

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЁВА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)»

А.Г. САНОЯН

ФИЗИКО-ТЕХНИЧЕСКИЕ ОСНОВЫ НАНОИНЖЕНЕРИИ

*Рекомендовано учебно-методическим объединением
высших учебных заведений Российской Федерации
по образованию в области прикладных математики
и физики в качестве учебного пособия для студен-
тов высших учебных заведений, обучающихся по
направлению «Прикладные математика и физика»
или по направлениям и специальностям в области
естественных наук, техники и технологии*

САМАРА
Издательство СГАУ
2011

УДК СГАУ: 539.1
ББК СГАУ: 31.4
С 186

Рецензенты: д-р физ.-мат. наук, проф. И. П. Завершинский;
д-р физ.-мат. наук, проф. В. А. Салеев

Саноян А. Г.

С 186 **Физико-технические основы наноинженерии:** учеб. пособие
/ А.Г. Саноян. – Самара: Изд-во Самар. гос. аэрокосм. ун-та,
2011. – 376 с.

ISBN 978-5-7883-0841-8

Представлен ряд актуальных физико-технических аспектов нанотехнологий и наноразмерных структур, имеющих непосредственное отношение к наноинженерии. Рассмотрена проблематика принципиальных и технических ограничений, связанных с созданием функциональных наноразмерных устройств. Показана определяющая роль квантово-механических представлений при анализе физико-технических параметров наноразмерных структур, элементов конструкций и вычислительных сред, выполненных на их основе. Подробно рассмотрены методы определения показателей качества нано- и микротехнологий, основанных на представлениях о степени упорядоченности наноразмерных атомно-молекулярных систем.

Пособие может быть рекомендовано студентам старших курсов, аспирантам, физикам-исследователям и инженерам различной отраслевой направленности, специализирующимся в области создания материалов и устройств наноинженерии.

Утверждено Редакционно-издательским советом университета в качестве учебного пособия

УДК СГАУ: 539.1
ББК СГАУ: 31.4

ISBN 978-5-7883-0841-8

© Самарский государственный
аэрокосмический университет, 2011

ОГЛАВЛЕНИЕ

Список использованных обозначений.....	9
Предисловие.....	13
Введение.....	17
Глава 1. Энергетические состояния микрочастиц в наноразмерных структурах с позиции квантово-механических представлений.....	23
1.1. Элементы квантовой механики.....	23
1.1.1. Корпускулярно-волновой дуализм.....	23
1.1.2. Квантово-механические принципы анализа состояния микрочастиц.....	24
1.1.3. Волновой вектор плоской монохроматической волны.....	27
1.1.4. Волновой пакет из вероятностных волн де Бройля.....	30
1.1.5. Квантование энергии микрочастицы, находящейся в потенциальном ящике.....	34
1.1.6. Волновая функция как характеристика состояния микрочастицы.....	38
1.1.7. Примеры квантования энергии микрочастиц в физических системах.....	44
а) Идеализация потенциальных атомных диаграмм.....	44
б) Движение микрочастицы в условиях отсутствия потенциальных полей.....	46
в) Квантование энергии микрочастицы, находящейся в потенциальном ящике.....	47
г) Квантование энергии электронов в водородоподобном атоме.....	49
1.2. Квантово-механические принципы анализа параметров электронных состояний в наноразмерных структурах.....	51
1.2.1. Общие принципы классификации наноразмерных структур.....	51
1.2.2. Энергетические состояния микрочастиц в наноразмерных структурах.....	54
а) Зонные энергетические диаграммы монокристаллических структур.....	54

б)	Спектр энергетических состояний в рамках модели слабосвязанных электронов.....	62
в)	Изоэнергетические поверхности состояния электронов в кристаллических структурах.....	68
г)	Спектр энергетических состояний в наноразмерных структурах типа (0D)	71
д)	Спектр энергетических состояний в наноразмерных структурах типа (1D)	75
е)	Спектр энергетических состояний в наноразмерных структурах типа (2D)	80
1.3.	Физические основы квантовой информации и квантовых компьютеров.....	84
1.3.1.	Физические основы квантовых вычислительных систем.....	84
а)	Проблематика и тенденции развития современной вычислительной техники.....	84
б)	Когерентная квантовая интерференция отдельных микрочастиц.....	89
в)	Общие принципы реализации квантовых вычислительных систем.....	94
г)	Символьный формализм представления квантовых состояний и процессов.....	98
д)	Суперпозиция квантовых состояний как основа вычислительных операций.....	105
е)	Общие физико-технические принципы преобразования информационного состояния кубитов.....	108
1.3.2.	Экспериментальные методы реализации квантовых компьютеров.....	114
а)	Проблематика экспериментальной реализации квантовых компьютеров.....	114
б)	Эксперименты по созданию вычислительных сред методом КЭР.....	116
1.4.	Теоретические и практические основы спинтроники.....	122
1.4.1.	Спин микрочастицы: определение и форма представления..	122
1.4.2.	Этапы становления и перспективы развития спинтроники...	124
1.4.3.	Магнитно-полупроводниковые спиновые нанотранзисторы...	127

Глава 2. Специфика механизмов электрической проводимости наноразмерных структур.....	131
2.1. Элементы теории электропроводности конденсированных сред.....	131
2.1.1. Проблематика и актуальность процессов транспорта носителей заряда в наноразмерных структурах.....	131
2.1.2. Элементы теории подвижности носителей заряда.....	135
а) Подвижность носителей заряда в атомно-молекулярных системах.....	135
б) Подвижность носителей заряда при механизме движения пролетного типа.....	139
в) Подвижность носителей заряда при механизме движения прыжкового типа.....	143
г) Механизм движения носителей заряда туннельного типа....	149
2.1.3. Концентрация носителей заряда в проводниках электронного типа.....	155
а) Механизм электропроводности кристаллических структур с позиции зонных энергетических диаграмм.....	155
б) Функция плотности и количества квантовых состояний.....	162
в) Вероятность заполнения квантовых состояний в кристаллических структурах.....	168
г) Полная функция распределения ансамбля микрочастиц по энергиям.....	174
д) Концентрация носителей заряда и электропроводность в кристаллических структурах металлического и полупроводникового типа.....	179
2.2. Элементы теории электропроводности наноразмерных структур.....	191
2.2.1. Общие ограничения применимости статистических методов анализа для случая наноразмерных структур.....	191
а) Статистические флуктуации характеристических параметров макро- и микросистем.....	191
б) Статистическая энтропия наноразмерных структур.....	196
2.2.2. Носители заряда для случая наноразмерных структур кристаллического типа.....	202

а) Специфика формализма зонных энергетических диаграмм для случая наноразмерных структур.....	202
б) Анализ возможных скоростей движения носителя заряда в электрическом поле.....	207
в) Влияние поверхности раздела физических сред на вид зонных энергетических диаграмм наноразмерных структур.....	215
г) Специфика баллистического транспорта носителей заряда в наноразмерных структурах.....	227
Глава 3. Энтропийные критерии качества и методы анализа нанотехнологий и наноразмерных структур.....	231
3.1. Замечания общего характера.....	231
3.2. Информационная энтропия.....	235
3.2.1. Элементы теории информации.....	235
а) Определение энтропии для случая равновероятных возможностей (дискретный случай).....	235
б) Энтропия для случая неравновероятных возможностей (дискретный случай)	238
в) Энтропия для случая неравновероятных возможностей (непрерывный случай)	242
г) Асимптотическая равновероятность и энтропийная устойчивость последовательности случайных событий.....	243
3.2.2. Основы информационной теории измерительных процессов.....	247
а) Применение основных положений теории информации для измерительных операций.....	247
б) Процесс измерения как фактор сужения интервала неопределенности.....	252
в) Практические методы определения энтропийного показателя погрешности измерения.....	257
г) Информационная емкость диагностических сигналов измерительного процесса.....	260
3.3. Статистические методы анализа кинетики протекания элементарных физико-химических процессов.....	262

3.3.1. Физическая сущность статистического метода анализа макроскопических систем (технологический аспект).....	262
3.3.2. Активационная модель протекания элементарных физико-химических процессов.....	265
3.3.3. Методика определения концентрации энергетически активных микрочастиц.....	266
3.3.4. Динамическое равновесие элементарных физико-химических процессов.....	267
3.3.5. Оценка срока службы материалов и изделий на основе методов физической надежности.....	270
3.4. Термодинамическая энтропия и упорядоченность структур.....	272
3.4.1. Элементы равновесной термодинамики.....	272
а) Основные понятия и определения.....	272
б) Внутренняя энергия термодинамической системы.....	274
в) Термодинамическая энтропия физико-химической системы.....	279
3.4.2. Элементы неравновесной термодинамики.....	284
а) Энтропия и необратимость физико-химических процессов..	284
б) Производство энтропии в неравновесной системе.....	286
в) Производство энтропии в линейных необратимых процессах.....	288
г) Вариационный принцип наименьшего рассеяния энергии в стационарных процессах.....	293
д) Принцип минимума производства энтропии.....	296
3.5. Энтропийные модели микро- и нанотехнологий и структур.....	298
3.5.1. Цели и задачи энтропийного подхода при анализе микро- и наноразмерных структур.....	298
3.5.2. Формальное представление технических объектов и технологии их создания.....	302
3.5.3. Показатели качества технологий на микроскопическом уровне рассмотрения.....	307
3.5.4. Показатели качества технологий на макроскопическом уровне рассмотрения.....	310

3.5.5. Показатели качества реальных микро- и нанотехнологий....	312
3.5.6. Информационный потенциал и информационный дефицит технологии.....	315
3.5.7. Информационный запас качества и его взаимосвязь со сроком службы изделия.....	317
3.5.8. Процесс создания микроструктур в рамках представлений алгоритмической энтропии.....	320
3.5.9. Алгоритмическая энтропия и самоорганизация сложных атомных структур.....	323
Заключение.....	330
Задачи и упражнения.....	333
Задачи и упражнения к разделу №1.....	333
Задачи и упражнения к разделу №2.....	339
Задачи и упражнения к разделу №3.....	350
Список использованной литературы.....	363
Приложение № 1. Справочные данные общего характера.....	369
Приложение № 2. Справочные данные о параметрах некоторых материалов.....	371
Приложение № 3. Общие рекомендации по решению задач и упражнений по разделу «контактные явления в твердотельных структурах».....	373

СПИСОК ИСПОЛЬЗОВАННЫХ ОБОЗНАЧЕНИЙ

- n_{n0} - равновесная концентрация электронов в полупроводнике донорного типа
- n_{p0} - равновесная концентрация электронов в полупроводнике акцепторного типа
- p_{p0} - равновесная концентрация дырок в полупроводнике акцепторного типа
- p_{n0} - равновесная концентрация электронов в полупроводнике акцепторного типа
- p_F - импульс электрона, обладающего энергией Ферми
- q_i - модуль значения электрического заряда для носителя заряда $i^{\text{го}}$ типа
- $f_{MB}(W)$ - функция распределения микрочастиц по энергиям (Максвелла-Больцмана)
- $f_{\Phi D}(W)$ - функция распределения микрочастиц по энергиям (Ферми-Дирака)
- k - модуль волнового вектора плоской монохроматической волны (де Бройля); постоянная Больцмана
- D - коэффициент прозрачности (туннельной) движения микрочастицы
- $\Delta\Gamma_P$ - элемент фазового объема в импульсном подпространстве
- $\Delta\Gamma_V$ - элемент фазового объема в координатном подпространстве
- $G(W)$ - функция количества квантовых состояний ансамбля микрочастиц
- $g(W)$ - функция плотности квантовых состояний ансамбля микрочастиц
- N - количество атомов, задействованных в процессе создания наноструктуры
- N_c - эффективное число энергетических состояний в зоне проводимости
- N_v - эффективное число энергетических состояний в валентной зоне

- N_a - концентрация атомов акцепторной примеси в полупроводнике
 N_o - концентрация атомов донорной примеси в полупроводнике
 $G(W)$ - полная функция распределения частиц по энергиям
 S - термодинамическая энтропия макроскопической системы
 H - информационная энтропия макроскопической системы
 I - количество информации (устраненной неопределенности состояния системы)
 E - напряженность электрического поля
 G_o - скорость генерации неравновесных (избыточных) носителей заряда
 P_n - вероятность рекомбинации электрона в зоне проводимости
 P_p - вероятность рекомбинации дырки в валентной зоне
 P_i - вероятность проявления определенного события (исхода)
 P_{i-j} - вероятность квантового перехода системы из состояния i в состояние j
 $|Q\rangle$ - условное обозначение квантового состояния кубита
 R - скорость рекомбинации носителей заряда
 S_n - площадь сечения взаимодействия электрона с дыркой (сечение захвата)
 W - энергия микрочастицы
 W_k - кинетическая энергия микрочастицы
 W_T - энергия теплового движения микрочастицы
 W_o - ширина запрещенной зоны полупроводника
 W_c - энергетический уровень дна зоны проводимости
 W_v - энергетический уровень потолка валентной зоны
 W_a - энергия активации элементарного физико-химического процесса
 N_a - концентрация энергетически активных микрочастиц
 ΔW - ширина запрещенной зоны собственного полупроводника

- W_F - энергия Ферми
 U - функция потенциальной энергии системы
 L - протяженность потенциальной ямы
 V - пространственный (или фазовый) объем системы
 Z - порядковый номер химического элемента
 $(0D)$ - условное обозначение квантовой точки
 $(1D)$ - условное обозначение квантовой проволоки
 $(2D)$ - условное обозначение квантовой пленки
 h - постоянная Планка
 p - механический импульс микрочастицы
 \bar{k} - волновой вектор монохроматической волны
 m - масса микрочастицы
 m_e - эффективная масса электрона
 n - квантовое число энергетического состояния
 n_i - концентрация носителей зарядов $i^{\text{го}}$ типа
 c - скорость распространения света в вакууме
 v - скорость движения микрочастицы
 s - удельная термодинамическая энтропия системы
 h_t - энтропия размещения единичного атома при наличии технологии
 h_c - энтропия размещения единичного атома при случайной технологии
 h_i - энтропия размещения единичного атома, обеспечивающая появление функционального качества изделия
 Ψ - волновая функция микрочастицы
 α_c - избирательность при отсутствии технологии как таковой
 α_t - избирательность при наличии технологии
 α_i - избирательность, обеспечивающая появление функционального качества изделия
 φ - электрический потенциал
 ε - относительная диэлектрическая проницаемость вещества
 ε_0 - электрическая постоянная

- γ - коэффициент рекомбинации избыточных носителей заряда
- ν - частота монохроматических колебаний
- λ - длина волны монохроматического оптического излучения
- τ_n - время жизни избыточных электронов (в зоне проводимости)
- τ_p - время жизни избыточных дырок (в валентной зоне)
- μ - подвижность носителей заряда
- σ - электрическая проводимость вещества
- ω - круговая частота монохроматической волны
- ϖ - функция плотности вероятности распределения микро-частицы в пространстве
- φ - фазовый угол монохроматической волны
- \mathcal{G} - фазовая скорость плоской монохроматической волны
- δ - начальная фаза монохроматической волны
- ω - круговая частота монохроматической волны
- ϕ - амплитудная волновая функция

ПРЕДИСЛОВИЕ

Настоящее учебное пособие является введением в теорию наноинженерии и предназначено для студентов, аспирантов и специалистов различных научных и технических направлений, желающих углубить свои знания в этой сравнительно новой и быстроразвивающейся области знаний. Пособие написано на основе материалов лекционного курса, изучаемого в Самарском государственном аэрокосмическом университете в рамках подготовки инженеров, бакалавров и магистров по различным научно-техническим направлениям и специализациям.

Для написания данного учебно-методического пособия по проблематике актуальных задач наноинженерии, в дополнение к широкому кругу изданий, имеющихся в настоящее время, существует несколько оснований.

Во-первых, по опыту преподавания ряда дисциплин, относящихся к направлению наноинженерии, стало известно, что новичку в этом инженерно-техническом направлении достаточно трудно порекомендовать какой-либо один литературный источник. Как правило, в каждом из этих источников с определенной полнотой рассмотрены отдельные вопросы, но ни один из них не затрагивает всех аспектов (с учетом динамики развития и современных достижений в области нанотехнологий и физики наноразмерных систем) с достаточно общей точки зрения. Таким же образом при формировании перечня рекомендованной дополнительной литературы возникает та же дилемма и обычно рекомендуют отдельные части различных книг. При этом основные расхождения между этими источниками, обусловленные различием интересов авторов и методик изложения учебного материала, существенно затрудняют процесс обучения.

Во-вторых, следует отметить тот негативный факт, что в сложившейся учебной практике наблюдается существенное расхождение между планируемым объемом трудозатрат и фактическим (в значительной мере превосходящим планируемый). Это расхождение обусловлено необходимостью использования широкого круга специальных литературных источников, ориентированных, как правило, на уровень хорошо подготовленного специалиста в данной конкретной области знаний.

В-третьих, в существующих литературных источниках (к ним, как правило, следует отнести монографии по ряду специальных разделов, относящихся к наноэлектронике и нанотехнологии) в недостаточной мере представлены учебно-методические аспекты. К ним можно отнести, в частности, такой важный, с точки зрения постановки образовательного процесса, момент, как комплексы задач и упражнений, адаптированные к тезаурусу студентов старших курсов и аспирантов технических вузов, специализирующихся по непрофильным научно-техническим направлениям.

Изложенные выше моменты нашли свое непосредственное отражение в настоящем учебно-методическом пособии как в части методических аспектов представления учебных материалов, так и в построении общей структуры пособия.

Основным требованием к базовой подготовке студентов, необходимым для успешной работы с пособием, является наличие базовых основ по ряду разделов физики (квантовая механика, статистическая физика и физика твердого тела – в объемах, предусмотренных учебным планом технических вузов) и теории информации. В то же время следует отметить, что опыт преподавания дисциплин по рассматриваемому направлению свидетельствует о недостаточной подготовке слушателей (особенно непрофильных специальностей: специалисты машиностроительного направления, химики, биологи) по указанным выше разделам физики и теории информации. В этой связи автор счел целесообразным и необходимым частично восполнить этот пробел в рамках учебного материала настоящего пособия.

Для успешного усвоения учебного материала и привития навыка к расчетной практике каждому разделу адресован комплекс задач и упражнений различной степени сложности. Тематика указанного комплекса сформирована таким образом, чтобы представлять интересы не только конкретной профессиональной группы учащихся и специалистов (наноэлектроника), но и более широкого (оговоренного выше) контингента учащихся.

Описанное выше построение учебного пособия позволяет читателю использовать его как вводный курс по наноинженерии в целом или об-

ращаться к его отдельным разделам по конкретным вопросам. Таким образом, пособие может оказаться полезным физикам, химикам, биологам и инженерам, специализирующимся в области исследования наноматериалов и создания наноразмерных структур и устройств, а также специалистам близких по профилю дисциплин, желающим получить начальную базовую подготовку по данному научно-техническому направлению.

Разумеется, что в рамках одного пособия невозможно охватить всю проблематику современных задач и дать развернутый анализ достижений в направлении– «наноинженерия». Стремительное и всеохватное развитие данного направления не оставляет возможности полноценно и однозначно отследить ее состояние на данный момент времени. В то же время общие теоретические положения и подходы, обладающие большей «инерцией» в плане динамики их трансформации, могут составить предмет рассмотрения в учебном плане. Именно этим принципом руководствовался автор (учитывая содержательную базу ранее изданных пособий других авторов по рассматриваемому научно-техническому направлению) при отборе материалов для настоящего пособия.

И конечно, при написании настоящего пособия автору не удалось избежать элементов личного предпочтения к материалам, отражающим его собственные научно-технические интересы по рассматриваемому научно-техническому направлению. В наибольшей мере эти моменты нашли свое отражение в разделе, посвященном энтропийным критериям и моделям оценки качества нанотехнологий и наноразмерных изделий. Целесообразность включения этих материалов в круг рассмотрения обусловлена тем обстоятельством, что в разряде наноразмерных технических устройств все чаще стали встречаться и такие устройства (и технические системы, построенные на их основе), которые содержат счетное количество индивидуальных атомов (в том числе есть и такие, в которых фигурирует всего один атом). В этой связи возникает вопрос, каковы должны быть требования к качеству проведения технологических процессов, чтобы производство этих изделий было целесообразным с позиции традиционных представлений о выходе годных изделий.

В рамках такой постановки вопроса технология создания микро- и наноустройств рассматривается как комплекс мер, направленных на

повышение вероятности размещения атомов определенного типа в наперед заданных точках физического пространства. Вероятностный подход при рассмотрении технологической проблематики реализуется с позиций методологии теории информации. В этой связи предлагаются энтропийные критерии оценки качества микро- и нанотехнологий, позволяющие сформировать требования к технологическим процессам во взаимосвязи со сложностью и выходом годных изделий, и вводятся в рассмотрение представления об информационном потенциале и дефиците микротехнологий. Приводятся результаты количественных оценок энтропийных показателей качества современных микро- и перспективных нанотехнологий, используемых при создании твердотельных структур.

Автор надеется, что представленные в пособии материалы позволят читателю ознакомиться с рядом основополагающих физико-технических аспектов наноинженерии с целью дальнейшего их использования в своей практической деятельности по указанному направлению.

Все отзывы и замечания по содержательным и методическим аспектам данного пособия будут приняты автором с благодарностью. Автор выражает признательность С.Б. Коныгину и И.Н. Козловой за оказанную помощь при подготовке ряда рисунков и упражнений настоящего учебного пособия.

*«Всякая вещь есть форма проявления
беспредельного разнообразия»
Козьма Прутков*

ВВЕДЕНИЕ

Впервые термин «нанотехнология» употребил Н. Танигутчи в 1974 году [1]. Он назвал этим термином технологические методы, используемые для производства изделий размером в несколько нанометров. В многих источниках, и в первую очередь англоязычных, первое упоминание методов, которые впоследствии были названы нанотехнологией, связывают с известным выступлением Р. Фейнмана «Там внизу много места» (англ. «*There's Plenty of Room at the Bottom*»), сделанным им в 1959 году в Калифорнийском технологическом институте, в котором он признал допустимой, с позиции физических законов, возможность механического перемещения одиночных атомов при помощи манипулятора соответствующего размера [2].

В международных документах нормативного плана (технический комитет – ISO/TK 229) под понятием нанотехнология принято подразумевать следующее определение:

- знание и управление процессами, как правило, в масштабе 1 нм, но не исключающее масштаб менее 100 нм, в одном или более измерениях, когда ввод в действие размерного эффекта (явления) приводит к возможности новых применений;
- использование свойств объектов и материалов в нанометровом масштабе, которые отличаются от характеристик свободных атомов или молекул, а также от объемных свойств вещества, состоящего из этих атомов или молекул, для создания более совершенных материалов, приборов, систем, реализующих эти свойства.

Согласно «Концепции развития в Российской Федерации работ в области нанотехнологий» (2004 – 2010 гг.), нанотехнология определяется как *«совокупность методов и приемов, обеспечивающих возможность контролируемым образом создавать и модифицировать объекты, включающие компоненты с размерами менее 100 нм хотя бы в од-*

ном измерении, и в результате этого получившие принципиально новые качества, позволяющие осуществлять их интеграцию в полноценно функционирующие системы большего масштаба».

Практический аспект нанотехнологий включает в себя процессы производства устройств, необходимых для создания, обработки и манипуляции атомами, молекулами и наночастицами [3-5]. Подразумевается, что не обязательно объект должен обладать хоть одним линейным размером менее 100 нм – это могут быть макрообъекты, атомарная структура которых контролируемым образом создаётся с разрешением на уровне отдельных атомов (к примеру, это может быть матрица на основе квантовых точек), либо же содержащие в себе нанообъекты. В более широкой трактовке этот термин охватывает также методы диагностики и исследования объектов наноразмерного масштаба.

Нанотехнологии качественно отличаются от традиционных подходов, поскольку на таких масштабах привычные макроскопические технологии обработки материалов зачастую неприемлемы, а физические явления, пренебрежительно слабые на привычных макроскопических масштабах, становятся намного значительнее. В более частной постановке под этим подразумевается проявление весьма сильного влияния квантовых эффектов на физические параметры и механизмы взаимодействия отдельных атомов, молекул и атомно-молекулярных комплексов [6]. Аналогичное заключение относительно ключевой роли квантовых эффектов можно распространить и на наноматериалы, в которых размерные ограничения существенным образом отражаются на свойствах материалов и элементов конструкций [7, 8]. В целом можно отметить, что наличие масштабных факторов существенно ограничивает возможность широкого использования традиционных термодинамических методов анализа структур и процессов, основанных на физике сплошных сред.

