптическая 739
 - полупроводниковая 684
- Фурье-процессор 667

Х

- Халпикон 130

Ш

- Шина 408, 437
- Шифратор 382, 432
- Ш приоритетный 458

Э

- Экситон 202, 672
 - Ванье-Мотта 687
 - Френкеля 687
- Экситрон 164
- Экстракция 224, 260
- Электрический разряд 146
 - дуговой 148
 - коронный 148
 - таунсендовский 146
 - тлеющий 148
- Электрон 211
- Электроника:
 - вакуумная 57
 - интегральная 198
 - квантовая 511
 - плазменная 146
 - СВЧ 110
 - функциональная 617, 622
- Электронно-дырочный переход:
 - обратное включение 224
 - прямое включение 223
- Электронный проектор 170
- Электронолитография 308
 - проекционная 310
 - скапирующая 309
- Электропроводность 211
- Элемент связи 561
 - призмный 562
 - решеточный 563
- Эмиссия:
 - автоэлектронная 59, 65
 - взрывная электронная 66
 - вторичная электронная 63, 97
 - горячих электронов 64
 - ионно-электронная 67
 - кинетическая ионно-электронная 64
 - термоэлектронная 59
 - фотоэлектронная 61
 - экзоэлектронная 65

- ◇ электронная 58
- ◇ электронно-ионная 97
- Эпитаксия 316
- ◇ газофазная 316
- ◇ гетеро- 316
- ◇ гомо- 316
- ◇ графо- 328
- ◇ жидкофазная 318
- ◇ молекулярно-пучковая 317
- ◇ твердофазная 319
- Эффект:
 - ◇ Ааронова-Бомы 486
 - ◇ Ганна 720

- ◇ диатронный 114
- ◇ Керра 676
- ◇ кулоновской блокады 485
- ◇ Погкельса 570, 676
- ◇ туннельный 171
- ◇ Фарадея 571
- ◇ фотогальванический 574
- ◇ Шоттки 60

Я

- Ячейка памяти 407

А. А. Щука

ЭЛЕКТРОНИКА

- Вакуумная и плазменная электроника
- Полупроводниковая электроника и микроэлектроника
- Квантовая и оптическая электроника
- Функциональная электроника
- Решение задач по всем разделам курса

УЧЕБНОЕ ПОСОБИЕ