Следует отметить, что нанотехнология (и в особенности молекулярная технология и технология атомной сборки) является новой, очень мало исследованной в теоретическом плане областью знаний, находящейся более на стадии академических, нежели прикладных исследований. Основные открытия, предсказываемые в этой области, пока еще не

сделаны. Тем не менее, проводимые исследования уже в настоящее время дали ряд ценных научно-практических результатов, которые позволяют отнести нанотехнологии к классу «высоких технологий».

Весьма характерно, что активный интерес к нанотехнологиям и наноматериалам проявляют специалисты различного профессионального профиля: физики, инженеры различных специальностей, химики, биологи, медики. Все они в указанных направлениях просматривают возможность реализации своих перспективных прорывных проектов, основанных на более высокой ступени развития научно-технического знания. Фронт работ по этим направлениям настолько широк и порой столь специфичен, что в настоящее время очень затруднительно, а подчас и невозможно провести однозначную классификацию частных направлений и методологических аспектов научно-исследовательских и прикладных исследований.

С точки зрения экономической целесообразности, реализация нанотехнологий в широких промышленных масштабах представляется проблематичной (имеется в виду использование технологии атомной сборки при создании изделий). В этой связи одной из важнейших проблем, стоящих перед нанотехнологией, является разработка теоретических и прикладных аспектов для обеспечения фактора самоорганизации наноразмерных структур в производственном цикле атомной сборки изделий.

Частично проблематика самоорганизации наноструктур составляет предмет интереса *супрамолекулярной химии* [9], изучающей не отдельные молекулы, а специфику межмолекулярных взаимодействий, способных реализовать фактор упорядоченности атомно-молекулярных систем. Обнадёживает то, что в природе действительно наблюдаются подобные атомно-молекулярные системы (нуклеотиды, белки, РНК, ДНК) и осуществляются подобные процессы управляемого синтеза сложных структур. Объекты супрамолекулярной химии – *супрамолекулярные ансамбли*, строящиеся самопроизвольно из комплементарных структур (т.е. имеющих геометрическое и химическое соответствие индивидуальных фрагментов), подобно самопроизвольной сборке сложнейших пространственных структур в живой клетке.

В более общей постановке следует говорить о необходимости использования *синергетических* (кооперативных) методов анализа процессов, способных реализовать феномен атомно-молекулярной упорядоченности пространственных структур в мезо- и макроскопических масштабах. Этот подход представляется весьма перспективным, однако его практическое использование сдерживается в настоящее время недостаточной теоретической проработкой фундаментальных принципов самоорганизации сложных систем [10, 11].

Наиболее близко к вопросу комплексного использования достижений в области нанотехнологий и наноматериалов (с позиции функциональной направленности работ) подошла современная микроэлектроника, в части обеспечения необходимой степени интеграции элементов. Это положение обусловлено тем, что классические (традиционные) методы производства устройств микроэлектроники подошли к своему естественному экономическому и технологическому барьеру, когда размер устройства уменьшается не намного, а экономические затраты возрастают экспоненциально. В этом смысле нанотехнология является следующим логическим шагом развития микроэлектроники. Уже в настоящее время новое поколение микропроцессоров основано на структурных элементах размером 45 нм, с перспективой освоения в ближайшее время размеров структурных элементов до 5 нм.

В плане дальнейшего развития микропроцессорной техники нельзя не упомянуть о *квантовом компьютере*, т.е. перспективном вычислительном устройстве [12], работа которого основана на использовании специфических квантово-механических эффектов, таких как *квантовый параллелизм* и *квантовая запутанность* [13]. Выигрыш в производительности (а следовательно и в быстродействии) квантовых алгоритмов обработки информации достигается за счёт того, что при применении одной квантовой операции большое число коэффициентов суперпозиции квантовых состояний, которые в виртуальной форме содержат классическую информацию, преобразуется одновременно. Таким образом, квантовое вычислительное устройство размером L кубит (некий аналог классического понятия – «бит») может выполнять параллельно 2^L вычислительных операций. Этот фактор порождает уверенность в

возможности существенного повышения тактико-технических данных вычислительных средств.

В свете рассмотренной выше технологической проблематики наноинженерии нельзя не упомянуть о направлении, непосредственно связанном с созданием наноразмерных роботов, размеры которых сопоставимы с молекулярными объектами (не более 10 нм). Нанороботы, обладающие функциями механического движения, обработки и передачи информации, способны к созданию своих копий, то есть к самопроизводству себе подобных устройств. В более общей трактовке под определением наноробота подразумевается макроскопическое устройство, способное точно взаимодействовать с наноразмерными объектами или – манипулировать объектами в наноразмерном масштабе перемещений. Вследствие этого даже крупные аппараты, такие как современные атомно-силовые микроскопы, можно считать в качестве нанороботов, так как они производят манипуляции объектами на наноскопическом уровне.

Из приведенного выше (далеко не полного) перечня работ, проводимых в направлении «наноинженерия» (включающем в своем составе как частные направления нанотехнологию и наноматериалы), можно сделать принципиальный вывод о том, что это направление научно-исследовательских и прикладных работ имеет ярко выраженный междисциплинарный характер, предполагающий для своего успешного развития:

- комплексный подход на этапах формирования постановки и реализации научно-практических работ;
- объединения усилий специалистов весьма широкого научно-технического профиля (физики, инженеры различных специальностей, химики, биологи, медики);
- привлечение всего комплекса достигнутых в настоящее время научных знаний практически из всех областей человеческой деятельности.

Из последнего обстоятельства следует, что к созданию перспективных изделий наноинженерии будут привлечены специалисты различных научно-технических направлений, с различным начальным уровнем физико-математической подготовки и профилем основных профес-

сиональных интересов. Наблюдая тенденцию широкомасштабного проникновения в соседние области научно-технических знаний, становится актуальной постановка вопроса о сближении теоретических подходов и методологических аспектов указанных специалистов. Трудность реализации подобного рода процесса сближения обусловлена тем немаловажным обстоятельством, что проблематика задач нанотехнологии предполагает использование ряда передовых рубежей знаний по физике твердого тела, физической химии, синергетике, теории информации. В то же время следует отметить недостаточную теоретическую подготовку инженерных кадров и специалистов химического и биологического профиля в части:

- использования традиционных методов анализа характерных параметров атомных (молекулярных) структур наноразмерного масштаба, основанных на квантово-механических моделях и представлениях;
- методов определения необходимой, с точки зрения достижения приемлемого функционального качества, меры упорядоченности физических сред, используемых при создании устройств наноразмерного масштаба;
- понимания физической специфики разнообразных размерных эффектов, имеющих место в микро- и наноразмерных системах.

В свете вышеизложенных моментов настоящее учебное пособие призвано восполнить существующий в настоящее время образовательный пробел в перечисленных выше областях инженерно-физических знаний.

Глава 1. ЭНЕРГЕТИЧЕСКИЕ СОСТОЯНИЯ МИКРОЧАСТИЦ В НАНОРАЗМЕРНЫХ СТРУКТУРАХ С ПОЗИЦИИ КВАНТОВО-МЕХАНИЧЕСКИХ ПРЕДСТАВЛЕНИЙ

1.1. Элементы квантовой механики

1.1.1. Корпускулярно-волновой дуализм

Зарождение квантовых представлений принято связывать с исторической статьей М. Планка «К теории распределения энергии излучения в нормальном спектре», в которой он ввел в рассмотрение универсальную физическую постоянную h . Суть квантовой гипотезы М. Планка состояла в том, что любая энергия поглощается или испускается светом только *дискретными порциями* с энергией W , пропорциональной частоте ν излучаемой световой волны:

$$W = h \cdot \nu, \quad (1.1)$$

где h – постоянная Планка, численно равная $6,62 \cdot 10^{-34}$ Дж·с.

Используя квантовую гипотезу М. Планка для объяснения явлений фотоэффекта, А. Эйнштейн (1905 г.) предположил, что свет также состоит из отдельных *квантов*, которые впоследствии были названы фотонами и рассмотрены в качестве своеобразных микрочастиц.

Для объяснения структуры атома Н. Бор (1913 г.) предположил существование *стационарных энергетических состояний* электрона в атомах, у которых энергия может принимать лишь ряд дискретных значений. Этот подход, развитый Зоммерфельдом и рядом других физиков, принято называть старой квантовой теорией, отличительной чертой которой является сочетание традиционных классических представлений и противоречащих им дополнительных квантовых предположений.

В 1923 году Луи де Бройль выдвинул гипотезу о двойственной природе вещества, основанную на том, что движение материальных частиц обладает как корпускулярными, так и волновыми свойствами. Л. де Бройль сопоставил движение микрочастицы с распространением некой волны, что получило экспериментальное подтверждение при исследовании дифракции электронов в кристаллах.

В 1927 году К. Дэвиссон и Л. Гермер демонстрируют дифракцию медленных электронов на кристаллах. При интерпретации угловой зависимости интенсивности отраженных электронов было показано её соответствие дифракции рентгеновских лучей. До принятия гипотезы де Бройля дифракция расценивалась как исключительно волновое явление, а любой дифракционный эффект – как волновой. Таким образом, гипотеза де Бройля была экспериментально подтверждена для случая микрочастиц (электрона). Этот факт стал поворотным моментом на пути дальнейшего становления как самой квантовой теории, так и развития фундаментальных представлений о физической природе материального мира.

Подобно тому как А. Комптон демонстрирует корпускулярную природу света, эксперимент Дэвиссона-Гермера подтвердил неразрывное «сосуществование» с микрочастицей её волновых представлений, т.е. присущность ей некой волновой природы. Подтверждение этой идеи для физики стало важным этапом, поскольку дало возможность не только характеризовать любую микрочастицу, присваивая ей определённую индивидуальную длину волны, но и полноправно используя её в виде объективного физического фактора в волновых представлениях эволюции состояния частиц.

Из вышеизложенного следует важный для последующего изложения вывод, что для всех физических объектов (в особенности это характерно для элементарных микрочастиц) экспериментально зафиксирован двойственный характер их поведения, что и получило название *корпускулярно-волнового дуализма* [1].

1.1.2. Квантово-механические принципы анализа состояния микрочастиц

Как отмечалось выше, в основе квантовой механики лежит представление о дискретном характере изменения энергии индивидуальных атомов и фотонов. основополагающей концепцией в квантовой механике явилось то, что корпускулярно-волновая двойственность свойств, установленная для света, имеет универсальный характер и должна проявляться для любых частиц, обладающих механическим импульсом. Те все частицы, обладающие конечным импульсом, проявляют волновые

свойства, и их движение можно представить с помощью некоего волнового процесса.

Гипотеза Л. де Бройля: движению свободной частицы (т.е. для случая отсутствия в пространстве каких-либо внешних по отношению к микрочастице физических полей) ставится в соответствие плоская монохроматическая волна (ПМВ), длина волны λ и частота ν которой определяются выражениями:

$$\nu = W/h \quad (1.2)$$

(аналог выражения (1.1) для случая микрочастиц);

$$\lambda = h/p = h/(m \cdot \nu), \quad (1.3)$$

где W и p – соответственно энергия и механический импульс микрочастицы; m – масса микрочастицы; ν – скорость движения микрочастицы.

Волны, о которых идет речь, принято называть *волнами де Бройля*. Нетрудно заметить, что длина волны де Бройля для частицы, имеющей кинетическую энергию W_k , равна

$$\lambda = \frac{h}{\sqrt{2 \cdot m \cdot W_k}}. \quad (1.4)$$

В практике использования электронно-лучевых технологий (в частности для оценки разрешающей способности электронной литографии) удобно использовать выражение вида

$$\lambda = \frac{12.25}{\sqrt{\varphi}} \text{ (ангстрем)}, \quad (1.5)$$

где φ – электрический потенциал (вольт), используемый для обеспечения процесса ускорения электронов.

Гипотеза де Бройля была экспериментально подтверждена опытами по рассеянию электронов и ряда других частиц на кристаллах и по прохождению частиц сквозь вещества. Признаком волнового процесса во всех таких опытах является дифракционная картина распределения микрочастиц в приемных устройствах.

Волны де Бройля имеют специфическую природу, не имеющую аналогии среди волн, изучаемых в классической физике. Физический смысл имеет *квадрат модуля амплитуды волны де Бройля* в данной

точке пространства, который является *мерой вероятности* того, что частица будет обнаружена в этой точке. Дифракционные картины, которые наблюдаются в опытах, являются проявлением статистической закономерности, согласно которой микрочастицы чаще попадают в определенные места в приёмниках, т.е. туда, где интенсивность волны де Бройля имеет наибольшее значение. Микрочастицы никогда не обнаруживаются в тех местах, где, согласно статистической интерпретации, квадрат модуля амплитуды «волны вероятности» обращается в нуль (только в этом моменте прослеживаются элементы проявления традиционного классического принципа детерминизма).

В этой связи весьма существенным моментом является то, что в квантовой механике вероятностные представления (проявление которых объяснялось в рамках классической механики как следствие неполноты информации об объектах и влияния метрологических погрешностей средств измерения) возведены в ранг объективной физической реальности.

Весьма характерным является то обстоятельство, что волновые свойства объектов не проявляются в явном виде у макроскопических тел. Этот факт связан с тем, что длина волны де Бройля для таких тел настолько мала, что обнаружение волновых проявлений оказывается практически невозможным даже с помощью современных экспериментальных методов анализа.

Например, движению объекта массой в один грамм со скоростью 100 м/с ставится в соответствие волна вероятности де Бройля со значением $\lambda = 6 \cdot 10^{-33}$ м.

В то же время для случая движения электрона с той же скоростью длина волны де Бройля составляет $\lambda = 6 \cdot 10^{-6}$ м, что:

- находится в пределах возможностей современных экспериментальных методов (10^{-10} м) определения величины протяженности физических объектов;
- более чем в 10^6 раз превышает «размер» электрона, если исходить из представлений классической физики о размерах индивидуальных атомов.

Следует отметить, что волны де Бройля (характерные, по определению, только для случая свободных частиц) представляют весьма огра-

ниченный практический интерес, поскольку в инженерно-физической практике приходится, как правило, иметь дело с частицами, локализованными в некоторой ограниченной области физического пространства (т.е. частицами, «запертыми» в потенциальном ящике). Сам факт наличия потенциального ящика заведомо ограничивает свободу перемещения частицы и, следовательно, она уже не может считаться свободной от воздействия внешних силовых полей, в роли которых выступают стенки (границы) потенциального ящика. Этот, более сложный и интересный для практических приложений случай поведения частиц будет рассмотрен ниже. Несколько забегаая вперед, отметим, что именно факт локализации микрочастиц в ограниченной области физического пространства приводит к появлению дискретных (квантовых) энергетических состояний у индивидуальных микрочастиц в конденсированных средах.

1.1.3. Волновой вектор плоской монохроматической волны

Плоская монохроматическая волна (ПМВ) для одномерного случая, т.е. при ее распространении вдоль направления оси x , может быть представлена в виде

$$U(x, t) = U_0 \cdot \cos[\omega \cdot (t \pm x/\mathcal{G}) + \delta], \quad (1.6)$$

где \mathcal{G} – фазовая скорость ПМВ, определяемая из условия $(t - x/\mathcal{G}) = \text{const}$; δ – начальная фаза волны при значении времени $t = 0$; U_0 – амплитуда бегущей волны.

Термин «плоская» означает, что фронт распространения волны может быть отображен в виде плоскости, а «монохроматическая» – что волна сохраняет при своем движении неизменную частоту ($\omega = \text{const}$) в пространстве и во времени.

Если направление распространения ПМВ составляет с осями декартовых координат соответственно углы α, β и γ , то выражение (1.6) преобразуется к виду

$$U(x, t) = U_0 \cdot \cos\left[\omega \cdot \left(t \pm \frac{x \cdot \cos \alpha + y \cdot \cos \beta + z \cdot \cos \gamma}{\mathcal{G}}\right) + \delta\right]. \quad (1.7)$$

Введем в рассмотрение понятие о *волновом векторе* \bar{k} , модуль которого равен $|\bar{k}| = 2 \cdot \pi / \lambda$, а его проекции на координатные оси соответственно равны:

$$\begin{aligned} k_x &= \left(2 \cdot \pi / \lambda\right) \cos \alpha; \quad k_y = \left(2 \cdot \pi / \lambda\right) \cos \beta; \\ k_z &= \left(2 \cdot \pi / \lambda\right) \cos \gamma. \end{aligned} \quad (1.8)$$

Используя понятие скалярного произведения волнового вектора \bar{k} и радиуса вектора \bar{r} (проекции которого равны: x , y и z), выражение (1.7) приобретает более лаконичный вид:

$$U(x, t) = U_0 \cdot \cos \left[(\omega \cdot t \pm \bar{k} \cdot \bar{r}) + \delta \right]. \quad (1.9)$$

Как правило, для облегчения последующих математических операций ПМВ часто представляют в комплексном виде, используя для этого формулу Эйлера:

$$e^{i\varphi} = \cos \varphi + i \cdot \sin \varphi. \quad (1.10)$$

В этом случае (на основании формального использования формулы Эйлера) выражение для ПМВ можно представить в комплексной форме:

$$U(x, t) = U_0 \cdot \operatorname{Re} \cdot e^{i[(\omega t \pm \bar{k} \cdot \bar{r}) + \delta]}, \quad (1.11)$$

где фрагмент (Re) свидетельствует о том, что в выражении (1.11) следует учитывать только действительную часть.

На практике, как правило, фрагмент (Re) не принято обозначать, но по умолчанию его присутствие подразумевается:

$$U(x, t) = U_0 \cdot e^{i[(\omega t \pm \bar{k} \cdot \bar{r}) + \delta]}. \quad (1.12)$$

Введя в рассмотрение комплексную амплитуду ПМВ ($U_0^* = U_0 \cdot e^{i[\delta]}$), выражение (1.12) приобретает более лаконичный вид:

$$U(x, t) = U_0^* \cdot e^{i[\omega t \pm \bar{k} \cdot \bar{r}]}. \quad (1.13)$$

Используя представление о волновом векторе ПМВ, выражение (1.3) преобразуется к виду

$$\bar{p} = \frac{h}{2 \cdot \pi} \cdot \bar{k}, \quad (1.14)$$

где $\vec{k} = (2\pi/\lambda) \cdot \vec{n}$ – волновой вектор ПМВ; \vec{n} – единичный вектор в направлении распространения фронта ПМВ.

Для случая свободной частицы (т.е. при отсутствии потенциальных полей) полная энергия частицы всецело определяется ее кинетической составляющей W_k :

$$W_k = \frac{m \cdot v^2}{2} = \frac{h^2}{8 \cdot m \cdot \pi^2} \cdot k^2, \quad (1.15)$$

где k – модуль волнового вектора ПМВ свободной частицы; m – масса частицы; v – скорость движения частицы.

Поскольку в выражении (1.15) на модуль волнового вектора k не накладывается никаких ограничений, то полная энергия свободной частицы W_k может принимать непрерывный ряд значений в диапазоне величин $(0 - \infty)$. Именно из этого обстоятельства следует тот факт, что *энергия свободной частицы не квантуется*.

Для последующего рассмотрения важно отметить, что ПМВ имеет *бесконечную* протяженность в пространстве, что, собственно, и следует из фактора ее монохроматичности. Последнее обстоятельство приводит к парадоксальной ситуации, когда речь идет о вероятностных волнах де Бройля. Согласно разделу (1.1.1.), квадрат модуля амплитуды волны де Бройля в определенной точке пространства является мерой вероятности того, что частица может быть обнаружена в этой точке. В то же время фактор бесконечной протяженности ПМВ в пространстве свидетельствует о том, что квадрат модуля амплитуды волны должен быть тождественно равен нулю. (Правомерность этого утверждения следует из необходимости обеспечения условия нормировки для вероятности нахождения частицы в заданной точке пространства, согласно которому суммарная вероятность нахождения микрочастицы в безграничном пространстве (неважно где) должна быть равна единице.) Следует отметить, что равенство нулю указанных вероятностей для свободной частицы является тривиальным фактом даже с точки зрения здравого смысла, и для установления этого обстоятельства едва ли было бы целесообразным вводить в рассмотрение волны де Бройля. В то же время волны де Бройля могут оказаться небесполезными, если на их основе

удастся «сконструировать» вероятностные представления для случая микрочастиц, движение которых происходит в локализованной области пространства (в потенциальном ящике). Эта задача решается ниже в концепции представлений о волновом пакете, образованном из совокупности волн де Бройля.

1.1.4. Волновой пакет из вероятностных волн де Бройля

Как отмечалось выше, ПМВ представляет собой бегущую волну, которая имеет *бесконечную протяженность* в пространстве. В то же время следует отметить, что даже квант света воспринимается как объект, вполне локализованный в ограниченной области пространства. Из этих соображений следует, что отдельный квант света не может быть представлен в виде единственной ПМВ. Ниже будет показано, что создание пространственно-локализованных объектов (микрочастиц или фотонов) возможно путем суперпозиции бесконечного количества индивидуальных ПМВ. Такие пространственно-локализованные объекты, образованные путем суперпозиции большого количества ПМВ, получили название *волнового пакета* [2].

Создадим прообраз волнового пакета путем *суперпозиции* бесконечного количества ПМВ, у которых волновой вектор k (в целях упрощения анализа рассматривается одномерный случай согласно выражению (1.6)) пробегает непрерывный ряд значений в интервале значений $(k_0 - \Delta k) \div (k_0 + \Delta k)$:

$$U(x, t) = \int_{k_0 - \Delta k}^{k_0 + \Delta k} U_0 \cdot \cos(\varpi \cdot t - k \cdot x) dk, \quad (1.16)$$

где U_0 – постоянная величина.

После интегрирования выражения (1.16) и элементарных математических преобразований имеем аналитическое выражение для волнового пакета в виде

$$U(x, t) = 2 \cdot U_0 \cdot (\Delta k) \cdot \left(\frac{\sin(\Delta k \cdot x)}{(\Delta k \cdot x)} \right) \cdot \cos(\varpi \cdot t - k_0 \cdot x). \quad (1.17)$$

Напомним фрагмент выражения (1.17):

$$\left(\frac{\sin(\Delta k \cdot x)}{\Delta k \cdot x} \right) = \begin{cases} 1 & \text{при условии:} \\ & (\Delta k \cdot x) \rightarrow 0; \\ 0 & \text{при условии:} \\ & (\Delta k \cdot x) = \pm n \cdot \pi. \end{cases} \quad (1.18)$$

В целях наглядности на рис. 1.1 представлены волновые пакеты, построенные для двух случаев: $(\Delta k / k_0) = 0.1$ и $(\Delta k / k_0) = 0.5$. Протяженность центрального фрагмента рисунка $2\Delta x$ (при $x = 0$) условно принимается равной ширине волнового пакета (т.е. области локализации объекта). Из рисунка следует, что увеличение протяженности интервала значений $(k_0 - \Delta k) \div (k_0 + \Delta k)$ приводит к повышению степени локализации объекта в пространстве.

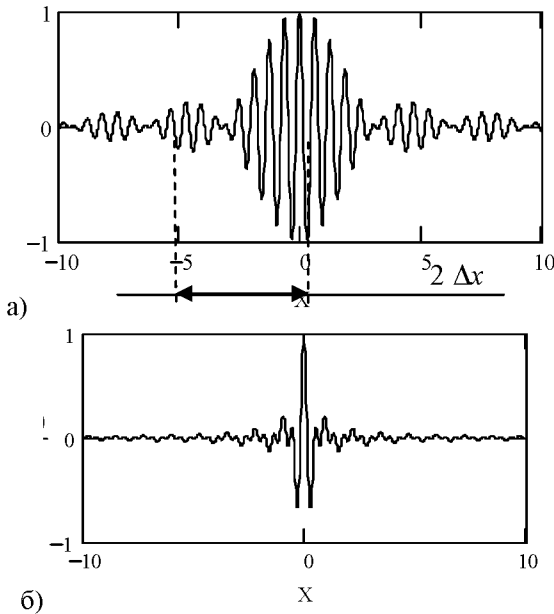


Рис. 1.1. Зависимость пространственной протяженности волнового пакета $2\Delta x$ от относительной величины интервала вариаций Δk :
а) – $(\Delta k / k_0) = 0.1$; б) – $(\Delta k / k_0) = 0.5$

С учетом выражения (1.18) (нижнее условие) для определения пространственной протяженности волнового пакета Δx имеем

$$(\Delta k \cdot \Delta x) = \pi. \quad (1.19)$$

Из выражения (1.19) следует, что при выполнении условия:

- $\Delta k \rightarrow \infty$, протяженность волнового пакета $\Delta x \rightarrow 0$, т.е. в данном случае имеет место реализация абсолютно локализованного в пространстве объекта (например, идеализированного кванта света);
- $\Delta k \rightarrow 0$, протяженность волнового пакета $\Delta x \rightarrow \infty$, т.е. в данном случае имеет место реализация одной единственной бесконечно протяженной ПМВ с волновым вектором, равным k_0 (аналог волны де Бройля для свободной частицы);
- $\Delta k = \text{const}$, протяженность волнового пакета Δx становится равной определенной конечной величине, т.е. в данном случае имеет место реализация цуга волн, отображающего область нахождения реального объекта.

Таким образом, можно считать, что методология формирования волновых пакетов является эффективным способом создания пространственно-локализованных структур (объектов) из бесконечно протяженных составляющих, имеющих волновую природу, вне зависимости от физической специфики рассматриваемых волновых процессов. В силу оговоренной выше универсальности подобного подхода он нашел весьма широкое распространение не только в оптике (электромагнитные волны), но в электронике (спектральный анализ электрических сигналов) и в квантовой механике (вероятностные волны де Бройля). Во всех вышеупомянутых случаях производится «конструирование» локализованных объектов путем суммирования (суперпозиции) большого количества (в пределе бесконечного) индивидуальных ПМВ.

Из выражения (1.14) следует, что интервалу значений Δk соответствует определенный интервал значений Δp (по каждой из координат $\{x, y, z\}$):

$$\Delta p = \left(\frac{h}{2 \cdot \pi} \right) \cdot \Delta k. \quad (1.20)$$

Из выражений (1.19) и (1.20) следуют соотношения, известные как *принцип неопределенностей В. Гейзенберга*:

$$\Delta p_x \cdot \Delta x = h/2; \quad \Delta p_y \Delta y = h/2; \quad \Delta p_z \cdot \Delta z = h/2. \quad (1.21)$$

Принцип неопределенности свидетельствует о *невозможности одновременного и точного* определения координат и импульсов микро-частиц (следует особо отметить, что классическая физика не налагает подобного рода принципиальных запретов), а следовательно, становится проблематичным определение потенциальной (фактор Δx) и кинетической (фактор Δp) энергии микро-частиц.

Выражение (1.21) может быть использовано для проведения грубых оценок (с точностью до порядка величины) кинетической энергии микро-частиц, «запертых» в пределах одномерного потенциального ящика, имеющего протяженность $2 \Delta x$ (см. рис. 1.1). Для этой цели, полагая, что $\Delta p \approx p$, можно провести оценку кинетической энергии частицы согласно выражению:

$$W_k = \frac{p^2}{2 \cdot m} = (\Delta p)^2 / (2 \cdot m) = h^2 / [8 \cdot m \cdot (\Delta x)^2]. \quad (1.22)$$

В табл. 1.1 представлены зависимости абсолютных W_k и относительных W_k / W_T значений кинетической энергии электрона (где W_T – средняя энергия теплового движения электрона) от протяженности потенциального ящика ($2 \cdot \Delta x$), полученные на основании выражения (1.22).

Таблица 1.1. *Зависимость параметров W_k и W_k / W_T от протяженности потенциального ящика ($2 \Delta x$) (масса покоя электрона: $m = 9.11 \cdot 10^{-31}$ кг)*

1	$2 \Delta x$, нм	0.2	0.5	1.0	2.0	5.0	10	20	50	100	
2	W_k , эВ	36	5.8	1.4	0.35	$5 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	$3 \cdot 10^{-3}$	$6 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	
3	$\frac{W_k}{W_T}$	$T=300$ К	1000	200	60	10	2	0.4	0.1	0.02	0.004
		$T=30$ К	1000	2000	600	100	20	4.0	1.0	0.2	0.04
		$T=3$ К	$1 \cdot 10^5$	$2 \cdot 10^4$	6000	1000	200	40	10	2.0	0.4

Примечание: заливкой обозначены ячейки, в которых W_T превосходит W_k более чем на порядок.

Данные табл. 1.1 позволяют в некоторой мере оценить влияние на энергетическое состояние микро-частицы масштабного фактора, свидетельствующего об ограничении свободы перемещения микро-частицы в пространстве. В частности из таблицы видно, что при размерах «потен-

циального ящика» более чем 100 нм влияние фактора ограничения свободы микрочастицы становится пренебрежимо мало относительно возможности проявления физико-химических эффектов, которые могут быть обусловлены проявлением тепловых флуктуаций в системе.

1.1.5. Квантование энергии микрочастицы, находящейся в потенциальном ящике

Как уже отмечалось выше, в квантовой механике движение свободной микрочастицы отображается посредством ПМВ, что свидетельствует об одинаковой вероятности ее нахождения в любой точке физического пространства (и, по сути, эта вероятность тождественно равна нулю – подробнее см. выше).

Движение микрочастицы в потенциальном ящике, с бесконечно высокими стенками, происходит в ограниченной области пространства (частица не имеет возможности выйти за пределы ящика), а следовательно, вероятностное представление ее местонахождения в пространстве будет иметь вид, схематически представленный для случая одномерного движения на рис. 1.2.

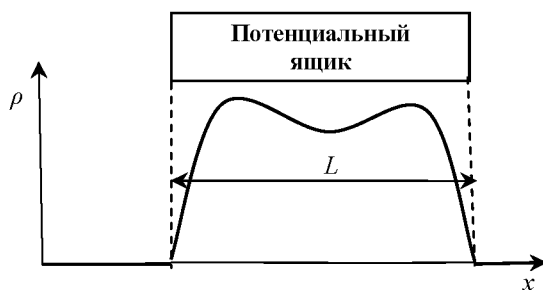


Рис. 1.2. Условное отображение функции плотности распределения вероятностей местонахождения микрочастицы в одномерном потенциальном ящике

Для начального рассмотрения важен не столько сам вид представленной кривой (этот аспект будет рассмотрен позже), а то обстоятельство, что вероятность нахождения микрочастицы непосредственно на границах потенциального ящика должна быть принята равной нулю. Это граничное условие следует из условия непрерывности поведения реальных физических объектов (т.е. при приближении к границе ящика

как справа, так и слева должна наблюдаться одинаковая вероятность местонахождения микрочастицы). Отсюда следует, что движению микрочастицы должны соответствовать не любые ПМВ, а лишь те из них, которые могут обеспечить выполнение этого граничного условия.

На рис. 1.3 представлены четыре (из бесконечной совокупности) варианта таких «правильных» ПМВ, суперпозиция которых образует волновой пакет, определяющий вид кривой на рис. 1.3.

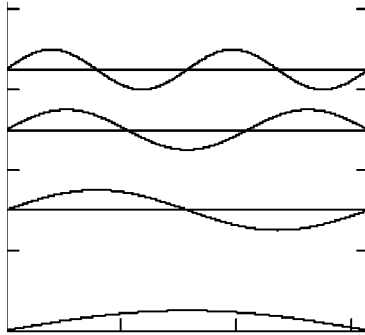


Рис.1.3. Иллюстрация четырех ПМВ ($n = 1, 2, 3, 4$), участвующих в образовании волнового пакета

Согласно волновому представлению, изображенному на рис. 1.3, к числу вышеуказанных «правильных» волн следует отнести только лишь те из ПМВ, для которых выполняется условие: $\lambda_n = (2 \cdot L) / n$, где параметр n пробегает бесконечный ряд значений из набора величин $\{1, 2, 3, \dots, \infty\}$.

Принимая во внимание результаты раздела 1.1.3, можно заключить, что и модуль волнового вектора волн де Бройля тоже может принимать бесконечный ряд дискретных значений, равных: $k_n = (\pi \cdot n) / L$. С учетом выражения (1.15) и вышеизложенных моментов относительно допустимых значений волнового вектора ПМВ можно каждому значению модуля вектора k_n поставить в соответствие вполне определенную волну де Бройля со значением энергии микрочастицы, равным:

$$W_n = \frac{h^2 \cdot n^2}{8 \cdot m \cdot L^2}. \quad (1.23)$$

Из выражения (1.23) следует фундаментальный вывод о том, что энергия микрочастицы, находящейся в потенциальной яме, может принимать только дискретный ряд значений, т.е. квантуется. Параметр n , характеризующий возможное значение энергии микрочастицы, получил название *квантового числа энергетического состояния* микрочастицы.

Определим относительное изменение энергии микрочастицы, представленное в виде отношения $W_{(n+1)}/W_{(n)}$, при ее переходе с одного квантового уровня (n) на соседний уровень ($n+1$). Согласно выражению (1.23) для искомого отношения имеем

$$\frac{W_{(n+1)}}{W_{(n)}} = \frac{(n+1)^2}{n^2} \rightarrow 1 \quad \text{при } n \rightarrow \infty. \quad (1.24)$$

Из выражения (1.24) следует, что наибольшие относительные изменения энергии при квантовых переходах имеют место при малых значениях квантовых чисел n и абсолютно не проявляются на практике при значениях $n \rightarrow \infty$. В последнем случае принято считать, что энергетический спектр микрочастицы при больших значениях параметра n имеет непрерывный (а не дискретный) характер.

Можно отметить также то обстоятельство, что в случае выполнения вышеуказанного условия ($n \rightarrow \infty$) количество максимумов (на графиках рис. 1.3) будет крайне велико и располагаться они между собой будут очень плотно. Отсюда можно сделать вывод, что вероятность местоположения микрочастицы будет практически постоянна в любой точке потенциального ящика, как это и принято считать в классической физике, для случая движения шарика по абсолютно ровной и горизонтальной поверхности.

Для последующего рассмотрения необходимо уточнить, что в выражении (1.23) в силу ограничений, накладываемых принципом неопределенности (см. раздел 1.1.4.), речь может идти только о полной энергии микрочастицы. О потенциальной энергии можно говорить только в контексте налагаемого на рассматриваемую систему в целом внешнего (относительно микрочастицы) силового поля, в условиях ко-

того происходит движение микрочастицы. Кинетическая энергия микрочастицы (определяемая, как это общепринято считать в классической физике, в виде разницы полной и потенциальной энергий частицы) в квантовой физике не является параметром, доступным процедуре непосредственного измерения, а следовательно, и не является физическим параметром системы.

В более общей трактовке можно сформулировать фундаментальный вывод о том, что координаты (x, y, z) и проекции импульса (p_x, p_y, p_z) микрочастицы уже не могут рассматриваться в качестве основных базовых параметров, однозначно характеризующих физическое состояние рассматриваемой системы. Это обстоятельство можно наглядно отобразить с помощью фазового портрета состояния микрочастицы.

На рис. 1.4 представлены фазовые портреты состояния индивидуальной частицы для случаев классического и квантового подходов (на рисунке 6-мерная система координат фазового пространства состояния системы условно представлена как трехмерная). В классическом подходе состояние микрочастицы отображается на фазовом портрете посредством одной единственной точки, с проекциями на оси координат – x, y, z, p_x, p_y, p_z .

В квантовом подходе следует учесть то обстоятельство, что в соответствии с выражениями (1.21) состояние микрочастицы в фазовом пространстве определяется с некоторой долей неопределенности, величина которой связана с наличием элементарного фазового объема ΔV (т.е. это уже не точка):

$$\Delta V = \Delta x \cdot \Delta y \cdot \Delta z \cdot \Delta p_x \cdot \Delta p_y \cdot \Delta p_z = h^3. \quad (1.25)$$

На рис. 1.4 положения точки **A** (классический подход) и объема ΔV_{t_0} (квантовый подход) характеризуют состояние микрочастицы в фазовом пространстве в начальный момент времени t_0 . С течением времени микрочастица перемещается в физическом пространстве и в момент времени t_1 фазовый портрет состояния микрочастицы отображается уже точкой **B** (классический подход) и объемом ΔV_{t_1} (квантовый подход).

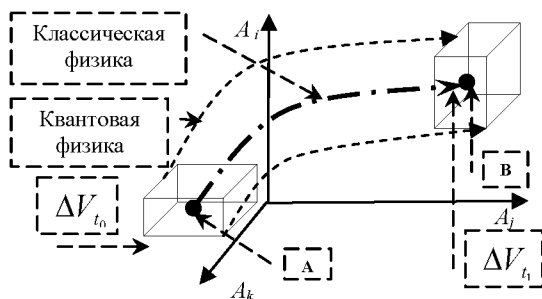


Рис. 1.4. Схематическое отображение фазовых траекторий движения микрочастицы в классическом и квантовом подходах

Процесс движения микрочастицы при классическом подходе отображается определенной фазовой траекторией (что условно обозначено толстой штрихпунктирной линией), что свидетельствует о жестко детерминированном характере процесса эволюции системы. В случае использования квантового подхода как начальное, так и конечное состояние движущейся микрочастицы содержит элемент неопределенности. В этой связи в последнем случае конечный результат процесса эволюции системы может быть спрогнозирован только на основе вероятностных представлений, представленных выше.

Из вышеизложенного следует, что детерминизм квантовой механики (в части своеобразного проявления причинно-следственных связей в физической системе) имеет своеобразный характер и исключает возможность формирования точного прогноза относительно состояния системы в будущем. Причем эта неопределенность состояния системы испытывает тенденцию к непрерывному возрастанию во времени.

1.1.6. Волновая функция как характеристика состояния микрочастицы

Основной задачей разд. 1.1.5 являлось рассмотрение феномена эффекта квантования энергии микрочастицы, локализованной в ограниченной области пространства (т.е. «запертой» в потенциальном ящике). При этом остался открытым вопрос, где именно находится частица в пределах упомянутого потенциального ящика.

Строгое рассмотрение этого вопроса предполагает использование понятия о *волновых функциях* микрочастицы [1]. Отметим, что волновые представления уже использовались ранее в вышеизложенных материалах в виде элементарных ПМВ де Бройля. Поэтому все нижеизложенные материалы можно рассматривать как дальнейшее развитие волновых представлений квантовой механики.

Волновая функция $\Psi(x, y, z, t)$ состояния микрочастицы формируется таким образом, чтобы квадрат её модуля $|\Psi(x, y, z, t)|^2$ представлял собой функцию плотности вероятности распределения микрочастицы ϖ . Для случая дискретных возможностей расположения частицы – это просто вероятность обнаружить микрочастицу в момент времени t_0 в точке пространства с координатами (x_0, y_0, z_0) :

$$\varpi = |\Psi(x_0, y_0, z_0, t_0)|^2 = \Psi \cdot \Psi^*, \quad (1.26)$$

где Ψ^* – комплексно-сопряженная с Ψ волновая функция.

Можно отметить значительную схожесть трактовки физического смысла волновой функции Ψ и ПМВ де Бройля (см. раздел 1.1.3).

Представляется очевидным, что волновая функция должна удовлетворять условию нормировки:

$$\iiint \Psi(x, y, z) \cdot \Psi^*(x, y, z) \cdot dx \cdot dy \cdot dz = 1, \quad (1.27)$$

где интегрирование производится по всему геометрическому пространству.

Равенство интеграла (1.27) единице отражает тот очевидный факт, что событие, состоящее в том, что микрочастица будет обнаружена в какой-либо точке геометрического пространства (при этом неважно, в какой именно точке пространства), является достоверным.

Целесообразность введения в рассмотрение функции плотности вероятности ϖ обусловлена появившейся возможностью определения вероятности нахождения микрочастицы в ограниченной области пространства (в частности, в пределах потенциального ящика). Действительно, из определения понятия функции плотности вероятности сле-

дует, что вероятность нахождения микрочастицы dP в бесконечно малом объеме пространства $dV = dx \cdot dy \cdot dz$ равна

$$dP = \varpi \cdot dV. \quad (1.28)$$

С учетом выражения (1.28), для вероятности нахождения микрочастицы $P(V)$ в любой области пространства конечного объема V имеем (в различных формах представления):

$$P(V) = \int_V dP \equiv \int_V \varpi \cdot dV \equiv \int (\Psi \cdot \Psi^*) \cdot dV. \quad (1.29)$$

Для полноты рассмотрения вопроса отметим, что набор координат, которые выступают в роли аргументов волновой функции в выражении (1.29), принято называть в квантовой механике *полным набором физических величин*, характеризующих состояние системы в соответствии со спецификой рассматриваемой задачи. В квантовой механике можно выбрать различные варианты полных наборов величин. Выбранный полный набор величин определяет *представление волновой функции*. Наибольшее распространение в инженерно-физической практике получило *координатное* представление ($dV = dx \cdot dy \cdot dz$) и *импульсное* представление ($dV = dp_x \cdot dp_y \cdot dp_z$). В первом случае определяется местоположение микрочастицы в обычном геометрическом пространстве, а во втором – в виртуальном импульсном пространстве.

Для волновых функций справедлив *принцип суперпозиции*, который уже был использован (в неявном виде) в разделе 1.1.5 при формировании волнового пакета из ПМВ де Бройля. Физическая суть принципа суперпозиции состоит в том, что если система может находиться в состояниях, описываемых волновыми функциями из набора индивидуальных функций $\{\Psi_1, \Psi_2, \dots, \Psi_n\}$, то она может пребывать и в обобщенном состоянии, описываемом волновой функцией Ψ_Σ , определяемой в виде

$$\Psi_\Sigma = \sum_{i=1}^n c_i \cdot \Psi_i, \quad (1.30)$$

где c_i – постоянные коэффициенты (в общем случае комплексные), удовлетворяющие условию нормировки $\sum |c_i|^2 = 1$.

Именно совокупное проявление индивидуальных функций $\{\Psi_1, \Psi_2, \dots, \Psi_n\}$ (как результат их суперпозиции в виде волновой функции – Ψ_Σ) определяет конкретный вид распределения вероятности нахождения микрочастицы в пределах потенциального ящика (см. рис. 1.2).

Из наличия у частиц волновых проявлений следует, что закон их движения также должен основываться на волновых принципах. Отсюда следует, что при анализе движения микрочастицы следует руководствоваться неким *волновым уравнением*. Для микрочастицы, движущейся в силовом поле с потенциальной энергией $U(x, y, z, t)$, это волновое уравнение имеет следующий вид (уравнение Шредингера, имеющее статус закона природы [2]):

$$-i \cdot \frac{h}{2\pi} \cdot \frac{\partial \Psi}{\partial t} = \frac{h^2}{8 \cdot \pi^2 \cdot m} \cdot \left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} \right) - U(x, y, z, t) \cdot \Psi, \quad (1.31)$$

где Ψ – рассмотренная выше волновая функция микрочастицы; m – масса микрочастицы; i – символ мнимой единицы, равный $\sqrt{-1}$.

Важность волнового уравнения (1.31) состоит в том, что оно отражает процесс эволюции состояний микрочастицы во времени. Следует отметить, что решением уравнения Шредингера являются только *комплексные* волновые функции. Это обстоятельство не должно вызывать недоразумение, поскольку физическим смыслом обладает только плотность вероятности распределения частиц ϖ (см. выражение (1.26)), а она является действительным числом.

Как правило, аналитическое решение уравнения Шредингера для случая реальных потенциальных функций $U(x, y, z, t)$ связано со значительными математическими трудностями. В этой связи в инженерно-физической практике часто, в целях упрощения анализа, рассматриваются:

- идеализированные модели функции потенциальной энергии $U(x, y, z, t)$, к примеру, рассмотренная выше модель потенциального ящика;

- стационарные задачи (интересные для многих случаев практического применения), в которых функция потенциальной энергии не зависит от времени, т.е. рассматривается функция – $U(x, y, z)$.

Рассмотрим общие принципы решения волнового уравнения Шредингера для *стационарного случая* на примере *одномерного* (во избежание громоздких выражений) движения микрочастицы. Условие стационарности позволяет представить волновую функцию $\Psi(x, y, z, t)$ в виде произведения независимых координатных и временных функций:

$$\Psi(x, y, z, t) = \phi(x, y, z) \cdot \varphi(t). \quad (1.32)$$

Подставив (1.32) в выражение (1.31) и разделив полученное выражение на фрагмент $\{\phi(x, y, z) \cdot \varphi(t)\}$, получим

$$\frac{\hbar^2}{8 \cdot \pi^2 \cdot m} \cdot \frac{1}{\phi(x)} \cdot \frac{d^2 \phi(x)}{dx^2} - U(x) = -i \cdot \frac{\hbar}{2 \cdot \pi} \cdot \frac{1}{\varphi(t)} \cdot \frac{d\varphi(t)}{dt}. \quad (1.33)$$

Выражение (1.33) интересно тем, что его левая часть является функцией только координаты местонахождения микрочастицы x , а правая – только времени t . Такая ситуация возможна только в том случае, если обе эти части равны одной и той же постоянной величине. Отметим также, что эта постоянная величина имеет размерность энергии. Можно показать [2], что эта постоянная равна $(-W)$, где W – полная энергия микрочастицы (замечания относительно возможности совместного определения потенциальной и кинетической энергии частицы приведены в разделе 1.1.5).

Приравняв правую и левую часть выражения (1.33) $(-W)$ и проведя необходимые алгебраические преобразования, получим два уравнения:

$$\frac{d^2 \phi(x)}{dx^2} + \frac{8 \cdot \pi^2 \cdot m}{\hbar^2} \cdot (W - U(x)) \cdot \phi(x) = 0, \quad (1.34a)$$

$$\frac{d\varphi(t)}{dt} = -i \cdot \frac{2 \cdot \pi}{\hbar} \cdot \varphi(t) \cdot W. \quad (1.34б)$$

Функция $\phi(x)$, или в общем случае $\phi(x, y, z)$, получила название *амплитудной* волновой функции. Как будет показано ниже (на примерах определения энергетических состояний для случая несвободных частиц), амплитудное уравнение (1.34а) имеет решение только при строго определенных значениях энергии микрочастицы $\{W_1, W_2, W_3, \dots, W_n\}$, которые называются *собственными значениями энергии* частицы. Волновые функции $\{\phi_1, \phi_2, \phi_3, \dots, \phi_n\}$, отвечающие этим значениям энергии, называются *собственными волновыми функциями*.

Решение уравнения (1.34а) не зависит в явном виде от функции $U(x)$ и находится посредством интегрирования по методу разделения переменных:

$$\varphi_n(t) = e^{-2 \cdot \pi \cdot i \cdot \frac{W_n \cdot t}{h}}, \quad (1.35)$$

где W_n – одно из собственных значений энергии микрочастицы.

Из выражения (1.35) следует, что временная зависимость имеет гармонический характер с частотой колебаний $\nu_n = W_n/h$ или $\omega_n = (2 \cdot \pi \cdot W_n)/h$ (обратить внимание на сходство с выражением (1.2)).

Таким образом, для рассматриваемого стационарного случая *полная волновая функция* $\Psi_n(x, t)$ для квантового состояния с номером n может быть представлена в виде

$$\Psi_n(x, y, z, t) = \phi_n(x, y, z) \cdot \varphi_n(t) = \phi_n(x, y, z) \cdot e^{-i \omega_n t}. \quad (1.36)$$

С учетом вышеизложенного вероятность обнаружения микрочастицы dP в элементе объема dV будет равна

$$dP = \phi(x, y, z) \cdot \phi^*(x, y, z) \cdot dV, \quad (1.37)$$

где совокупность координат $\{x, y, z\}$ характеризует расположение элементарного объема dV в геометрическом пространстве. Эта веро-

ятность dP для рассмотренного стационарного случая не зависит от времени.

1.1.7. Примеры квантования энергии микрочастиц в физических системах

а) Идеализация потенциальных атомных диаграмм

В настоящем разделе приведены примеры решения уравнения Шредингера для ряда конкретных физических систем, широко используемых в современной инженерно-физической практике. Ограниченный объем пособия не предполагает проведения детального рассмотрения данных материалов. В этой связи основной целью раздела является изучение основных методологических принципов анализа квантовых *стационарных* систем на примерах движения микрочастиц в условиях [8]:

- отсутствия потенциальных полей как таковых (специальный случай, адекватный вероятностным волнам де Бройля);
- нахождения в потенциальном ящике конечной протяженности с бесконечно высокими стенками;
- электростатического механизма взаимодействия микрочастиц (электрона с протоном в водородоподобном атоме);

Как отмечалось выше, практическое использование уравнения Шредингера сопряжено со значительными математическими трудностями. Эффективность процесса минимизации указанных математических трудностей напрямую связана с решением вопроса «оптимизации» используемых потенциальных функций $U(x, y, z)$. Под оптимизацией принято понимать процедуру замены реальных потенциальных функций взаимодействия микрочастиц на их упрощенные (идеализированные) модели, которые, тем не менее, адекватны целям рассматриваемой задачи.

В инженерно-физической практике анализа микроскопических (атомно-молекулярных) и макроскопических систем (конденсированные среды) в качестве базовой потенциальной функции $U(x, y, z)$ часто используется выражение вида [3]

$$U(r) = 4 \cdot \varepsilon \cdot \left[\left(\frac{\sigma}{r} \right)^n - \left(\frac{\sigma}{r} \right)^m \right], \quad (1.38)$$

где r – межцентровое расстояние микрочастиц; n и m – постоянные величины, численное значение которых определяется типом механизма взаимодействия микрочастиц (выполняется условие: $n > m$); ε и σ – постоянные, определяемые экспериментальным путем для каждого типа взаимодействующих микрочастиц (в частности, для случая атомно-молекулярных взаимодействий (*потенциал Леннарда-Джонса* [3]): $n = 12$; $m = 6$).

Первое слагаемое в выражении (1.38) характеризует проявление сил взаимного отталкивания между микрочастицами, второе – сил взаимного притяжения. На рис. 1.5 представлены примеры графического отображения процедуры проведения идеализации (варианты Б) и (В)) потенциальной функции $U(x, y, z)$ вида (1.38).

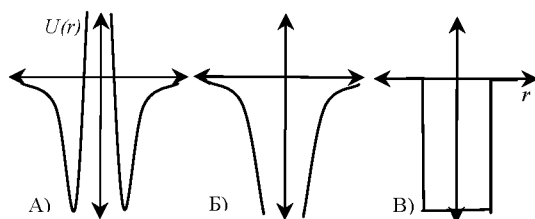


Рис. 1.5. Варианты представления потенциальной функции $U(x, y, z)$
(А – реальная функция, по (1.38); Б – идеализация № 1;
В – идеализация № 2)

Из материалов рис.1.5 следует, что на некоторых участках оси r функция $U(x, y, z)$ принимает отрицательные значения. Последнее обстоятельство является следствием наложенного требования о равенстве $U(x, y, z) = 0$ при $r \rightarrow \infty$. Факт появления отрицательных значений энергии не должен вызывать особого недоразумения, поскольку в расчетной практике представляют интерес не абсолютные значения энергий (уровни квантования энергии), а их разностные формы, определяющие энергетические параметры конкретных элементарных физико-химических процессов.

б) Движение микрочастицы в условиях отсутствия
потенциальных полей

В отсутствие потенциального поля (т.е. при условии: $U(x) = 0$) движение микрочастицы следует рассматривать как абсолютно свободное и стационарное. Уравнение Шредингера для этого типа движения микрочастицы в одномерном приближении имеет вид

$$-i \cdot \frac{h}{2\pi} \cdot \frac{\partial \Psi(x,t)}{\partial t} = \frac{h^2}{8 \cdot \pi^2 \cdot m} \cdot \left(\frac{\partial^2 \Psi(x,t)}{\partial x^2} \right). \quad (1.39)$$

С учетом стационарного характера движения микрочастицы решением этого уравнения является полная волновая функция (см. выражение (1.36)):

$$\Psi(x,t) = \phi(x) \cdot e^{-i\varpi t}, \quad (1.40)$$

где $\phi(x)$ – амплитудная волновая функция, удовлетворяющая амплитудному уравнению (1.34).

При выполнении условия $U(x, y, z) = 0$ для амплитудного волнового уравнения $\phi(x)$ имеем

$$\frac{d^2 \phi(x)}{dx^2} + \frac{8 \cdot \pi^2 \cdot m}{h^2} \cdot W \cdot \phi(x) = 0. \quad (1.41)$$

Поскольку в данном случае полная энергия микрочастицы равна ее кинетической энергии, то с учетом выражения (1.15) для амплитудного уравнения имеем

$$\frac{d^2 \phi(x)}{dx^2} + k^2 \cdot \phi(x) = 0. \quad (1.42)$$

Частными решениями амплитудного уравнения (1.42) являются:

$$\phi_1(x) = A \cdot e^{+ikx} \text{ и } \phi_2(x) = B \cdot e^{-ikx}, \quad (1.43)$$

где параметры, стоящие перед экспонентами, A и B – постоянные коэффициенты.

С учетом выражения (1.40), решая уравнение (1.39), для полной волновой функции $\Psi(x,t)$ имеем

$$\Psi(x,t) = A \cdot e^{i(k \cdot x - \varpi t)} + B \cdot e^{-i(k \cdot x - \varpi t)}. \quad (1.44)$$

Уравнение (1.44) отражает суперпозицию двух ПМВ де Бройля (как это и следовало ожидать, согласно материалам раздела 1.1.2), распространяющихся в противоположных направлениях вдоль оси x геометрического пространства.

Для последующего изложения необходимо отметить, что энергия движения свободной микрочастицы (то есть при выполнении условия: $U(x, y, z, t) = 0$) не квантуется (поскольку нет никаких ограничений на допустимые значения модуля волнового вектора $-k$), ее изменение может иметь непрерывный характер, а ее величина пропорциональна квадрату модуля волнового вектора (подробнее см. выражение (1.15)).

в) Квантование энергии микрочастицы, находящейся в потенциальном ящике

Определим энергетический спектр микрочастицы при ее движении в пределах потенциального ящика (используется модельное представление согласно рис. 1.5,В). Для этого рассмотрим случай движения микрочастицы в условиях внешнего силового поля, определяемого потенциальной функцией $U(x)$ вида

$$U(x) = \begin{cases} U_0, & \text{при } x < 0 \text{ – (область № I);} \\ 0, & \text{при } 0 \leq x \leq L \text{ – (область № II);} \\ U_0, & \text{при } x > 0 \text{ – (область № III).} \end{cases} \quad (1.45)$$

Уравнения Шредингера (см. (1.34а)) для вышеуказанных областей оси x имеют соответственно вид

$$\frac{d^2 \phi_{I,III}(x)}{dx^2} + \frac{8 \cdot \pi^2 \cdot m}{h^2} \cdot (W - U_0) \cdot \phi_{I,III}(x) \equiv \frac{d^2 \phi_{I,III}(x)}{dx^2} + k_1^2 \cdot \phi_{I,III}(x) = 0, \quad (1.46)$$

где $k_1 = \frac{2 \cdot \pi}{h} \cdot \sqrt{2 \cdot m \cdot (W - U_0)}$;

$$\frac{d^2 \phi_{II}(x)}{dx^2} + \frac{8 \cdot \pi^2 \cdot m}{h^2} \cdot W \cdot \phi_{II}(x) \equiv \frac{d^2 \phi_{II}(x)}{dx^2} + k^2 \cdot \phi_{II}(x) = 0, \quad (1.47)$$

где $k = \frac{2 \cdot \pi}{h} \cdot \sqrt{2 \cdot m \cdot W}$.

Для упрощения рассмотрения задачи положим, что стенки потенциального ящика являются бесконечно высокими, а следовательно, решение уравнения (1.46) имеет вид

$$\phi_I(x) = \phi_{III}(x) \equiv 0. \quad (1.48)$$

Общим решением уравнения (1.47), характерного для случая нахождения микрочастицы в пределах потенциального ящика, является:

$$\phi_{II}(x) = A \cdot e^{+i \cdot k \cdot x} + B \cdot e^{-i \cdot k \cdot x}. \quad (1.49)$$

Воспользовавшись условием непрерывности волновой функции $\phi(x)$ на левой границе ($x = 0$) потенциального ящика, имеем

$$\phi_I(0) = \phi_{II}(0) = 0. \quad (1.50)$$

Из выражения (1.50) следует, что постоянные коэффициенты A и B , используемые в выражении (1.49), удовлетворяют условию: $A = -B$. С учетом этого обстоятельства для волновой функции $\phi_{II}(x)$ находим:

$$\phi_{II}(x) = A \cdot (e^{+i \cdot k \cdot x} - e^{-i \cdot k \cdot x}) = C_1 \cdot \sin(k \cdot x). \quad (1.51)$$

Воспользовавшись условием непрерывности волновой функции $\phi(x)$ на правой границе ($x = L$) потенциального ящика, имеем

$$\phi_{III}(L) = \phi_{II}(L) = 0. \quad (1.52)$$

Из выражений (1.51) и (1.52) следует весьма существенное следствие:

$$C_1 \cdot \sin(k \cdot L) = 0, \quad (1.53)$$

что представляется возможным (если $C_1 \neq 0$) только лишь в том случае, если выполняется условие: $(k_n \cdot L) = n \cdot \pi$, где $n = 1, 2, 3, \dots, \infty$.

С учетом этого условия для амплитудной волновой функции имеем

$$\phi_n(x) = C_n \cdot \sin\left(n \cdot \frac{\pi \cdot x}{L}\right), \quad (1.54)$$

где значения постоянных коэффициентов C_n определяются из условия нормировки амплитудной волновой функции.

Искомое энергетическое состояние микрочастицы, описываемое волновой функцией (1.54), можно получить, подставив дискретные величины k_n в выражение (1.15):

$$W_k = \frac{h^2}{8 \cdot m \cdot L^2} \cdot n^2. \quad (1.55)$$

Из выражения (1.55) следует, что при больших значениях L , намного превосходящих атомные масштабы, расстояние между соседними энергетическими уровнями может оказаться столь незначительным, что позволит считать функцию $W_n = \varphi(n)$ непрерывной (случай свободной микрочастицы).

г) Квантование энергии электронов в водородоподобном атоме

В водородоподобном атоме вокруг ядра с электрическим зарядом, равным $Z \cdot e$ (где Z – порядковый номер химического элемента; e – заряд электрона), движется всего один – единственный электрон (предполагается, что отсутствие остальных $(Z - 1)$ электронов в атоме обусловлено проявлением процессов ионизации). В этом случае, потенциальная энергия электростатического взаимодействия этого электрона с зарядом ядра равна (используется модельное представление согласно рис. 1.5,Б):

$$U(r) = -\frac{Z \cdot e^2}{4 \cdot \pi \cdot \epsilon_0 \cdot r}, \quad (1.56)$$

где r – расстояние между геометрическими центрами ядра и электрона; ϵ_0 – электрическая постоянная, численно равная $8.854 \cdot 10^{-12}$ Ф/м.

Поскольку рассматриваемая зарядовая система обладает центральной симметрией, уравнение Шредингера необходимо представить в сферической системе координат:

$$\Delta \phi(r) + \frac{8 \cdot m \cdot \pi^2}{h^2} \cdot \left(W + \frac{Z \cdot e^2}{4 \cdot \pi \cdot \epsilon_0 \cdot r} \right) \cdot \phi(r) = 0, \quad (1.57)$$

где Δ – оператор Лапласа, представленный в сферической системе координат.

Решение уравнения (1.57) приводит к спектру разрешенных энергетических состояний для электрона в атоме в виде (подробный вывод рассмотрен в источнике [7]):

$$W_n = -\frac{Z^2 \cdot m \cdot e^4}{8 \cdot \varepsilon_0^2 \cdot h^2} \cdot \frac{1}{n^2}. \quad (1.58)$$

Весьма характерно то обстоятельство, что фрагмент n^2 в выражении (1.58) находится в знаменателе (в выражении (1.55) он расположен в числителе). Это расхождение обусловлено тем обстоятельством, что в рассматриваемом случае не используется представление о бесконечной глубине потенциальной ямы, в которой находится микрочастица.

Решение в виде (1.58) является достаточно грубым приближением, поскольку в нем не учитывается фактор орбитального движения электрона, приводящего к нарушению сферической симметрии у функции плотности вероятности расположения электрона (обратить внимание на тот факт, что потенциальная функция (1.56) задана в симметричной форме относительно параметра r). Приемлемая точность выражения (1.58) имеет место только при значениях параметра $n > 100$.

Несмотря на оговоренные выше недостатки, выражение (1.58) нашло широкое распространение в научно-технической практике, поскольку оно способствует пониманию (на качественном уровне рассмотрения) физических принципов эффекта квантования энергии электронов в центрально-симметричных системах, в качестве которых выступают индивидуальные атомно-молекулярные объекты.

Рассмотренные выше примеры практического использования уравнения Шредингера позволяют понять основные методологические принципы квантово-механического анализа энергетического спектра состояний микрочастиц.

Следует отметить, что метод анализа состояния систем, основанный на использовании уравнения Шредингера, не является единственным в теоретическом арсенале квантовой механики. В частности, широкое распространение получил *метод вторичного квантования* [4], основанный на использовании (в качестве базовой концепции) операторов рождения и уничтожения элементарных квантовых состояний. Суть

метода вторичного квантования состоит в том (подробное рассмотрение выходит за рамки настоящего пособия), что в нем, вместо волновых функций микрочастиц в координатном или в импульсном представлении (см. разд. 1.1.5) вводится в рассмотрение новый класс волновых функций в представлении чисел заполнения различных квантовых состояний одной частицы. При этом переходы между различными состояниями одной частицы описываются как уменьшение числа заполнения, соответствующего одной волновой функции, на единицу, и соответственного увеличения числа заполнения другого состояния на единицу.

Несомненным достоинством метода вторичного квантования является то, что он позволяет упростить процедуру анализа физико-технических функциональных (в плане практических интересов) систем, содержащих большое количество индивидуальных микрочастиц, используемых при формировании твердотельных микро- и наноразмерных атомно-молекулярных структур.

1.2. Квантово-механические принципы анализа параметров электронных состояний в наноразмерных структурах

1.2.1. Общие принципы классификации наноразмерных структур

В настоящее время существует чрезвычайно большое количество физических и биологических объектов, которые, исходя из их размерных характеристик, могут быть отнесены к наноразмерным структурам. В рамках настоящего пособия рассматривается довольно ограниченный класс наноразмерных структур, а именно структур с упорядоченным расположением атомов (их ближайшим аналогом в макромире являются кристаллы с электронной проводимостью). В части практических интересов и целевых установок основные результаты этого раздела более всего относятся к области создания и анализа потенциальных возможностей наноразмерных устройств электронной техники. Тем не менее значительный объем рассмотренных ниже материалов может оказаться полезным и при проведении анализа свойств неупорядоченных (например: веществ с атомной структурой аморфного типа) или квазиупорядоченных (т.е. структур, обладающих хотя бы некоторыми эле-

ментами симметрии в расположении атомных группировок, что характерно для значительного круга биологических объектов) наноструктур.

В современной научно-технической практике принято (с некоторой долей условности) подразделять наноразмерные структуры на три класса по критерию пространственной конфигурации [5]:

- **квантовая точка** – пространственная, нуль-мерная структура (условное обозначение – (0D)), в которой движение носителей зарядов (электронов) *ограничено* по всем *трем* геометрическим координатам;
- **квантовая проволока** – пространственная, одномерная структура (условное обозначение – (1D)), в которой движение носителей зарядов (электронов) *ограничено* по *двум* геометрическим координатам, которые направлены по нормали по отношению к направлению раскрытия проволоки;
- **квантовая пленка** – пространственная, двумерная структура (условное обозначение – (2D)), в которой движение носителей зарядов (электронов) *ограничено* по *одной* координате, направленной по нормали к плоскости поверхности структуры.

Относительно приведенных выше определений, использованных при классификации наноразмерных структур, необходимо сделать ряд замечаний общего и частного характера:

а) под термином «*пространственная структура*» следует понимать тот факт, что в общем случае характеристические размеры наноразмерной структуры (по всем координатам или по некоторым из них) могут по порядку величин быть более чем атомного (характерный размер – 1 ангстрем) масштаба;

б) под словосочетанием «*движение ограничено*» следует понимать то обстоятельство, что для характеристических размеров структуры должно быть приемлемым условие квантования энергии микрочастиц, представленное в виде выражения (1.23) (последнее означает, что параметр L в выражении (1.23) не должен быть настолько большим, чтобы энергетический спектр микрочастиц можно было бы считать непрерывным);

в) условные обозначения $\{(0D), (1D), (2D)\}$ относятся к структурам, графическое отображение которых представлено на рис. 1.6;

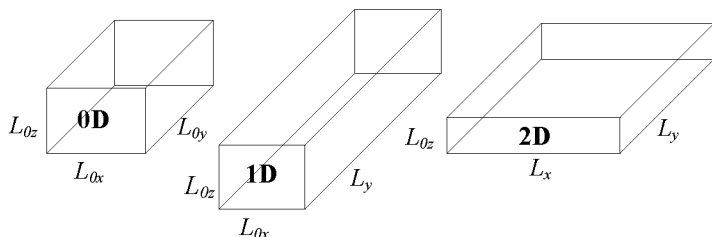


Рис. 1.6 Графическое отображение структур различной размерности: $\{(0D), (1D), (2D)\}$ (индекс «0» при L относится к направлениям с ограниченным движением)

г) в более общей постановке, речь может идти о факторах ограничения движения микрочастиц других типов (помимо электронов), обусловленных проявлением эффекта квантования энергетических состояний микрочастиц в наноразмерных структурах (с позиции представлений, развитых в разделе № 1, т.е. в потенциальном ящике, имеющем конечную протяженность в соответствии с пунктом (б) настоящих замечаний);

д) на вид функциональной зависимости потенциальной энергии от координат в пределах потенциального ящика (в котором находится микрочастица) не накладывается особых ограничений, однако в последующих материалах предполагается, что эта функция отслеживает фактор упорядоченного расположения атомов, характерного для монокристаллических структур.

Резюмируя вышеизложенное, отметим, что целью дальнейшего рассмотрения является анализ влияния квантовых эффектов на характер движения микрочастиц в пространственных структурах, ограниченных (хотя бы по одной из координат) потенциальными барьерами. В конкретном плане подлежит рассмотрению следующий круг вопросов:

- специфика спектров энергетических состояний микрочастиц, находящихся в пределах пространственно упорядоченных наноразмерных атомных структур, с геометрическими параметрами систем типа: $\{(0D), (1D), (2D)\}$;

- специфика процессов транспорта носителей заряда в твердотельных структурах, с характеристическими размерами, предполагающими учет факторов размерного и квантово-механического характера.

Указанные моменты играют существенную роль в вопросах создания перспективных электронных и электронно-оптических вычислительных сред (на основе традиционных планарных технологий производства интегральных схем), обладающих чрезвычайно высокой степенью интеграции элементов ($10^{13} - 10^{15}$ элементов/см³) и уровнем быстродействия ($10^{10} - 10^{12}$ Гц).

Освоение оговоренных практических рубежей позволит вплотную приблизить теоретическую и конструкторско-технологическую базу к постановке задач по созданию квантовых компьютеров, которые, благодаря специфике реализации вычислительных операций (этот аспект рассмотрен в разд. 1.3), позволят обеспечить дальнейшее повышение тактико-технических данных вычислительных средств.

1.2.2. Энергетические состояния микрочастиц в наноразмерных структурах

а) Зонные энергетические диаграммы монокристаллических структур

В разд. 1.1.7 на примере атома водородоподобного типа показано, что полная энергия микрочастицы (в частности, электрона), находящейся в потенциальной яме, квантуется. Последнее означает, что энергия микрочастицы может принимать ряд (в пределе бесконечный) строго определенных дискретных значений.

Рассмотрим вопрос: во что трансформируется спектр разрешенных энергетических состояний электрона, находящегося в атоме (потенциальной яме), при образовании кристаллического тела из большой совокупности индивидуальных атомов одного и того же типа. При этом следует иметь в виду то обстоятельство, что расстояние между соседними атомами в твердом теле настолько мало (порядка 1 ангстрем), что каждый атом находится в силовом поле, образованном соседними атомами. Несколько забегая вперед, отметим, что именно факт наличия

оговоренного выше атомного силового поля вызывает трансформацию энергетических диаграмм (спектров) индивидуальных атомов при образовании ими кристаллических структур. В рамках настоящего раздела этот вопрос обсуждается на качественном уровне рассмотрения.

На рис. 1.7 представлены энергетические диаграммы и волновые функции $\phi(x)$ для случая двух однотипных атомов, разнесенных друг от друга на столь значительное расстояние r , что их атомные силовые поля можно считать не взаимодействующими.

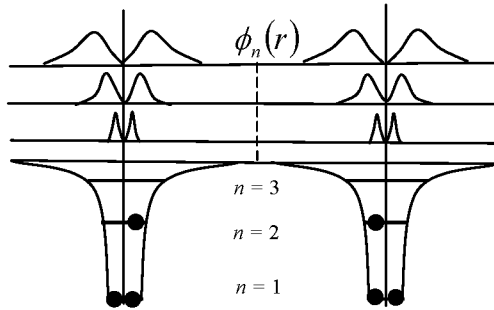


Рис. 1.7. Амплитудные волновые функции электронов при $n = 1, 2, 3$ для двухатомной системы (случай отсутствия взаимодействия атомных полей)

Для дальнейшего рассмотрения вопроса необходимо отметить ряд характерных деталей (рис. 1.7):

- атомные потенциальные функции $U(r)$ пространственно разнесены и в этой связи не претерпевают искажений, обусловленных присутствием соседнего атома;
- посредством горизонтальных отрезков, расположенных в потенциальной яме, отображены разрешенные энергетические состояния (уровни) электронов в атоме;
- чем глубже расположен электрон в потенциальной яме (например: при $n = 1$), тем он более локализован в пространстве (т.е. более несвободен в своем движении);
- при очень больших значениях параметра n разрешенные энергетические состояния подходят близко к потолку потенциальной ямы и волновые функции $\phi_n(r)$ становятся все более протяжен-

ными, а следовательно, поведение электрона приобретает черты свободного движения микрочастиц (т.е. волн де Бройля);

- суперпозиция всевозможных амплитудных волновых функций $\phi_n(r)$ при значениях параметра $n = 1, 2, 3, \dots, \infty$ образует волновой пакет, реализующий фактор локализации микрочастицы в пределах потенциального ящика (атома);
- жирными точками условно обозначено положение электронов, находящихся на разрешенных энергетических уровнях (согласно принципу Паули их должно быть не более двух на одном энергетическом уровне).

Сблизим теперь атомы друг к другу до такого малого расстояния, чтобы их атомные поля начали перекрываться, и рассмотрим возможные варианты заселения энергетических уровней имеющимися у атомов электронами (в данном случае это шесть электронов, отображенных на рис. 1.7). Здесь надо иметь в виду то обстоятельство, что если ранее имели место два независимых объекта (атомов), то теперь имеется одна общая атомная система, подчиняющая «интересы» индивидуальных атомов «интересам» системы в целом.

Под последним утверждением понимается то фундаментальное обстоятельство, что любая физическая система в процессе ее самопроизвольной эволюции стремится в конечном итоге принять состояние, обладающее минимальной энергией. (В более строгой постановке, здесь следует говорить о выполнении принципа *минимальной свободной энергии* системы, который является своеобразным компромиссом между принципами *минимальной энергии* и *максимальной энтропии* системы [6]).

На рис. 1.8 рассмотрены возможные варианты трансформации энергетической диаграммы индивидуальных атомов при образовании ими кристаллической структуры.

В первом варианте «сложения» энергетических спектров индивидуальных атомов (на рис. 1.8 – зачеркнут) энергетическая диаграмма оставлена неизменной, и электроны при образовании кристаллической структуры начинают заполнять более высоко отстоящие свободные энергетические уровни (напомним, что согласно принципу Паули, на

каждом энергетическом уровне может быть расположено не более двух электронов). Но в этом случае, при большом количестве атомов, образующих кристалл (в 1 см^3 кристалла порядка 10^{23} атомов), электроны начнут заполнять энергетические уровни, весьма близко отстоящие от нулевого уровня «0», в результате чего работа выхода электронов из кристалла становится функцией размера образуемого кристалла. Поскольку на практике это свойство кристалла не проявляется, рассмотренный вариант энергетической диаграммы является ошибочным.

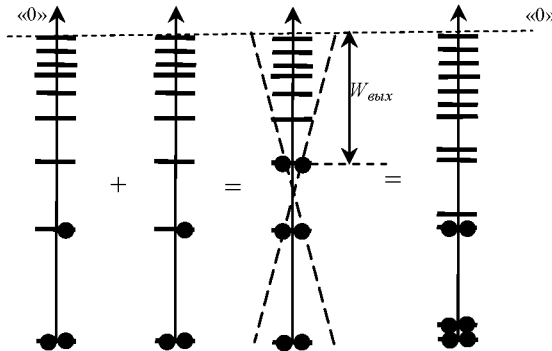


Рис. 1.8. Схематическое отображение трансформации энергетического спектра состояний электронов при образовании кристалла ($W_{\text{вых}}$ – работа (энергия) выхода электрона из кристалла)

В действительности, в природе наблюдается ситуация, представленная на рис. 1.8 справа. Ей соответствует тот факт, что *при объединении атомов в кристаллическую структуру каждый индивидуальный энергетический уровень атома расщепляется на совокупность $N/2$ дополнительных уровней, где N – количество атомов, задействованных в процессе создания кристалла.*

Для дальнейшего рассмотрения важно акцентировать внимание на трех принципиальных моментах:

- при образовании кристалла каждый энергетический уровень электрона в атоме вырождается в *энергетическую зону*, содержащую большое количество ($N/2$) разрешенных энергетических состояний и имеющую протяженность (по энергетической шкале исчисления) порядка 1 эВ;

- расстояние между соседними уровнями в энергетической зоне составляет по порядку величины – 10^{-23} эВ. Столь малое значение энергетического интервала позволяет считать, что энергетический спектр разрешенных состояний в пределах разрешенной энергетической зоны является (практически) *непрерывным*, а не *дискретным*, как это имеет место в индивидуальных атомах. (Для сравнения отметим, что средняя тепловая энергия микрочастицы, приходящаяся на одну степень свободы, составляет (при температуре окружающей среды $T = 300$ К) 0.025 эВ. Приведенное значение средней тепловой энергии более чем на двадцать порядков превышает энергетический интервал между соседними уровнями, расположенными в пределах разрешенной энергетической зоны кристалла. В этой связи при рассмотрении *макроскопических тел дискретный характер энергетического спектра в разрешенной энергетической зоне практически не проявляется*;

- значительная протяженность (1 эВ) энергетических зон при относительно небольшой полной глубине атомной потенциальной ямы (2 – 28 эВ), как правило, приводит к *эффекту перекрытия зон*, т.е. к их наложению друг на друга, что еще больше способствует образованию непрерывного спектра допустимых энергетических состояний, но уже в пределах объединенных зон.

Появление энергетических зон в кристаллических структурах существенным образом сказывается на их электрических и оптических характеристиках, что и обуславливает практический интерес к этому вопросу. На рис. 1.9 представлена обобщенная энергетическая диаграмма фрагмента кристаллической структуры, построенная с учетом моментов, приведенных на рис. 1.7 и 1.8.

Уточним и конкретизируем ряд характерных графических фрагментов, представленных на рис. 1.9:

- посредством штриховки отображены факторы наличия дискретных энергетических уровней и конечной протяженности (по энергетической шкале) индивидуальной энергетической зоны кристаллической структуры, соответствующей определенному значению квантового числа n ;

- точками, расположенными в пределах энергетической зоны, отражен тот факт, что в данной зоне индивидуальные энергетические уровни могут быть полностью либо частично заполнены электронами;
- точка «А» характеризует местоположение максимального значения потенциальной энергии кристаллической системы, имеющей периодический характер по пространственной координате x (рассматривается одномерный случай);
- штриховая линия на верхнем фрагменте графика функции $\phi_n(x)$ отображает суммарное значение амплитудной волновой функции в пределах этого интервала (в рамках модельного представления о суперпозиции индивидуальных вероятностных волн де Бройля, характерных для свободной частицы);
- функция потенциальной энергии (в нижней части рисунка) претерпела значительное изменение относительно невозмущенной функции потенциальной энергии (см. рис. 1.7), что обусловлено фактором сильного сближения соседних атомов (до расстояния порядка 1 ангстрем) в процессе образования кристаллической структуры.

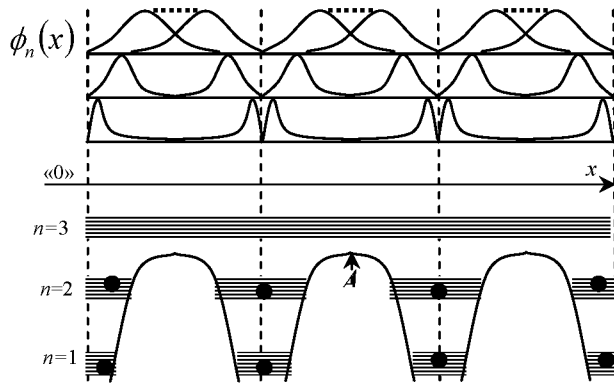


Рис. 1.9. Схематическое отображение энергетических зон в кристалле и соответствующих им амплитудных волновых функций $\phi_n(x)$

На основании схематического отображения, представленного на рис. 1.9, можно сформулировать ряд важных, с практической точки зрения, выводов относительно характера поведения электронов, находя-

щихся в атомных структурах с периодической потенциальной функцией:

- амплитудные волновые функции электронов, расположенных в глубоко лежащих энергетических зонах ($n = 1$), имеют ярко выраженный локализованный характер, а следовательно, возможность их перемещения в пределах кристаллической структуры полностью исключается (т.е. их следует считать абсолютно несвободными с позиции представлений о ПМВ де Бройля);
- электроны, находящиеся в энергетической зоне, условно характеризуемой квантовым числом $n = 2$, имеют более протяженные в пространстве амплитудные волновые функции и для них существует определенная вероятность движения в пределах кристаллической структуры (т.е. их следует считать ограниченно свободными с позиции представлений о ПМВ де Бройля);
- электроны, находящиеся в энергетической зоне, условно характеризуемой квантовым числом $n = 3$ (обратить внимание на штриховую линию на фрагменте амплитудной волновой функции), могут рассматриваться как свободные микрочастицы, движение которых полностью соответствует поведению элементарных ПМВ де Бройля.

Для полноты рассмотрения необходимо отметить, что согласно рис. 1.9 взаимное расположение атомов принято статичным, т.е. предполагается, что межатомное расстояние является неизменной величиной. В реальности же наличие фактора теплового колебательного движения атомов (получивших название *фононов*) приводит к появлению вариаций этого расстояния, а следовательно, положение точки A на энергетической шкале (см. рис. 1.9) тоже непрерывно меняется (происходит ее перемещение вверх или вниз) во времени. В этой связи амплитудные волновые функции испытывают аналогичные динамические вариации, что может внести свои негативные коррективы в характер движения электронов (в частности может иметь место рассеяние поступательного движения электронов на фононах). Представляется очевидным, что интенсивность процессов рассеяния поступательного движения электронов в кристаллической структуре должна подчиняться общим статистическим закономерностям процессов образования фононов.

Энергетическая диаграмма, представленная на рис. 1.9, не используется в современной инженерно-физической практике, поскольку она обременена большим количеством деталей. На практике используется гораздо более лаконичное отображение, получившее название *зонной диаграммы* кристаллической структуры, представленное на рис. 1.10.

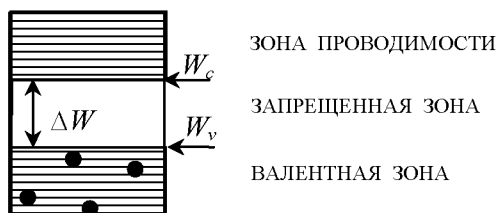


Рис. 1.10. Зонные энергетические диаграммы кристаллических структур

Для построения зонной диаграммы (рис. 1.10) используются следующие фрагменты рис. 1.9:

- самая высоко расположенная энергетическая зона, заполненная электронами ($n = 2$, на рис. 1.9), – *валентная зона*;
- самая низко расположенная энергетическая зона, свободная от электронов при температуре кристалла равной нулю ($n = 3$, на рис. 1.9), – *зона проводимости*;
- интервал между вышеуказанными зонами – *запрещенная зона*.

На рис. 1.10 использованы символы, нашедшие широкое распространение в современной научно-технической практике:

- W_c – значение энергии для самого низко расположенного энергетического уровня в зоне проводимости;
- W_v – значение энергии для самого высоко расположенного энергетического уровня в валентной зоне;
- ΔW – ширина запрещенной зоны.

Несмотря на лаконичность рассмотренной зонной диаграммы, ее вполне достаточно для объяснения практически всех электрофизических и оптических характеристик кристаллических структур с электронным типом проводимости. В частности, с ее помощью проводится классификация кристаллических тел по критерию их электропроводности: $\Delta W = 0$ – металлы (энергетические зоны перекрываются);

$\Delta W < 1.5$ эВ – собственные полупроводники; $\Delta W > 1.5$ эВ – диэлектрики. Из приведенного принципа классификации веществ по критерию их уровня электропроводности следует, что разделение материалов на диэлектрики и полупроводники имеет несколько условный характер.

б) Спектр энергетических состояний в рамках модели слабосвязанных электронов

В вышеизложенном пункте (а) настоящего раздела проведен анализ (на качественном уровне рассмотрения) энергетических диаграмм конденсированных сред, что позволило уяснить основные физические принципы появления практически непрерывных энергетических спектров разрешенных состояний микрочастиц в регулярных многоатомных структурах и целесообразность их отображения в виде зонных энергетических диаграмм. Ниже этот же круг вопросов будет рассмотрен более корректно, с использованием математического аппарата волновой квантовой механики.

Последующий анализ проводится в рамках модели, основанной (в целях упрощения математических выкладок) на приближении *слабосвязанных электронов* [7]. Согласно этому приближению, потенциальная диаграмма кристаллической структуры представляется в виде потенциального ящика со слаборифленным дном, отражающим периодический характер функции потенциальной энергии электрона в кристаллической структуре. При этом считается, что потенциал решетки мало отличается от потенциала решетки в приближении свободных электронов. Последнее приближение можно допустить в случае, если глубина потенциального ящика намного превышает амплитуду рифления дна ящика.

Привлекательность приближения слабосвязанных электронов состоит в том, что в этом случае представляется возможным искать решение уравнения Шредингера (1.34а) в виде совокупности элементарных ПМВ де Бройля. Однако следует иметь в виду то обстоятельство, что периодические вариации потенциальной энергии (рифление дна ящика) должны найти свое отражение в общем решении уравнения Шредингера. Представляется вполне естественным, что указанная периодичность потенциала в кристаллической структуре должна привес-

ти к появлению фактора *дополнительной периодичности* амплитудной волновой функции $\phi(x, y, z)$.

С учетом вышеизложенных моментов амплитудная волновая функция электрона, находящегося в периодическом потенциале, ищется в виде выражения (для одномерного случая)

$$\phi(x) = u(x) \cdot e^{i \cdot k \cdot x}, \quad (1.59)$$

где $u(x)$ – некоторая периодическая функция с периодом, равным *постоянной* (межцентровому расстоянию между соседними атомами) кристаллической структуры.

Выражение (1.59) в современной научно-технической практике получило название *функции Блоха*. Нетрудно заметить, что по своей сути функция Блоха представляет собой рассмотренную ранее ПМВ де Бройля (фрагмент $e^{i \cdot k \cdot x}$), амплитуда которой модулирована периодической функцией $u(x)$. Конкретный вид функции $u(x)$ определяется видом периодической потенциальной функции кристаллической структуры $U(x)$, используемой в уравнении Шредингера.

На рис. 1.11 представлена потенциальная функция кристаллической структуры (*модель Кронига-Пенни*), которая будет использована в дальнейшем для нахождения спектра разрешенных энергетических состояний в регулярной кристаллической структуре.

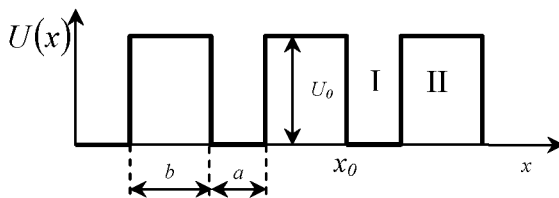


Рис. 1.11. Фрагмент потенциальной диаграммы кристаллической структуры в рамках модели Кронига-Пенни

Допустим, что энергия электрона W меньше высоты потенциально-го барьера U_0 (в этом случае можно считать, что он заперт в ограниченных пределах потенциального ящика с рифленным дном). В общем случае состояние электрона (амплитудная волновая функция) описывается уравнением Шредингера (см. (1.34a)):

$$\frac{d^2 \phi(x)}{dx^2} + \frac{8 \cdot \pi^2 \cdot m}{h^2} \cdot (W - U(x)) \cdot \phi(x) = 0. \quad (1.60)$$

Для областей (I), для которых $U(x) = 0$, решение уравнения Шредингера (1.60) имеет вид (подробнее см. раздел № 1.1.7, пункт «б»)

$$\phi_I(x) = A \cdot e^{i \cdot \alpha \cdot x} + B \cdot e^{-i \cdot \alpha \cdot x}, \quad (1.61)$$

где A и B – постоянные коэффициенты; α – характеристический параметр движения микрочастицы в области (I), равный

$$\alpha = \sqrt{\frac{8 \cdot m \cdot \pi^2 \cdot W}{h^2}}.$$

Для областей (II), для которых $U(x) = U_0$, решение уравнения Шредингера (1.60) имеет вид (подробнее см. раздел № 1.1.7, пункт «в»)

$$\phi_{II}(x) = C \cdot e^{\beta \cdot x} + D \cdot e^{-\beta \cdot x}, \quad (1.62)$$

где C и D – постоянные коэффициенты; β – характеристический параметр движения микрочастицы в области (II), равный

$$\beta = \sqrt{\frac{8 \cdot m \cdot \pi^2 \cdot (U_0 - W)}{h^2}}.$$

Подставив амплитудные волновые функции $\phi_I(x)$ и $\phi_{II}(x)$ из выражений (1.61) и (1.62) в выражение для функции Блоха (1.59), можно определить вид периодической функции $u(x)$ для соответствующих пространственных областей (I) и (II):

$$u_I(x) = A \cdot e^{i(\alpha-k)x} + B \cdot e^{-i(\alpha+k)x}; \quad (1.63)$$

$$u_{II}(x) = C \cdot e^{(\beta-i \cdot k)x} + D \cdot e^{-(\beta+i \cdot k)x}. \quad (1.64)$$

Для определения постоянных коэффициентов A , B , C и D воспользуемся условиями:

- Условие № 1: Функция $u(x)$ и ее первая производная считаются непрерывными, в том числе в местах скачка потенциальной энергии (например в точке x_0), поскольку реальные природные объекты не имеют бесконечных скоростей в процессах эволюции:

$$u_I(x_0) = u_{II}(x_0); \quad \left(\frac{du_I(x)}{dx} \right)_{x=x_0} = \left(\frac{du_{II}(x)}{dx} \right)_{x=x_0}. \quad (1.65)$$

• Условие № 2: Функция $u(x)$ обладает свойством периодичности по геометрическим координатам с периодом, равным $(a+b)$:

$$u_I(x_0) = u_{II}(x_0 + a + b);$$

$$\left(\frac{du_I(x)}{dx} \right)_{x=x_0} = \left(\frac{du_{II}(x)}{dx} \right)_{x=x_0+a+b}. \quad (1.66)$$

Подставив в (1.65) и (1.66) функции $u_I(x)$ и $u_{II}(x)$ (согласно выражениям (1.63) и (1.64)), получим систему из четырех уравнений, необходимых для определения значений постоянных коэффициентов: A , B , C , D . Чтобы рассматриваемая система уравнений имела решения для указанных коэффициентов A , B , C , D , отличные от нуля, необходимо, чтобы определитель системы уравнений был равен нулю. Последнее требование приводит к уравнению

$$\frac{\beta^2 - \alpha^2}{2 \cdot \alpha \cdot \beta} \cdot \sin(\alpha \cdot a) \cdot (sh(\beta \cdot b)) +$$

$$+ \cos(\alpha \cdot a) \cdot (ch(\beta \cdot b)) = \cos k \cdot (a + b). \quad (1.67)$$

В целях упрощения дальнейшего анализа рассмотрим широко используемый предельный частный случай, в котором:

- ширина атомного потенциального барьера b (согласно рис. 1.11) устремляется к нулю ($b \rightarrow 0$);
- высота атомного потенциального барьера синхронно устремляется к бесконечности ($U_0 \rightarrow \infty$), причем таким образом, чтобы произведение $(b \cdot U_0)$ оставалось постоянной величиной (некий аналог дельта-функции).

Введем в рассмотрение характеристический параметр P – *прозрачность потенциального барьера* для реализации процесса движения микрочастицы по периодической структуре согласно рис. 1.11:

$$P = \frac{4 \cdot \pi^2 \cdot m \cdot a}{h^2} \cdot b \cdot U_0. \quad (1.68)$$

С учетом сделанных упрощений ($b \rightarrow 0; U_0 \rightarrow \infty$) и использованием (1.68) выражение (1.67) преобразуется к виду

$$P \cdot \frac{\sin(\alpha \cdot a)}{\alpha \cdot a} + \cos(\alpha \cdot a) = \cos(k \cdot a). \quad (1.69)$$

Выражение (1.69) отражает в неявном виде зависимость энергии электрона W (которая входит в состав параметра α , см. выражение (1.61)) от волнового вектора k и прозрачности потенциального барьера P .

На рис. 1.12 представлено графическое решение уравнения (1.69). По оси абсцисс отложен комплексный параметр $(\alpha \cdot a)$, а по оси ординат – фрагмент выражения (1.69):

$$\left(P \cdot \frac{\sin(\alpha \cdot a)}{\alpha \cdot a} + \cos(\alpha \cdot a) \right).$$

Поскольку функция $\cos(k \cdot a)$, стоящая в правой части выражения (1.69), не может быть (по определению) больше +1 и меньше -1, то параметр $(\alpha \cdot a)$ может принимать только такие значения, при которых левая часть уравнения (1.69) не может выходить за пределы значений ± 1 .

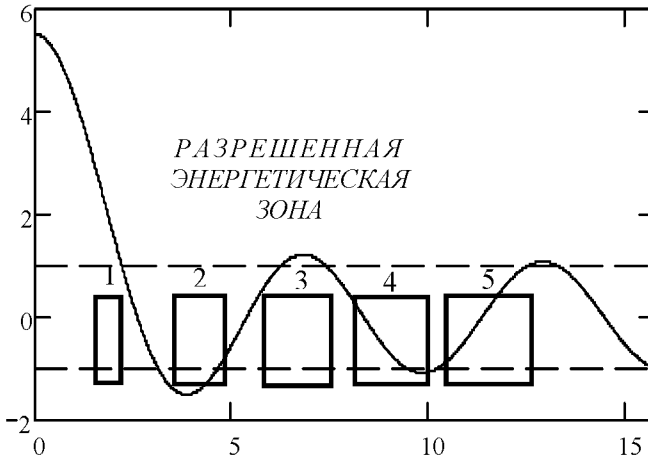


Рис. 1.12. Схематическое отображение процесса образования энергетических зон в структуре с периодической потенциальной энергией

Из рис. 1.12 следует, что вся шкала энергий (ось абсцисс) разделяется на ряд дискретных интервалов:

- *разрешенные энергетические зоны*, в которых значения энергий лежат в интервалах, ограниченных прямоугольниками;
- *запрещенные энергетические зоны*, в которых значения энергий лежат в интервалах, составляющих промежутки между прямоугольниками.

С увеличением энергии электрона (т.е. параметра $(\alpha \cdot a)$) ширина разрешенных энергетических зон увеличивается, а ширина запрещенных зон уменьшается.

Как и следовало ожидать, ширина разрешенных и запрещенных энергетических зон является функцией прозрачности потенциального барьера P . При выполнении условия $P \rightarrow \infty$ разрешенные зоны сужаются, превращаясь в дискретные уровни, соответствующие условию $\alpha \cdot a = n \cdot \pi$, где параметр $n = \pm 1, \pm 2, \dots$, т.е. к значениям энергий, соответствующим изолированной потенциальной яме. При выполнении условия $P \rightarrow 0$ возникает противоположная ситуация, при которой исчезают запрещенные зоны и движение электрона становится практически свободным (аналог ПМВ де Бройля).

Согласно вышеизложенным моментам, внутри каждой разрешенной энергетической зоны (см. рис. 1.12) энергия электрона непрерывно возрастает с увеличением волнового вектора k . При значениях последнего, равных

$$k = n \cdot \frac{\pi}{a} ; (n = \pm 1, \pm 2, \dots), \quad (1.70)$$

энергия электронов претерпевает разрыв, приводящий к появлению на энергетической диаграмме запрещенных зон (для преемственности анализа сравнить с материалами раздела 1.1.7,в).

Принимая во внимание, что волновой вектор $k = 2\pi/\lambda$, выражение (1.70) можно преобразовать к более явному виду с позиции представлений о ПМВ де Бройля и заодно получить известную формулу Вульфа-Брегга, выражающую условие дифракции волн де Бройля для случая, когда угол падения электронного луча на плоскость кристаллической решетки равен 90° :

$$n \cdot \lambda = 2 \cdot a, \quad (n = \pm 1, \pm 2, \dots). \quad (1.71)$$

Для последующего рассмотрения важно отметить, что зависимость $E(k)$ вблизи нижней и верхней границ разрешенной зоны является квадратичной функцией волнового вектора, отсчитанного от численного значения этой границы [7]:

$$W = W_0 + A \cdot k^2, \quad (1.72)$$

где W_0 – энергия, соответствующая границе зоны; A – постоянный коэффициент.

У дна разрешенной зоны парабола (1.72) является *нормальной* – $A > 0$, а у потолка зоны парабола является *обращенной* – $A < 0$ (подробный анализ этого вопроса выходит за рамки настоящего пособия и приведен в источнике [8]).

*в) Изоэнергетические поверхности состояния электронов
в кристаллических структурах*

На рис. 1.12 представлено графическое отображение зависимости энергии электронов от волнового вектора, испытывающего вариации значений от 0 до $\pm\infty$ (на рисунке изображена только положительная ветвь (!)). Периодический характер поведения волнового вектора (см. выражение (1.70)) позволяет выразить энергию электронов, находящихся в различных разрешенных энергетических зонах, с помощью волновых векторов, лежащих в пределах первой зоны, т.е. принимающих значения в интервале от $-\pi/2$ до $+\pi/2$. Эта процедура представляется тем более возможной, если учесть тот факт, что выражению (1.60) удовлетворяет не только функция $\phi(x) = U(x) \cdot \exp(i \cdot k' \cdot x)$, где k' – волновой вектор, лежащий в пределах первой зоны, но и функция $\phi(x) = U(x) \cdot \exp(i \cdot k \cdot x)$, где вектор k , отображающий состояние электронов в любой другой энергетической зоне под номером n (см. рис 1.12.), удовлетворяет условию

$$k = k' + n \cdot \frac{2\pi}{a}. \quad (1.73)$$

Процедура построения всех энергетических зон в пределах первой энергетической зоны называется *приведением зон*, а сами зоны – *приведенными зонами*. Смысл процедуры приведения энергетических зон состоит в перемещении фрагментов зависимости функции $W(k)$, относящихся к различным зонам, к отрезку оси энергии с координатами $-\pi/2$ и $+\pi/2$.

На рис. 1.13 представлено схематическое отображение функций $W(k)$, приведенных в нормальном (сплошные линии) и приведенном (пунктирные линии) виде. Штрихпунктирной линией изображена зависимость $W(k)$ для случая свободного электрона (см. выражение (1.15)). Из рисунка следует, что у всех зон с нечетными значениями параметра n в центре расположены *минимумы энергии*, а с четными значениями n – *максимумы энергии*. Характерным моментом энергетической диаграммы является увеличение ширины разрешенных зон и уменьшение ширины запрещенных зон с возрастанием волнового вектора k , что нашло свое отражение на фрагменте, расположенном справа на рис. 1.13 для частных случаев $n = 1, 2, 3$.

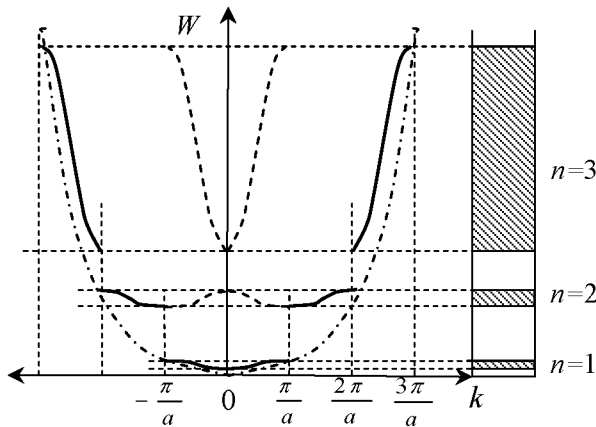


Рис. 1.13. Схематическое отображение функций $W(k)$, представленных в нормальном (сплошные линии) и приведенном (пунктирные линии) виде

Приведенные выше материалы могут быть распространены и на трехмерный случай, характерный для реальных кристаллических струк-

тур. Так как параметры решетки (т.е. периодичности), как правило, не одинаковы по разным направлениям в кристалле, то и значения волнового вектора k , при которых наступают разрывы в разрешенных энергетических зонах, будут также различны (см. выражение (1.70)):

$$k_x = n \cdot \frac{\pi}{a}; \quad k_y = n \cdot \frac{\pi}{b}; \quad k_z = n \cdot \frac{\pi}{c}, \quad (n = \pm 1, \pm 2, \dots), \quad (1.74)$$

где a, b, c – параметры кристаллической решетки соответственно по координатным направлениям x, y, z .

По аналогии с (1.73), для трехмерной кристаллической структуры вблизи дна разрешенной зоны, характеризуемой нечетными значениями параметра $n = 1, 3$, для функции $W(k)$ имеем

$$W(k) = W_0 + A \cdot k_x^2 + B \cdot k_y^2 + C \cdot k_z^2, \quad (1.75)$$

где W_0 – энергия электрона, соответствующая местоположению дна разрешенной зоны; A, B, C – постоянные коэффициенты, которые в случае минимума (т.е. нечетных значений параметра n) являются положительными величинами.

Для кристаллов кубического типа, в которых ввиду симметрии $A = B = C$, выражение (1.75) трансформируется к виду

$$W(k) = W_0 + A \cdot (k_x^2 + k_y^2 + k_z^2) = W_0 + \frac{\hbar^2}{8 \cdot \pi^2 \cdot m_n} \cdot (k_x^2 + k_y^2 + k_z^2). \quad (1.76)$$

В выражении (1.76) постоянный коэффициент A представлен в виде, аналогичном случаю движения свободного электрона (см. выражение (1.15)), с тем лишь отличием, что масса свободного электрона заменена на *эффективную массу* электрона, значение которой характерно для каждого конкретного типа кристаллической структуры.

Для четных значений параметра n (в центре приведенной зоны находится максимум энергии) функция $W(k)$ также определяется с помощью выражений (1.75) и (1.76), однако в этом случае постоянные коэффициенты A, B, C являются *отрицательными* величинами. Интересно отметить, что для этого случая и эффективная масса электрона m_n также приобретает отрицательное значение, в силу того что фрагмент $(\hbar^2) / (8 \cdot \pi^2)$ в выражении (1.76) является положительной величиной.

Физическая сущность и вытекающие практические следствия из этого парадоксального феномена рассмотрены в источнике [7].

Поверхность в координатном представлении $\{k_x, k_y, k_z\}$, отображающая энергетическое состояние системы и удовлетворяющая условию $W(k) = \text{const}$, называется *изоэнергетической поверхностью*. С учетом вышеизложенных моментов изоэнергетические поверхности в пространстве состояний $\{k_x, k_y, k_z\}$, в общем случае имеют вид *эллипсоидов* вращения (согласно выражению (1.75)), а в частном случае, для кристаллических структур кубического типа, – *сферический* вид (согласно выражению (1.76)).

На рис. 1.14 представлены характерные примеры отображения изоэнергетических поверхностей кристаллических структур.

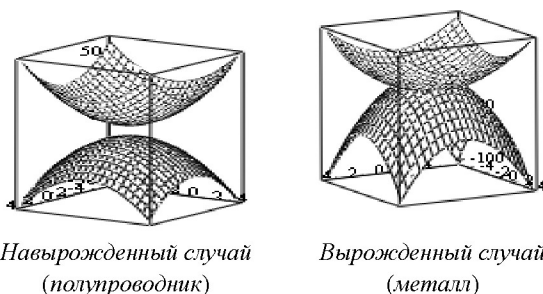


Рис. 1.14. Изоэнергетические поверхности состояния электрона (квантовые числа $n = 2$ и 3 выбраны согласно рис. 1.13)

При этом следует иметь в виду то обстоятельство, что экстремумы функции $W(k)$ не обязательно должны располагаться в центрах приведенных энергетических зон, поскольку последнее имеет место лишь в том случае, если граничные значения волновых векторов k_x, k_y, k_z в пределах индивидуальных зон имеют одинаковые величины.

г) Спектр энергетических состояний в наноразмерных структурах типа (0D)

В наноразмерных структурах типа (0D) движение электронов ограничено по всем трем координатным направлениям. Эти ограничения соответствуют случаю нахождения микрочастицы в трехмерном потенциальном ящике, который имеет весьма малую протяженность по всем

координатным осям – $\{x, y, z\}$. В современной научно-технической практике диапазон этой протяженности сопоставим с интервалом значений характеристических длин порядка 1 – 100 нм. Именно исходя из этого обстоятельства такие структуры получили название *квантовой точки*, несмотря на то что они могут содержать достаточно большое количество атомов.

На рис. 1.15 представлено модельное отображение квантовой точки с позиции общих представлений о трехмерном потенциальном ящике.

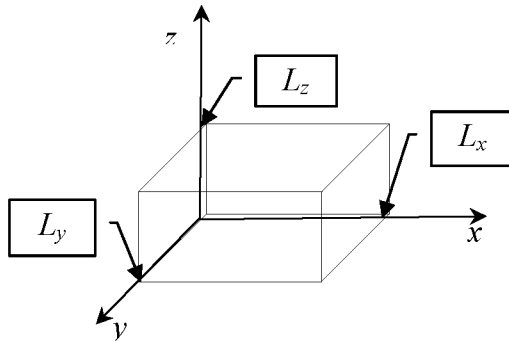


Рис. 1.15. Пространственное отображение квантовой точки с характеристическими размерными параметрами $\{L_x, L_y, L_z\}$

Беря за основу одномерную модель потенциального ящика (см. выражение (1.55)), для функции $W_{(0D)}$ применительно к квантовой точке в общем случае имеем

$$W_{(0D)}(n_x, n_y, n_z) = \frac{h^2}{8 \cdot m_e} \cdot \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right), \quad (1.77)$$

где L_x, L_y, L_z – пространственная протяженность квантовой точки соответственно по координатным направлениям $\{x, y, z\}$; n_x, n_y, n_z – квантовые числа из набора целочисленных величин $\{1, 2, \dots, \infty\}$, характеризующие локальное состояние микрочастицы соответственно по координатным направлениям $\{x, y, z\}$.

Для квантовой точки кубического типа (т.е. при выполнении условия $L_x = L_y = L_z = L$) выражение (1.77) трансформируется к более симметричному виду:

$$W_{(0D)}(n_x, n_y, n_z) = \frac{h^2}{8 \cdot m_e \cdot L^2} \cdot (n_x^2 + n_y^2 + n_z^2). \quad (1.78)$$

Из выражений (1.77) и (1.78) следует, что энергия электронов, находящихся в пределах квантовой точки, может принимать только ряд дискретных значений, разделенных друг от друга конечными интервалами энергии. Этот феномен дискретизации значений энергии электронов, находящихся в квантовой точке наглядно представлен на рис. 1.16 с помощью двумерного отображения энергетического состояния электронов.

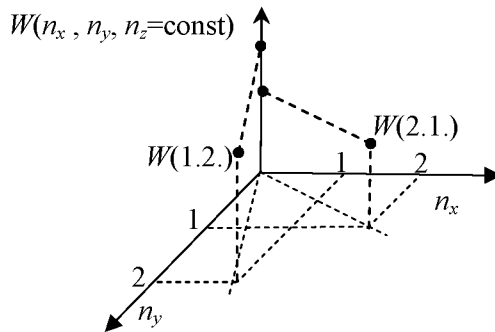


Рис. 1.16. Схематическое отображение фактора дискретизации энергии электронов в квантовой точке ($\{W(1,2): n_x = 1; n_y = 2\}$; $\{W(2,1): n_x = 2; n_y = 1\}$)

Выражение (1.78) свидетельствует о возможности появления в квантовых точках кубического типа вырожденных энергетических состояний, обусловленных вариациями значений совокупности квантовых чисел $\{n_x, n_y, n_z\}$ при одновременном выполнении условия

$$(n_x^2 + n_y^2 + n_z^2) = \text{const.}$$

Энергетические спектры (1.77) и (1.78) можно представить как функциональные зависимости от волновых векторов. Для этого вос-

пользуемся представлением (согласно рис. 1.3) $L_i/n_i = \lambda_i/2$ и определением волнового вектора $k_i = (2 \cdot \pi)/\lambda_i$, где параметр i относится к геометрическим координатам $\{x, y, z\}$. После очевидных преобразований получим выражение $L_i/n_i = \pi/k_i$, на основании которого преобразуем выражение (1.77) к виду

$$W_{(0D)}(k_x, k_y, k_z) = \frac{h^2}{8 \cdot m_e \cdot \pi^2} \cdot (k_x^2 + k_y^2 + k_z^2), \quad (1.79)$$

где волновые векторы $\{k_x, k_y, k_z\}$ могут принимать строго фиксированные численные значения, соответствующие квантовым числам $\{n_x, n_y, n_z\}$.

Интересно отметить, что выражение (1.79) справедливо в одинаковой мере к любым квантовым точкам, независимо от их геометрической конфигурации (в отличие от представления, основанного на квантовых числах).

Важным аспектом при проведении физико-химического анализа наноразмерных систем является функция плотности квантовых состояний электронов. По определению, плотность квантовых состояний $g(W)$ – отношение количества квантовых состояний ΔG , приходящихся на малый интервал энергий ΔW (в пределе бесконечно малый), к величине этого энергетического интервала ΔW : $g(W) = \Delta G / \Delta W$. В частности, функция $g(W)$ необходима для определения концентрации носителей зарядов в связи с рассмотрением вопроса электропроводности в наноразмерных системах.

Применительно к наноразмерным системам типа (0D) функция плотности квантовых состояний электронов имеет вид [5]

$$g_{(0D)}(W) = \frac{2}{L_x \cdot L_y \cdot L_z} \cdot \sum_{i,j,k} \delta(W - W_{i,j,k}), \quad (1.80)$$

где индексы $i, j, k = 1, 2, \dots, \infty$.

На рис. 1.17 представлено графическое отображение функции плотности квантовых состояний $g_{(0D)}(W)$, характерной для наноразмерной структуры типа квантовой точки.

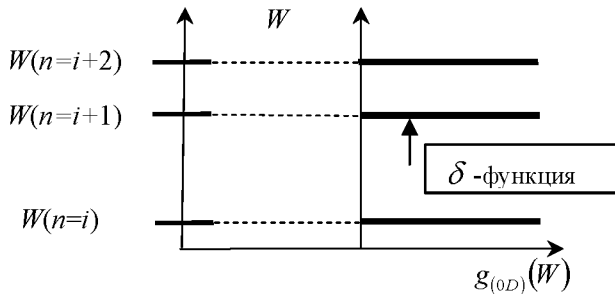


Рис.1.17. Функция плотности квантовых состояний для случая квантовой точки (правый фрагмент) и соответствующий ей спектр разрешенных квантовых состояний (левый фрагмент)

Присутствие в выражении (1.80) δ – функции весьма характерно и свидетельствует о том, что энергетический спектр электронов представляет собой совокупность индивидуальных дискретных квантовых состояний. Именно благодаря подобному сходству энергетических спектров индивидуальных атомов и квантовых точек, эти наноразмерные структуры в ряде практических случаев называются *искусственными атомами*.

д) *Спектр энергетических состояний в наноразмерных структурах типа (1D)*

В наноразмерных структурах типа (1D) движение электронов ограничено по двум координатным направлениям. Эти ограничения соответствуют случаю нахождения микрочастицы в трехмерном потенциальном ящике, который имеет весьма малую протяженность по двум координатным осям $\{y, z\}$. В современной научно-технической практике диапазон этой протяженности так же, как было уже изложено выше, сопоставим с интервалом значений характеристических длин порядка 1 – 100 нм. Именно исходя из этого обстоятельства такие структуры получили название *квантовой проволоки (нити, шнура)*. Квантовые проволоки, как и квантовые точки, могут содержать в своем составе

достаточно большое (в пределе бесконечное, при бесконечной протяженности проволоки) количество атомов.

На рис. 1.18 представлено модельное отображение квантовой проволоки с позиции общих представлений о трехмерном, протяженном в направлении одной из координат потенциальном ящике. Беря за основу анализа одномерную модель потенциального ящика для координатных направлений $\{y, z\}$ (см. выражение (1.55)) и модельное представление о движении свободного электрона по координатному направлению $\{x\}$ (см. выражение (1.15)), для функции $W_{(1D)}$ применительно к квантовой проволоке имеем (сравнить приводимое выражение(1.77):

$$W_{(1D)}(n_y, n_z, k_x) = \frac{h^2}{8 \cdot m_e} \cdot \left(\frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right) + \frac{h^2}{8 \cdot m_e \cdot \pi^2} \cdot k_x^2, \quad (1.81)$$

где m_e – эффективная масса электрона при его поступательной форме движения по координатному направлению $\{x\}$; L_y, L_z – пространственная протяженность квантовой точки соответственно по координатным направлениям $\{y, z\}$; n_y, n_z – квантовые числа из набора целочисленных величин $\{1, 2, \dots, \infty\}$, характеризующие локальное состояние микрочастицы соответственно по координатным направлениям $\{y, z\}$; k_x – проекция волнового вектора поступательной формы движения электрона по направлению оси $\{x\}$.

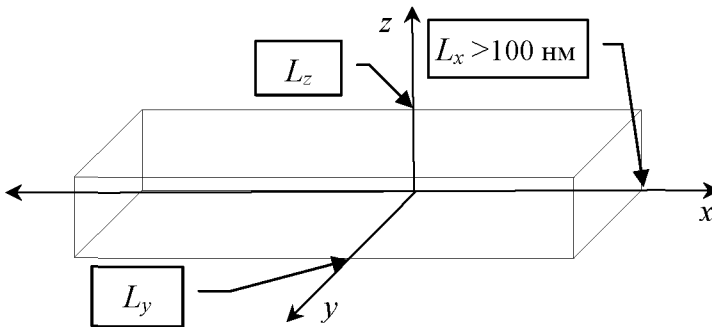


Рис. 1.18. Пространственное отображение квантовой проволоки с характеристическими размерными параметрами $\{L_x, L_y, L_z\}$ (L_x – микро-, мезо- или макроскопическая величина)

Из выражений (1.81) следует, что координатные составляющие допустимых значений энергии электронов, находящихся в пределах квантовой проволоки, могут принимать при движении электронов:

- по осям $\{y, z\}$ – только ряд дискретных значений, разделенных друг от друга конечными интервалами энергии (как это уже имело место для случая квантовой точки);
- по оси $\{x\}$ – практически непрерывный ряд значений (ввиду значительной протяженности наноразмерной структуры по данному направлению) в интервале величин $\{(2 \cdot \pi) / L_x - (2 \cdot \pi) / a\}$, где a – параметр кристаллической структуры в направлении оси $\{x\}$.

На рис. 1.19 представлено схематическое отображение энергетических спектров электронов в квантовой проволоке.

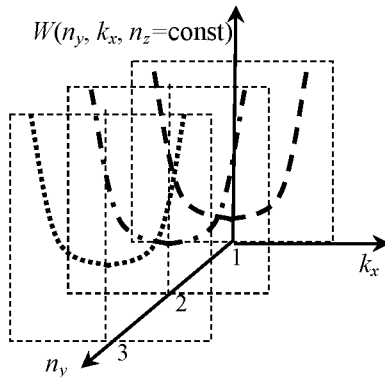


Рис.1.19. Схематическое отображение фрагмента энергетической диаграммы состояния электронов в квантовой проволоке

Для пояснения энергетической диаграммы рис. 1.19 уместно привести следующие комментарии:

- все энергетические зависимости параболического типа приведены для случая фиксированного значения квантового числа n_z ;
- пунктирные прямоугольники характеризуют плоскости, в которых сохраняется постоянное значение координатного квантового числа n_y , пробегающего от плоскости к плоскости ряд значений в диапазоне чисел (1, 2, 3);

- смещение индивидуальных параболических зависимостей вверх, наблюдаемое с ростом координатного квантового числа n_y , обусловлено постоянным вкладом (при одновременных вариациях значения волнового вектора k_x в диапазоне значений $\{(2 \cdot \pi)/L_x - (2 \cdot \pi)/a\}$) в энергию микрочастицы от фрагмента $[(\hbar^2/8m_e) \cdot (n_y^2/L_y^2)]$ формулы (1.81);

- различие в начертании приведенных параболических зависимостей использовано с целью облегчения зрительного восприятия изображения.

Из сравнительного анализа рис. 1.16 и 1.19 следует то обстоятельство, что в наноразмерных структурах типа квантовой проволоки имеет место частичный отход от эффекта полной дискретизации энергии электронов, как это имело место в наноразмерных структурах типа квантовой точки.

Энергетический спектр (1.81) можно представить как функциональную зависимость от волновых векторов $\{k_x, k_y, k_z\}$ (процедура подобного перехода аргументов представлена выше, при рассмотрении структур типа (0D)):

$$W_{(0D)}(k_x, k_y, k_z) = \frac{\hbar^2}{8 \cdot m_e \cdot \pi^2} \cdot (k_x^2 + k_y^2 + k_z^2), \quad (1.82)$$

где волновые векторы $\{k_y, k_z\}$ могут принимать строго фиксированные численные значения, соответствующие квантовым числам $\{n_y, n_z\}$, а волновой вектор $\{k_x\}$ – непрерывный ряд значений в диапазоне величин $\{(2 \cdot \pi)/L_x - (2 \cdot \pi)/a\}$, где a – параметр кристаллической решетки по координатному направлению $\{x\}$.

Применительно к наноразмерным системам типа (1D) функция плотности квантовых состояний электронов $g(W)$ в направлениях, связанных с проявлением квантовых ограничений, имеет вид [5]

$$g_{(1D)}(W) = \frac{2 \cdot (2 \cdot m_e)^{1/2}}{\hbar \cdot L_y \cdot L_z} \cdot \sum_{i,j} (W - W_{i,j})^{-1/2}, \quad (1.83)$$

где индексы i, j – пробегает ряд дискретных значений в диапазоне величин $\{1, 2, \dots, \infty\}$.

На рис. 1.20 представлено схематическое отображение фрагмента функции плотности квантовых состояний электронов $g(W)$, находящегося в геометрических пределах наноразмерных структур типа (1D).

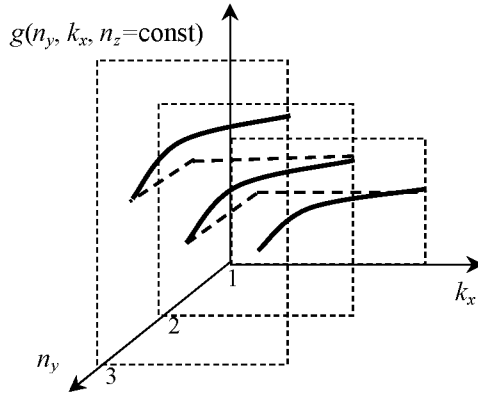


Рис. 1.20. Схематическое отображение фрагмента функции плотности квантовых состояний электронов $G(W)$ для структуры типа (1D)

Для пояснения энергетической диаграммы рис. 1.20 уместно привести следующие комментарии:

- все три фрагмента функции $g(W)$ (сплошные толстые линии) приведены для случая фиксированного значения квантового числа n_z ;
- пунктирные прямоугольники характеризуют плоскости, в которых сохраняется постоянное значение координатного квантового числа n_y , пробегающего от плоскости к плоскости ряд значений в диапазоне чисел (1, 2, 3);
- смещение индивидуальных параболических зависимостей вверх, наблюдаемое с ростом координатного квантового числа n_y , обусловлено постоянным вкладом (при одновременных вариациях значения волнового вектора k_x в диапазоне значений $\{ (2 \cdot \pi) / L_x - (2 \cdot \pi) / a \}$) в энергию микрочастицы от фрагмента $[(\hbar^2 / 8m_e) \cdot (n_y^2 / L_y^2)]$ формулы (1.81);

- пунктирными линиями условно отображен фактор перехода от совокупности энергетических состояний, характеризуемых квантовым числом, от $n_y = j$ до состояний $n_y = j+1$.

Приведенное на рис. 1.20 отображение функции $g(W)$ соответствует случаю отсутствия вырождения энергетических состояний, поскольку при построении графика принято, что квантовое число $n_z = \text{const}$. В противном случае может иметь место пересечение индивидуальных фрагментов зависимости функции $g(W)$, что приведет к появлению фактора вырождения энергетических состояний электронов.

Совокупность электронов в квантовых проволоках принято называть *одномерным электронным газом*.

е) Спектр энергетических состояний в наноразмерных структурах типа (2D)

В наноразмерных структурах типа (2D) движение электронов ограничено по одному координатному направлению. Это ограничение соответствует случаю нахождения микрочастицы в трехмерном потенциальном ящике, который имеет весьма малую протяженность по двум координатным осям $\{y, z\}$ (рис 1.21). В современной научно-технической практике диапазон этой протяженности так же, как и было уже отмечено выше, сопоставим с интервалом значений характеристических длин порядка 1 – 100 нм. Именно исходя из этого обстоятельства такие структуры получили название *квантовой пленки*. Квантовые пленки, как и квантовые точки и проволоки, могут содержать в своем составе достаточно большое (в пределе бесконечное, при бесконечной протяженности пленки по координатным направлениям $\{x, y\}$) количество атомов.

Беря за основу одномерную модель потенциального ящика для координатного направления $\{z\}$ (см. выражение (1.55)) и модельное представление о движении свободного электрона по координатным направлениям $\{x, y\}$ (см. выражение (1.15)), для функции $W_{(2D)}$ применительно к квантовой пленке имеем (сравнить приводимое ниже выражение(1.81)

$$W_{(2D)}(k_x, k_y, n_z) = \frac{\hbar^2}{8 \cdot m_e} \cdot \left(\frac{n_z^2}{L_z^2} \right) + \frac{\hbar^2}{8 \cdot m_e \cdot \pi^2} \cdot k_x^2 + \frac{\hbar^2}{8 \cdot m_e \cdot \pi^2} \cdot k_y^2, \quad (1.84)$$

где m_e – эффективная масса электрона при его поступательной форме движения по координатному направлению $\{x\}$; L_z – пространственная протяженность квантовой пленки по координатному направлению с ограничением на фактор движения электрона $\{z\}$; n_z – квантовое число из набора целочисленных величин $\{1, 2, \dots, \infty\}$, характеризующих локальное энергетическое состояние электрона, по координатному направлению $\{z\}$; k_x, k_y – проекции волнового вектора поступательной формы движения электрона по координатным направлениям $\{x, y\}$.

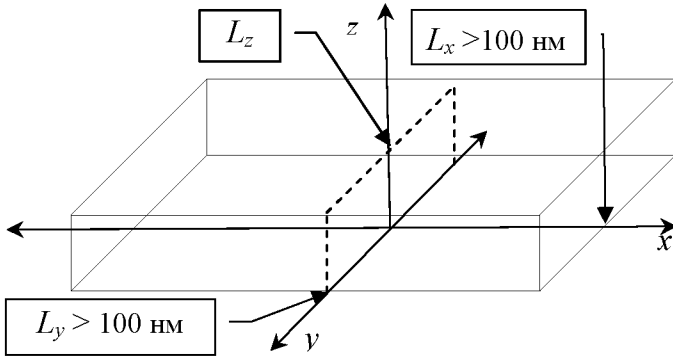


Рис. 1.21. Пространственное отображение квантовой пленки с характеристическими размерными параметрами $\{L_x, L_y, L_z\}$ (L_x, L_y, L_z – микро-, мезо- или макроскопическая величина)

Энергетический спектр (1.84) можно представить как функциональную зависимость от волновых векторов $\{k_x, k_y, k_z\}$ (процедура подобного перехода аргументов представлена выше, при рассмотрении структур типа (0D) и (1D)):

$$W_{(2D)}(k_x, k_y, k_z) = \frac{\hbar^2}{8 \cdot m_e \cdot \pi^2} \cdot (k_x^2 + k_y^2 + k_z^2), \quad (1.85)$$

где волновой вектор $\{k_z\}$ может принимать строго фиксированные численные значения, соответствующие квантовому числу $\{n_z\}$, а волно-

вые вектора $\{k_x, k_y\}$ – соответственно непрерывный ряд значений в диапазоне величин

$$\{(2 \cdot \pi)/L_x - (2 \cdot \pi)/a\} \text{ и } \{(2 \cdot \pi)/L_y - (2 \cdot \pi)/b\},$$

где a и b – параметры кристаллической решетки соответственно по координатным направлениям $\{x, y\}$.

На рис. 1.22 представлено схематическое отображение характерных моментов (фрагментов) энергетических спектров электронов, находящихся в пределах квантовой пленки.

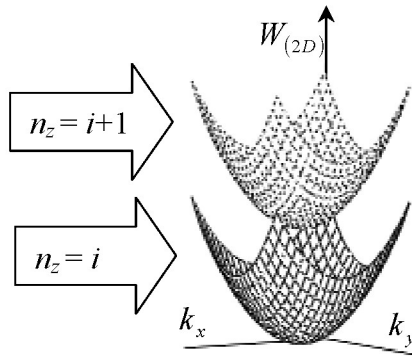


Рис.1.22. Схематическое отображение фрагмента энергетической диаграммы состояния электронов в квантовой пленке

Применительно к наноразмерным системам типа (2D) функция плотности квантовых состояний электронов $g(W)$ в направлениях, связанных с проявлением квантовых ограничений, имеет вид [5]

$$g_{(2D)}(W) = \frac{4 \cdot \pi \cdot m_e}{L_z} \cdot \sum_{k_z} \Theta(W - W_{k_z}), \quad (1.86)$$

где k_z – индекс, пробегающий ряд дискретных значений при изменении квантового числа n_z в диапазоне величин $\{1, 2, \dots, \infty\}$; $\Theta(W - W_{k_z})$ – условное обозначение ступенчато (k_z) – непрерывной (k_x, k_y) функции специальной формы.

На рис. 1.23 представлено схематическое отображение характерных фрагментов функции плотности квантовых состояний электронов $g(W)$, находящихся в геометрических пределах наноразмерных структур типа (2D).

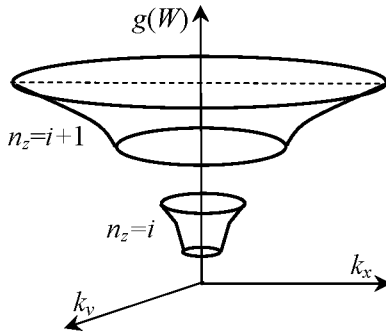


Рис. 1.23. Схематическое отображение фрагмента функции плотности квантовых состояний электронов $g(W)$ для структуры типа (2D)

Из рис. 1.23 видно, что в кристаллических структурах типа (2D) может иметь место появление запрещенной зоны на энергетической диаграмме состояния электронов, если размеры квантовой пленки и по пространственным направлениям $\{x, y\}$ будут в определенной мере ограничены (в этом случае следует обратить внимание на фрагменты $(2 \cdot \pi)/L_x$ и $(2 \cdot \pi)/L_y$).

Ввиду отмеченного выше, своеобразия энергетических диаграмм совокупность электронов в квантовых пленках в ряде случаев принято в научно-технической практике называть *двумерным электронным газом*.

Квантовые пленки уже в настоящее время нашли достаточно широкое практическое применение при создании широкого круга гетеропереходов на основе полупроводниковых и диэлектрических материалов, используемых в элементах конструкций микроэлектронных устройств.

1.3. Физические основы квантовой информации и квантовых компьютеров

1.3.1. Физические основы квантовых вычислительных систем

а) Проблематика и тенденции развития современной вычислительной техники

Ключевые вехи в развитии технического прогресса связаны, прежде всего, с открытием новых ресурсов освоения природы, таких как новые конструкционные материалы и источники энергии. В двадцатом веке, в связи с появлением компьютеров, позволяющих проводить обработку больших информационных массивов вне человеческого мозга, к числу этих ресурсов добавилась *информация*.

История развития компьютерной технологии обработки информации, в свою очередь, состояла из перманентного перехода от простых физических реализаций к более сложным (шестерни → реле → электровакуумные приборы → транзисторы → микросборки → интегральные схемы → сверхбольшие интегральные схемы (СБИС)). На всем этом пути прослеживается стремление к непрерывному уменьшению габаритов вычислительных устройств, вплоть до применения элементов конструкций субатомного масштаба (в настоящее время достигнут технологический уровень производства, позволяющий в промышленных масштабах реализовать минимальный линейный размер элемента конструкции порядка 45 нм).

Следует отметить, что с точки зрения человеческого (макроскопического) восприятия природы классические законы физики являются фактически безупречными. Однако в атомных масштабах, недоступных чувственному восприятию, начинают доминировать совершенно другие, квантовые законы. Квантовые технологии обработки информации призваны дополнить, а при необходимости заменить существующие классические технологии. Уже в настоящее время можно утверждать, что эти (квантовые) технологии могут предложить гораздо больше, нежели тривиальное повышение массово-габаритных показателей вычислительных устройств. Оказывается, что они могут реализовать абсолютно новые принципы вычислительных алгоритмов, не имеющих классических аналогов.

Рассмотрим в более частной постановке проблематику классического подхода при реализации высокопроизводительных вычислительных систем. Согласно современной тенденции развития вычислительной техники тактовая частота микропроцессоров увеличивается примерно вдвое через каждые 18 месяцев (это положение известно в научно-технической практике как «закон Мура»). Увеличение быстродействия на основе методов проектирования и теоретических подходов традиционной планарной микроэлектроники достигается путем соответствующего уменьшения габаритов вычислительных устройств. Поясним это обстоятельство на примере вычислительного устройства, схематически отображенного на рис. 1.24.

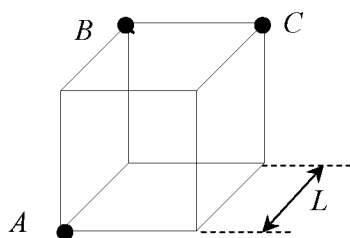


Рис. 1.24. Схематическое отображение вычислительного устройства

Не умаляя общности рассмотрения, будем считать, что вычислительное устройство в целом имеет кубическую форму, а его ребро имеет протяженность L . В точке C расположена вычислительная ячейка, реализующая элементарную вычислительную операцию (например типа: «И»; «ИЛИ») с информационными сигналами, приходящими из ячеек, расположенных в точках A и B . Идеальным, с точки зрения отсутствия сбоев в работе, представляется случай, когда оба сигнала (из точек A и B) приходят в ячейку C одновременно (в частном аспекте под этим следует понимать одновременный приход переднего фронта электрических импульсов).

Для обеспечения синхронизации (по времени или фазе) процессов выхода информационных сигналов из ячеек A и B и их прихода в ячейку C используется генератор тактовых импульсов с частотой ν_0 . Не-

трудно заметить, что при одновременном выходе сигналов из источников A и B сигнал от точки A приходит в точку C позже, чем от точки B . Максимальное расхождение во времени прибытия информационных сигналов составит: $\Delta\tau = (3 \cdot L)/c$, где c – скорость распространения электромагнитной волны в рассматриваемой физической среде.

Если принять линейный размер куба $L = 0.01$ м (отметим, что в таком кубе, при существующем уровне развития технологий, размещается $10^8 - 10^9$ элементов конструкций – структур транзисторного типа), то указанное выше значение расхождения во времени составит по порядку величины: $\Delta\tau = 10^{-8}$ с. Из последнего обстоятельства следует, что для обеспечения временного (фазового) расхождения приходящих в точку C информационных сигналов на уровне не более 10% необходима тактовая частота синхронизации вычислительных процессов порядка $\nu_0 = 10^9$ Гц (в настоящее время это значение представляется скромным показателем и рассматривается возможность перехода к частотам 10^{10} Гц).

Таким образом, можно заключить, что с позиции уровня требований к быстродействию современных вычислительных устройств электромагнитные волны распространяются в пространстве крайне «медленно». Именно ввиду наличия этого ограничительного фактора развитие быстродействующих вычислительных устройств идет по пути уменьшения размеров вычислительных ячеек, выполняющих элементарные логические операции, и сокращения длины элементов коммутации между индивидуальными ячейками (дополнительные численные данные, отражающие динамику процесса повышения степени интеграции элементов конструкций в устройствах вычислительной техники, представлены в табл. 2.1).

Первоначальный интерес к квантовым компьютерам был обусловлен традиционным стремлением использовать специфику квантовых наноразмерных систем с целью уменьшения размеров элементарных логических ячеек, вплоть до субатомного масштаба. Однако при более пристальном изучении указанных квантовых систем выяснилось, что на этом пути существуют и другие резервы повышения быстродействия вычислительных устройств, основанные на использовании специфиче-

ских (не имеющих классических аналогов) квантово-механических эффектов, таких как *квантовый параллелизм* и *квантовая запутанность* (этот круг вопросов рассмотрен подробно ниже).

Выигрыш в производительности (и как следствие, в быстродействии) квантовых алгоритмов обработки информации достигается за счёт того, что при применении одной квантовой операции большое число коэффициентов суперпозиции квантовых состояний, которые в виртуальной форме содержат классическую информацию, *преобразуется одновременно* (классический аналог более скромного порядка – технология параллельных вычислений). В итоге квантовое вычислительное устройство размером L кубит (некий аналог классического понятия – «бит») может выполнять одновременно (параллельно) 2^L вычислительных операций. Этот фактор порождает уверенность в возможности существенного повышения тактико-технических данных вычислительных средств [9]. Эта новая квантовая технология обработки информации находится в настоящее время в теоретическом плане на стадии становления. Достижения в экспериментальном плане подтверждают возможность обеспечения процессов хранения и обработки информации, закодированной в индивидуальных квантовых системах атомного масштаба.

Квантовая теория вычислений играет более фундаментальную роль в части формирования общего мировоззрения на процессы обработки информации, чем ее классический аналог, нашедший широкое распространение в современной компьютерной практике. На ее основе, возможно, будут созданы новые методы обработки таких огромных объемов информации, которые принципиально недоступны существующим классическим методам. В этой связи можно говорить о зарождении новой области научно-прикладных знаний – квантовой теории информации, представляющей собой сплав квантовой физики, теории информации и теории алгоритмов обработки информации в вычислительных средах. Ее охват простирается от установления новых представлений на природу физических законов до практического использования результатов в компьютерной и коммуникационной технике.

Такая глубина проникновения в вопросы взаимосвязи физических и информационных процессов позволит в обозримой перспективе перей-

ти к конструктивному рассмотрению мыслительной деятельности биологических устройств, т.е. приложить полученные научно-практические наработки к проблематике установления механизмов работы мозга как вычислительной среды. В настоящее время мнения ученых существенно расходятся в вопросе о том, насколько методы хранения, отыскания и обработки информации в мозгу соответствуют методам, используемым в классической вычислительной технике. С одной стороны, имеется точка зрения, согласно которой мозг работает по заранее заданным алгоритмам, близким по своей сути к алгоритмам, применяемым в цифровых машинах. С другой же стороны, высказывается мнение, что мозг функционирует не на основе детерминированных алгоритмов, и функции его мало сходны с известными логическими и математическими алгоритмами в цифровых машинах, а наиболее существенными являются вероятностные (в том числе квантово-механические) методы анализа и модели физических сред. Прогресс в этих вопросах открывает новые горизонты в направлении создания специфических биотехнических вычислительных сред, содержащих в своем составе фрагменты известных биологических структур.

Следует отметить принципиальное отличие от ранее существовавшего, современного отношения исследователей к постулатам квантовой механики с позиции существования возможных ограничительных факторов (принцип неопределенности Гейзенберга, вероятностная трактовка волновых функций), накладываемых ею на физические системы. Когда физико-технические аспекты вычислительных операций систематически исследовались в семидесятых годах прошлого века, имелись серьезные опасения в том, что квантово-механические эффекты могут привести к фундаментальным ограничениям на точность и однозначность, с которой реальные физические объекты могут реализовать свойства битов информации, логических ячеек и т.д., составляющих основу абстрактной теории информации. Как выяснилось, эти опасения оказались беспочвенными. Квантовая механика не только не устанавливает никаких ограничений на то, какие классические вычисления могут быть реализованы в природе, а более того, разрешает их все и предоставляет новые способы вычислений, которые не могут быть выполнены с помощью классических схем реализации компьютеров.

Нельзя не упомянуть еще об одном принципиальном аспекте квантовых вычислений, имеющем непосредственное отношение к практике математических доказательств. Считается, что сама процедура проведения вычисления по классической схеме, приводящая к однозначному конечному результату, эквивалентна доказательству того, что *наблюдаемый результат* является одним из возможных вариантов результатов вычислений. Это представляется очевидным лишь в том случае, когда операции компьютера описываются математически строго и в этой связи могут быть переориентированы в доказательство некой математической теоремы.

Такое жестко детерминированное представление вычислительных процессов было уместным в классическом случае (при полном отсутствии интерференционных эффектов, свойственных квантовой механике), *когда можно было пошагово проследить за ходом процесса вычисления*. Подобный шаговый режим контроля позволяет, в принципе, произвести строгое математическое доказательство в виде конечной последовательности предложений, каждое из которых является либо аксиомой, либо следует из предыдущих предложений в соответствии со стандартными правилами формирования логических умозаключений (*детерминированная логика*). К сожалению, в настоящее время приходится отказаться от такого привлекательного в логическом плане представления и рассматривать сам *процесс квантового вычисления как форму доказательства* математических концепций и положений. Это заключение следует из того, что в будущем квантовые компьютеры будут доказывать теоремы такими специфическими методами, которые человеческий мозг (представляющий собой макроскопическую систему) не будет в состоянии проверить, следуя традиционному методу «шаг за шагом» [9].

б) Когерентная квантовая интерференция отдельных микрочастиц

Ключевым фундаментальным понятием классической теории информации является представление об элементарной единице информации – «бит». С точки зрения физико-технической реализации понятие «бит» отражает систему, которая может находиться в двух *устойчивых*

и различных состояниях. Эти состояния должны быть разделены высоким потенциальным барьером, чтобы исключить возможность спонтанных переходов системы из одного состояния в другое (например, под воздействием различных дестабилизирующих факторов флуктуационного характера, возникающих в процессе штатных режимов эксплуатации изделия).

В качестве типичного представителя единичного элемента такой информационной системы можно рассматривать двухпозиционный (бинарный) механический переключатель, в котором отдельным его позициям предписывается строго определенное состояние, условно обозначаемое посредством ограниченного набора символьных позиций, например: «0» – «1»; «нет» – «да»; «верно» – «неверно»; «А» – «Б»; «закрыто» – «открыто» и т.д.

В принципе аналогичным переключателем, но атомного масштаба, можно считать и единичный атом, если он может находиться в двух различных квантовых состояниях. Однако в связи с тем, что данная система является типичным представителем квантово-механических объектов, здесь появляется определенный нюанс, получивший название *когерентной суперпозиции* квантовых состояний. Под этим понимается возможность возникновения новых состояний, которые не имеют классических аналогов и в которых система *одновременно представляет* оба состояния – «0» и «1». Именно в этом квантовом феномене состоит принципиальное отличие элементарной единицы квантовой информации «кубит» от традиционной единицы информации «бит», нашедшей широкое распространение в классических вычислительных устройствах. В качестве носителя «кубита» может выступать практически любая квантовая система, имеющая, по меньшей мере, два взаимодополняющих состояния (т.е. удовлетворяющая условию нормировки вероятностей появления отдельных исходов). Можно предложить множество вариантов практической реализации подобных систем, часть из которых уже успешно прошла экспериментальную апробацию [9].

Такая ситуация, реализующая «кубит» в квантовой системе, при которой физический параметр может принимать одновременно два разных значения, настолько непривычна (с позиций традиционной класси-

ческой физики), что целесообразно пояснить этот принципиально важный момент на примере простого мысленного эксперимента, схема которого представлена на рис. 1.25,а.

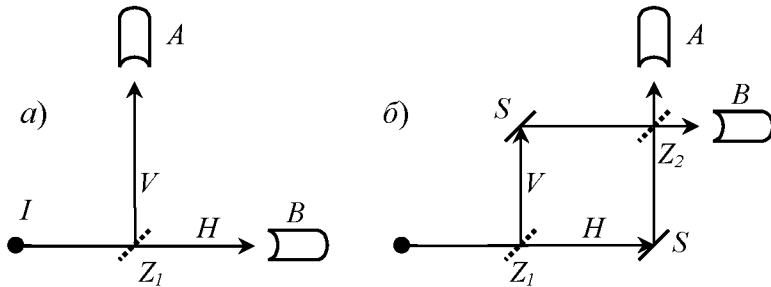


Рис. 1.25. Схематическое отображение интерференции одной микрочастицы

Пусть фотон (в принципе это может быть любая микрочастица) вылетает из источника I и попадает на полупрозрачное зеркало Z_1 , после чего появляются две возможные траектории его последующего движения – V и H . Что можно сказать точно, так это то, что фотон не расщепляется на зеркале на два объекта и будет зарегистрирован с *равной вероятностью* одним из детекторов A или B . Получается, что фотон пролетает по направлениям V или H случайным образом (в соответствии с постулатами классической физики). Однако в действительности это не так! Может показаться очевидным, что фотон в *каждом конкретном эксперименте* находится или на пути V , или на пути H . Но это тоже не соответствует действительности. На самом деле фотон одновременно движется по двум путям V и H , что может быть продемонстрировано с помощью эксперимента, схема которого отображена на рис. 1.25,б.

В этом новом эксперименте два обычных зеркала S расположены так, что оба пути следования фотона (V и H) пересекаются на втором полупрозрачном зеркале Z_2 . Несколько забегаю вперед, отметим, что с помощью оптической схемы рис. 1.25,б можно наблюдать весьма неожиданный, исключительно квантовый эффект – *одночастичной интерференции*.

Предположим, что конкретный фотон (см. рис. 1.25,б) двигался после прохождения полупрозрачного зеркала Z_1 вдоль пути H . Тогда из

сравнения с рис. 1.25,а следует, что оба детектора должны регистрировать фотон с одинаковой вероятностью. То же самое наблюдалось если бы фотон двигался вдоль пути V . Следовательно, если бы фотон двигался строго по одному пути (неважно V или H), каждый из детекторов A и B срабатывал бы в половине случаев прохождения фотона. Однако в реальности наблюдается совершенно иная картина: фотон *всегда* попадает на детектор A и *никогда* не попадает на детектор B .

Если один из путей (V или H) перекрыть непроницаемым экраном, то попадание фотона в детекторы A и B становится равновероятным. Отсюда следует неизбежный вывод, что фотон перемещается одновременно по двум путям.

Таким образом, *перекрытие любого из путей* (рис. 1.25,б) *приводит к освещению детектора B* . Когда же оба пути открыты (V и H), фотон неведомым способом получает информацию, которая не позволяет ему однозначно попасть в детектор B . *Эта информация распространяется вдоль другого пути со скоростью света*, отражаясь от зеркал точно так же, как распространяется фотон. Это свойство *квантовой интерференции*, которое можно объяснить существованием гипотетических невидимых «двойников» фотона, оказывающих влияние на движение наблюдаемых фотонов, можно распространить и на все типы известных микрочастиц.

Для последующего рассмотрения важно отметить тот факт, что *вероятностный* характер исхода первой оптической системы (рис. 1.25,а) полностью устраняется с помощью второй оптической схемы (рис. 1.25,б), обеспечивающей предельно *детерминированное* поведение микрочастицы. Проблематика квантовых вычислительных процессов непосредственно связана с созданием таких условий, при которых обеспечивается синхронизация индивидуальных эффектов квантовой интерференции (протекающих с участием упомянутых выше, не наблюдаемых воочию, «двойников» микрочастиц) для случая более чем одной микрочастицы.

С точки зрения реализации квантовых вычислительных процессов, весьма показательным является эксперимент, схематическое отображение которого представлено на рис. 1.26. По своей сути это тот же эксперимент, что и отображенный на рис. 1.25,б, с тем лишь отличием, что

на одном из путей движения фотона расположен элемент задержки во времени в виде тонкой стеклянной пластинки P заданной толщины. Так как исход процесса интерференции между фотоном и его невидимым «двойником» зависит от точных времен их прибытия в детектор, то можно подобрать толщину стеклянной пластинки (и соответственно время задержки движения фотона) таким образом, чтобы фотон достоверно оказывался в детекторе B , а не как это было ранее в детекторе – A . Таким образом можно реализовать систему, в которой состояние квантовой системы становится управляемым, т.е. приобретает детерминированный характер в части обеспечения необходимого исхода. Последнее обстоятельство представляет большой практический интерес.

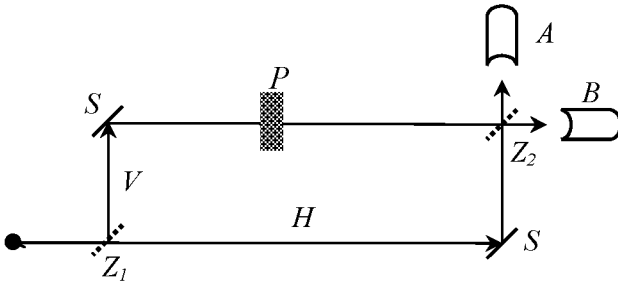


Рис. 1.26. Схематическое отображение квантовой интерференции с элементом задержки во времени

Точно так же, как фотон может находиться в когерентной суперпозиции оптических путей (V и H), так и «кубит» может быть *синтезирован* в виде суперпозиции двух логических (квантовых) состояний, условно представленных в виде «0» и «1». Отметим еще раз суть принципиального различия между «битом» и «кубитом». В «бите» состояние системы *однозначно определено*: «0» или «1». В «кубите» же состояние системы «размазано» в том смысле, что в нем одновременно содержатся оба состояния «0» и «1», представленные в *известной пропорции*. Масштаб представительства состояний «0» и «1» в «кубите» описывается исключительно на основе вероятностных представлений. Именно с этими значениями вероятностей путь движения фотона будет обнару-

жен детекторами A или B в оптической схеме, представленной на рис. 1.25,а. В этом смысле проявление вероятностного способа описания состояния системы можно рассматривать в качестве негативного сдерживающего психологического фактора на пути построения квантовых вычислительных систем.

Не имеет никакого смысла сетовать на вероятностный характер представления состояния квантово-механических систем, поскольку это представление отслеживает природу физической картины мира. Скорее всего, следует принять во внимание ту точку зрения, что реалии физических проявлений намного обширнее, чем те, которые доступны непосредственному наблюдению.

в) Общие принципы реализации квантовых вычислительных систем

Любой физический процесс, использующий квантовые эффекты для выполнения определенных вычислительных операций, является по своей сути *квантовым алгоритмом* обработки информации. Отличительная особенность квантовых алгоритмов от классических алгоритмов состоит в том, что в квантовом случае:

- элементы памяти в компьютере содержатся не в форме записи в традиционных информационных единицах «бит», а в форме «кубит», обладающих свойством *когерентной одночастичной интерференции*, что приводит к вероятностной интерпретации параметров состояния ячейки памяти;

- способ обработки информационных массивов в своей методологической основе автоматически реализует принцип параллельных вычислений путем *суперпозиции чисел*, что позволяет рассчитывать на существенное повышение, в перспективе, параметров быстродействия и производительности средств вычислительной техники;

- каждый конкретный «кубит» в квантовой вычислительной системе, содержащей совокупность взаимосвязанных «кубитов», не несет сам по себе никакой информации, пока один из них не будет подвергнут процедуре измерения его состояния в базисе « $0 - 1$ ». (Т.е. индивидуальные «кубиты» не обладают изначально никаким смысловым содержанием, аналогичным понятию «бит»);

- в качестве основы математического аппарата при анализе логических операций используются *унитарные преобразования*, а не Булева алгебра (Преобразование считается унитарным, если оно сохраняет норму. В упрощенной интерпретации речь идет о соблюдении принципа нормировки параметров, характеризующих вероятность появления индивидуальных исходов (квантовых состояний): $\sum P_i = 1$, где P_i – вероятности появления событий (исходов), составляющих полную группу).

С учетом вышеизложенных моментов квантовый компьютер можно рассматривать как своего рода «*программируемый интерферометр*», в котором различные вычислительные пути (в частности, в этом качестве выступают оптические пути типа, представленного на рис. 1.25) спланированы таким образом, что их управляемая интерференция приводит к однозначному конечному результату.

В развитие идеи об использовании фактора суперпозиции чисел, который выступает в качестве одного из ключевых позитивных моментов квантовых вычислительных устройств, рассмотрим конкретный пример реализации 3-«битного» регистра памяти на основе классических и квантовых представлений. Классический 3-битный регистр памяти может находиться в одном из восьми устойчивых состояний и тем самым хранить в точности *одно* из восьми различных чисел – {000, 001, 010, 011, ..., 111}. В квантовом 3-битном регистре памяти может находиться *одновременно* до восьми чисел в состоянии квантовой суперпозиции элементарных состояний. При добавлении «кубитов» в регистр памяти его емкость, с точки зрения объема хранения информации, возрастает экспоненциально. В общем случае N «кубитов» могут хранить *одновременно* до 2^N чисел. Таким образом, регистр памяти, выполненный всего лишь на основе 50 «кубитов», может хранить одновременно количество чисел, эквивалентное объему памяти современных (классических) компьютеров 100 Гб. (При $N = 250$ регистр памяти может хранить одновременно количество чисел, превосходящее количество атомов в видимой части Вселенной!).

С точки зрения специфики методологии выполнения квантовых вычислительных операций следует отметить следующие характерные моменты:

- обращение к регистру памяти квантового компьютера выделяет одно число из всей совокупности хранящихся в ней 2^N чисел (информационных единиц), находящихся в состоянии квантовой суперпозиции;
- нетривиальные квантовые вычислительные операции, основанные на принципах когерентной интерференции, производятся одновременно на всем массиве, хранящихся в памяти компьютера чисел.

Как упоминалось выше, индивидуальные «кубиты» выполняются на основе физических объектов атомного масштаба. Перевод их состояния (в частности энергетического), условно представляемого в символах (0 – 1), осуществляется посредством использования факторов внешнего возмущающего воздействия (например – лазерными импульсами). В результате такого воздействия начальная суперпозиция закодированных чисел переходит в новую суперпозицию. В процессе эволюции *каждое число* из общего количества совокупности чисел в квантовой суперпозиции одновременно подвергается управляющему воздействию, что обеспечивает проведение значительного объема вычислений за один шаг работы в *параллельном режиме* обработки информации.

Таким образом, квантовый компьютер может за один шаг вычислений произвести операцию над массивом объемом в 2^N чисел, результатом которой будет новая суперпозиция всех этих чисел, представляющих совокупность выходных данных процесса вычисления. Для того чтобы выполнить аналогичную задачу на компьютере классического типа, ему необходимо последовательно произвести 2^N вычислительных операций или использовать (для обеспечения одношагового режима обработки всего массива начальных и промежуточных информативных данных) такое же количество параллельно работающих процессоров. Именно исходя из вышеизложенных моментов следуют те ожидаемые преимущества в сокращении времени обработки информации и необходимого объема памяти при использовании квантовых вычислительных устройств.

Структуры атомного масштаба, реализующие «кубиты», являются очень уязвимыми с точки зрения обеспечения их помехоустойчивости к

воздействию внешних неблагоприятных факторов окружающей среды. Эти факторы могут вызвать дестабилизацию процессов когерентной интерференции (*квантовая декогерентность*) и тем самым полностью нарушить как вычислительные операции, так и сохранность информации в виде квантовой суперпозиции состояний «кубита». Фактор декогерентности состояния квантовой системы можно представить как следствие своеобразного «перепутывания» квантовых состояний рассматриваемой системы с квантовыми состояниями окружающей среды (с позиции классической физики, выступающей в качестве термостата).

Проблематика помехоустойчивости информационных и измерительных средств имеет фундаментальный характер и является актуальной и для классических компьютеров, однако ее масштабы в этом случае будут гораздо меньше, чем для квантовых компьютеров. Последнее находит свое объяснение в том, что технические устройства, реализующие понятие «бит», по своей сути (несмотря на малые размеры), являются все же макроскопическими системами (к примеру, транзисторная структура с линейным размером 0.1 мкм содержит в своем составе порядка 10^9 атомов). Ввиду того обстоятельства, что элементы конструкций классических компьютеров имеют, как правило, макроскопические объемы, то у них энергия управляющих воздействий, на много порядков превосходит энергию шумовой (дестабилизирующей) составляющей, сравнимой с энергетическим масштабом тепловых флуктуаций: $kT = 0.025$ эВ (при $T = 300$ К). В квантовых компьютерах в качестве исполнительных устройств, реализующих «кубиты», выступают индивидуальные микрочастицы (электрон, ион, атом, фотон и т.д.), и в этом случае влияние шумовой составляющей может быть весьма значительным (поэтому неслучайно, что ряд экспериментальных апробаций элементов конструкций квантовых компьютеров производится при температурах, близких к абсолютному нулю).

Эффективным направлением решения проблемы декогерентности является использование методов квантового *исправления* или *коррекции* набегающих ошибок. Они основаны на том принципе, что состояние единичного «кубита» можно закодировать в перепутанных состояниях нескольких «кубитов» (этот подход близок по своей принципиальной сути к методам резервирования, достаточно широко используемым в

классических компьютерах). Однако в случае квантовых компьютеров, в связи с большей чувствительностью самих перепутанных состояний к фактору декогерентности, необходим более тонкий компромисс между процессами неизбежного добавления дополнительных ошибок и их исправления.

Существует ряд других альтернативных способов борьбы с проявлением процесса декогерентности квантовых систем, но они имеют достаточно сложную физическую природу и их понимание требует специальную физико-математическую подготовку (гораздо шире традиционных курсов подготовки инженерных кадров технических направлений). В этой связи они не рассматриваются в рамках настоящего пособия.

г) Символьный формализм представления квантовых состояний и процессов

Как показано выше, при рассмотрении физических основ квантовой механики в качестве основополагающих представлений фигурируют амплитуда вероятности ПМВ де Бройля и волновые функции $\Psi(x, y, z, t)$ (см. разд. 1.1.). Поскольку в дальнейшем изложении эти волновые представления используются достаточно широко, в целях удобства рассмотрения будет применено их сокращенное символьное обозначение – «скобки Дирака». Использование скобок Дирака приводит к существенному сокращению формы записи и повышению наглядности представления промежуточных и конечных результатов анализа.

Рассмотрим сущность и методологию практического использования скобок Дирака [1] на ряде характерных примеров, представляющих интерес в свете тематики квантовых компьютеров.

Пример № 1. «Прохождение одной микрочастицы через экран с одной целью».

Допустим, что микрочастица испускается из точки s и в конечном итоге попадает в точку x . Согласно общим принципам квантовой механики, для описания этого процесса необходимо определить *амплитуду вероятности* сложного события – «частица покидает s и частица попадает в x ». В символьном отображении с помощью скобок Дирака этот

же процесс (как совокупность независимых событий) отображается в виде

$$\varphi(s, x) = \left\langle \begin{array}{l} \text{Частица попадает в точку } x \mid \text{Частица} \\ \text{покидает точку } s \end{array} \right\rangle. \quad (1.87)$$

Следует отметить, что выражение (1.87) представляет собой всего-навсего *некоторое комплексное число, квадрат которого равен вероятности реализации данного сложного события*. Для «расшифровки» смыслового значения выражения (1.87) его следует читать справа – налево, т.е. микрочастица сначала покидает точку s , а затем переходит в точку x .

Скобки $\langle \rangle$ в выражении (1.87) являются символьным эквивалентом тривиальной фразы – «амплитуда вероятности того, что микрочастица вылетает из точки s и попадает в точку x ». Выражение *справа* от вертикальной черты всегда задает *начальное условие* рассматриваемого процесса, а выражение *слева* – *конечное условие* процесса. В более лаконичном представлении амплитуду вероятности, отображенную в (1.87), принято записывать в виде

$$\langle x|s \rangle. \quad (1.88)$$

Если на пути движения микрочастицы имеется достаточно узкая щель, через которую она должна непременно пройти для последующего попадания в точку x , то весь путь микрочастицы можно разбить на два независимых участка: участок « s – щель»; участок «щель – x » (рис. 1.27).

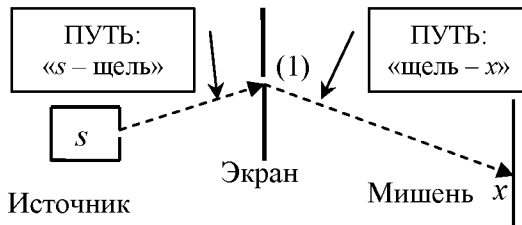


Рис. 1.27. Прохождение микрочастицы через одиночную щель

Когда микрочастица перемещается по такой определенной двухступенчатой траектории, то амплитуда (т.е. вероятность) для этого пути будет равна *произведению амплитуд* отдельных участков траектории:

$$\langle x | s \rangle_{(1)} = \langle x | 1 \rangle \cdot \langle 1 | s \rangle, \quad (1.89)$$

где $\langle 1 | s \rangle$ – амплитуда прохождения участка пути «источник – щель»; $\langle x | 1 \rangle$ – амплитуда прохождения участка пути «щель – мишень».

Степень детализации описания процесса может быть расширена путем включения в рассмотрение более полного круга факторов, стоящих на пути реализации процесса прохождения микрочастицы. К примеру, выражение (1.89) можно представить и в таком развернутом виде:

$$\langle x | s \rangle_{(1)} = \langle x | 1 \rangle \cdot \langle 1 | 1 \rangle \cdot \langle 1 | s \rangle, \quad (1.90)$$

где $\langle 1 | 1 \rangle$ – амплитуда прохождения микрочастицей щели в экране (см. рис. 1.27).

Пример № 2. «Прохождение одной микрочастицы через экран с двумя щелями».

Согласно [1] следует тот факт, что если микрочастица может достигнуть некоторой точки пространства с помощью двух различных независимых путей (условно «1» и «2») прохождения, то вероятность P_{1-2} ее обнаружения в данной точке равна *квадрату модуля от суммы амплитуд вероятностей* этих вариантов путей. Следует отметить, что исходя из представлений классической физики, следует ожидать сумму вероятностей прохождения отдельных вариантов путей. Физическую сущность этого расхождения составляет учет явления интерференции индивидуальных вероятностных волн де Бройля:

$$P_{1-2} = |\varphi_1(x, y, z, t) + \varphi_2(x, y, z, t)|^2. \quad (1.91)$$

С помощью символов Дирака для случая двух щелей на экране амплитуду процесса перехода микрочастицы от точки s к точке x можно отобразить в следующей лаконичной форме:

$$\langle x | s \rangle_{(1-2)} = \langle x | s \rangle_1 + \langle x | s \rangle_2, \quad (1.92)$$

или, с учетом выражения (1.90), можно представить этот процесс и в более развернутом (детализированном) представлении:

$$\begin{aligned} \langle x|s \rangle_{(1-2)} = & \langle x|1 \rangle \cdot \langle 1|1 \rangle \cdot \langle 1|s \rangle + \\ & + \langle x|2 \rangle \cdot \langle 2|2 \rangle \cdot \langle 2|s \rangle. \end{aligned} \quad (1.93)$$

В зависимости от целевых постановок практических задач, процесс детализации рассмотрения может быть продолжен дальше. В частности, принимая во внимание фактор конечной протяженности щелей, можно поставить вопрос о конкретизации интервала щели, через который проходит микрочастица.

Пример № 3. «Прохождение n микрочастиц через экран с одной щелью».

Пусть имеется n не взаимодействующих между собой микрочастиц, которые испускаются из различных точек пространства s_i (где i – натуральное число, пробегающее ряд значений в диапазоне величин $\{1, \dots, n\}$) и, проходя через одну щель на экране, попадают соответственно в различные точки пространства x_i . В этом случае амплитуда вероятности сложного процесса, состоящего из совокупности отдельных независимых процессов перемещения отдельных микрочастиц, будет равна произведению амплитуд вероятности отдельных процессов:

$$\langle x_1, x_2, \dots, x_n | s_1, s_2, \dots, s_n \rangle_{(1)} = \prod_{i=1}^n [\langle x_i | 1 \rangle \cdot \langle 1 | s_i \rangle]. \quad (1.94)$$

В выражении (1.94) фактор отсутствия взаимодействия между микрочастицами нашел свое отражение в отсутствии интерференционных составляющих.

Пример № 4. «Представление квантового состояния единичного кубита».

В предыдущих примерах рассмотрено использование формализма скобок Дирака для описания не только состояний системы (начального и конечного), но и отображения пути ее эволюции в процессе перехода от начального состояния к конечному состоянию системы. В рамках задач вычислительной техники используется более узкая постановка вопроса, а именно представление только информационного состояния системы (элементарной ячейки памяти).

Информационное состояние ячейки памяти в современных вычислительных устройствах описывается в «битовом» представлении, что предполагает *однозначную* трактовку состояния системы, условно отображаемую символами «1» или «0». Эта однозначность при определении состояния информационной ячейки обусловлена тем обстоятельством, что эти ячейки памяти являются (по физической сути используемых принципов записи информации) объектами рассмотрения классической физики, основанной на детерминированных подходах.

В случае использования в качестве носителя информации «кубита» указанная однозначность состояния ячейки памяти не имеет места ввиду специфики квантово-механических принципов записи информации. Поскольку кубит является типичным представителем квантово-механических объектов (здесь уместно напомнить как о принципе неопределенностей Гейзенберга, так и вероятностном (волновом) представлении квантовых объектов и процессов), его состояние может быть описано на основе вероятностных представлений.

Если состояние кубита (условно «0» или «1») определено со сто-процентной достоверностью (предельный случай), то такое идеализированное состояние рассматривается как *базисное состояние*, и его принято обозначать в символьном виде с помощью модифицированных скобок Дирака, соответственно $|0\rangle$ или $|1\rangle$. В общем случае (т.е. в реальных квантово-механических системах) состояние кубита $|Q\rangle$ располагается где-то между идеализированными базовыми состояниями $|0\rangle$ и $|1\rangle$ и в этой связи его принято представлять в символьном виде:

$$|Q\rangle = a \cdot |0\rangle + b|1\rangle, \quad (1.95)$$

где $|a|^2$ и $|b|^2$ – соответственно вероятности обнаружения (посредством проведения процедуры диагностики состояния кубита) альтернативных исходов: $|0\rangle$ или $|1\rangle$.

Поскольку базовые состояния $|0\rangle$ и $|1\rangle$ составляют полную группу исходов при проведении процедуры диагностики состояния кубита, то

параметры a и b должны удовлетворять очевидному условию нормировки:

$$|a|^2 + |b|^2 = 1. \quad (1.96)$$

Важно отметить, что после проведения процедуры диагностики состояния кубита он может перейти во вполне определенное квантовое состояние, т.е. происходит устранение начальной неопределенности и поэтому кубит переходит в разряд информационных объектов, описываемых подходами классической физики. Таким образом, можно считать, что сам факт проведения процедуры «правильной» диагностики состояния кубита может привести к *устранению начальной неопределенности* физической системы, на основе которой реализован кубит.

Для иллюстрации этого момента рассмотрим конкретный пример кубита. Пусть начальное квантовое состояние кубита (т.е. до проведения процедуры диагностики его состояния) описывается с помощью выражения

$$|Q\rangle = \frac{2}{6} \cdot |0\rangle + \frac{4}{6} |1\rangle, \quad (1.97)$$

где $\frac{2}{6} = a$; $\frac{4}{6} = b$ (см. выражение (1.95)).

Запись вида (1.97) отражает тот факт, что базисное состояние $|0\rangle$ представлено в кубите с вероятностью $\left(\frac{2}{6}\right)^2 \approx 11\%$, а базисное состояние $|1\rangle$ – с вероятностью $\left(\frac{4}{6}\right)^2 \approx 89\%$. После проведения процедуры

диагностики может получиться так, что кубит перейдет в новое квантовое состояние, которое будет определяться выражением

$$|Q\rangle = 1 \cdot |0\rangle + 0 \cdot |1\rangle, \quad (1.98)$$

т.е. при проведении последующей процедуры диагностики базовое состояние $|0\rangle$ будет обнаружено со 100 % вероятностью, а базовое состояние $|1\rangle$ будет иметь нулевую вероятность обнаружения.

Пример № 5. «Представление квантового состояния системы из n кубитов».

Рассмотрим общие принципы анализа информационной емкости и n кубитов на примере простой системы, состоящей всего из двух кубитов. Согласно вышеизложенному, диагностика квантового состояния каждого из кубитов позволит в итоге определить его состояние в классическом приближении «0» или «1». В целом у системы из двух кубитов имеется четыре классических состояния («00», «01», «10», «11»), которым можно поставить в соответствие четыре квантовых базисных состояния: $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. Таким образом, общее квантовое состояние системы, состоящей из двух кубитов, можно выразить в виде суммы базисных состояний:

$$|Q\rangle = a \cdot |00\rangle + b \cdot |01\rangle + c \cdot |10\rangle + d \cdot |11\rangle, \quad (1.99)$$

где $|a|^2, |b|^2, |c|^2, |d|^2$ – соответственно вероятности обнаружения (до проведения процедуры диагностики) базовых состояний квантовой двухкубитной информационной системы: $|00\rangle, |01\rangle, |10\rangle, |11\rangle$.

После проведения процедуры диагностики состояния рассматриваемой системы ее квантовое состояние может быть представлено в виде (приведен один из возможных вариантов исхода процедуры диагностики):

$$|Q\rangle = 0 \cdot |00\rangle + 1 \cdot |01\rangle + 0 \cdot |10\rangle + 0 \cdot |11\rangle. \quad (1.100)$$

Полученные результаты можно без труда распространить на случай информационной системы (блока памяти вычислительного устройства), содержащей в своем составе n кубитов. Необходимо отметить, что процедура диагностики затрагивает сразу весь массив исходных данных емкостью 2^n , находящихся в вычислительной среде. Последнее обстоятельство обеспечивает беспрецедентный параллелизм вычислительных операций, который не имеет своего аналога в классических системах обработки информации. Этот параллелизм открывает новые возможности в повышении ключевых тактико-технических показателей вычислительной техники, в частности, он позволяет:

- существенно повысить быстродействие вычислительных процессов при обработке чрезвычайно больших массивов информации

ных потоков, а следовательно, повысить параметры производительности вычислительных систем;

- существенно *снизить количество индивидуальных ячеек памяти*, необходимых для хранения информации, а следовательно, и для практической реализации блоков памяти перспективных высокопроизводительных вычислительных систем.

Приведенные выше примеры позволяют перейти к рассмотрению общих принципов и концепции квантовых вычислительных процессов.

д) Суперпозиция квантовых состояний как основа вычислительных операций

Общая теория алгоритмов квантовых вычислительных процессов находится на стадии своего становления. В настоящее время можно лишь с уверенностью говорить только о некоторых общих подходах к решению этой проблематики и ряде оригинальных алгоритмов частного характера, которые, тем не менее, способны пролить свет на концептуальную основу квантовых вычислительных процессов.

Рассмотренный выше принцип когерентной квантовой интерференции (суперпозиции) квантовых состояний играет ключевую роль как в вопросах постановки большинства мысленных экспериментов квантовой механики, так и квантовой информации и реализации вычислительных алгоритмов. С позиции рассматриваемой проблематики к подобному мысленному эксперименту следует отнести эксперимент с «двумя щелями». Необходимыми составляющими этого базового эксперимента являются: источник микрочастиц, экран с двумя щелями и экран, на котором формируется интерференционная картина. Физическая природа этой интерференционной картины отражает волновые квантово-механические свойства микрочастиц, испускаемых источником. С позиции практических интересов создания квантовых вычислительных устройств важно отметить, что эффект интерференции наблюдается для микрочастиц различного типа: фотоны, электроны, атомы, нейтроны. С точки зрения представлений квантовой механики, рассматриваемая интерференционная картина является примером представленной выше *когерентной интерференции* (см. пункт (б) настоя-

щего раздела), которая может быть отображена с помощью волнового представления в виде

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\psi_1\rangle + |\psi_2\rangle), \quad (1.101)$$

где фрагменты $|\psi_1\rangle$ и $|\psi_2\rangle$ – описывают квантовые состояния в том случае, если открыта только щель (1) или щель (2); $\left(\frac{1}{\sqrt{2}}\right)^2 = 0.5$ – от-

ражает тот факт, что вероятности прохождения микрочастицей щелей (1) и (2) приняты одинаковыми (50 %).

Важным обстоятельством в эксперименте с двумя щелями является тот факт, что интерференционная картина наблюдается и в случае прохождения всего лишь одной микрочастицы (*одночастичная интерференция* – при которой частица «интерferирует» сама с собой). Этот феномен не имеет классического аналога и в этой части расходится со «здравым смыслом», который, воспринимая ее в качестве локализованного объекта, пытается ответить на вопрос – «через какую же щель проходит микрочастица?». С позиции квантовой механики эксперимент, поставленный с целью ответить на этот вопрос, приводит к взаимодействию «измерительного (диагностического) прибора» с объектом, в результате которого происходит эффект *декогерентности*, т.е. к потере интерференции. Таким образом, ответ на поставленный вопрос устраняет *начальную неопределенность* состояния квантовой системы и переводит ее в разряд *детерминированных* объектов классической физики. Последнее обстоятельство представляется важным с позиции взаимодействия макроскопического наблюдателя (оператора) с вычислительным устройством, использующим квантово-механические принципы обработки информации.

В этой связи рассмотренное выше символьное представление кубита (см. выражение (1.95)) описывает *когерентную суперпозицию* квантовых состояний, для которых существует понятие *базиса* ($|a|^2 + |b|^2 = 1$), а не просто «механическую» смесь независимых состояний – (0) или (1).

Рассмотрим процесс реализации декогерентности состояния квантовой системы на конкретном примере использования *преобразования Адамара* – одного из основных алгоритмических приемов в науке о квантовой информации [9]. Пусть имеется одиночный кубит, у которого начальное информативное состояние определяется выражением

$$|Q\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle). \quad (1.102)$$

Как показано выше (см. выражение (1.101)), подобная форма описания кубита означает, что индивидуальные состояния (0) и (1) представлены в кубите с вероятностью обнаружения 50 %. По своей сути преобразование Адамара является специфическим оператором H , воздействие которого на кубит проявляется следующим образом:

$$\begin{aligned} H|0\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \\ H|1\rangle &\rightarrow \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \quad (1.103)$$

Знак минус во втором выражении не должен вызывать недоразумение, поскольку физическим смыслом обладает только квадрат амплитуды $1/\sqrt{2}$.

Применим преобразование Адамара в форме (1.103) непосредственно к самому кубиту (1.102):

$$\begin{aligned} H|Q\rangle &\rightarrow \frac{1}{\sqrt{2}} \cdot (H|0\rangle + H|1\rangle) = \\ &= \frac{1}{\sqrt{2}} \cdot (|0\rangle + |1\rangle) + \frac{1}{\sqrt{2}} \cdot (|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|0\rangle. \end{aligned} \quad (1.104)$$

Из выражения (1.104) следует, что в результате воздействия преобразования Адамара на кубит последний (с *неопределенным* начальным состоянием) переходит в новое конечное квантовое состояние, которому соответствует 100% содержание состояния $|0\rangle$, т.е. имеет место *полная определенность* состояния кубита.

Следует обратить особое внимание на тот факт, что к конечной форме выражения (1.104) *неприменимы обычные «арифметические» операции*, и в частности сложение математических фрагментов:

$$\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|0\rangle \neq \frac{2}{\sqrt{2}}|0\rangle. \quad (1.105)$$

Последнее обусловлено тем обстоятельством, что это выражение отражает всего лишь фактор суперпозиции индивидуальных квантовых состояний. В этой связи смысловое содержание выражения (1.104) можно трактовать следующим образом: «да, действительно имеют место два независимых исхода измерения, но они оба имеют одинаковый конечный результат».

*е) Общие физико-технические принципы преобразования
информационного состояния кубитов*

Рассмотрим общие физико-технические принципы преобразования информационного состояния кубитов на примере ряда базовых элементов, составляющих основы квантовых вычислительных сред.

Базовый элемент на основе деления входных сигналов

Рассмотрим одну из базовых экспериментальных операций в физике квантовой информации на примере простого делителя света, который делит луч света в отношении 50/50. В качестве такого делителя используется полупрозрачное зеркало, которое с вероятностью 50 % может как отразить, так и пропустить наклонно падающий на зеркальную поверхность единичный фотон (отметим, что такие делители в настоящее время реализованы не только для фотонов, но и для широкого круга других микрочастиц).

Для общности рассмотрения возьмем случай *двух наклонно входящих* в делитель фотонов (падающих сверху и снизу относительно зеркальной плоскости), как это представлено на рис. 1.28.

Микрочастица, падающая по наклонной (относительно полупрозрачной зеркальной поверхности делителя «50/50») траектории на делитель сверху или снизу, может в дальнейшем появиться либо в верхнем, либо в нижнем полупространстве, с одной и той же вероятностью ее обнаружения, а именно, в точности равной 50 %. Подобрал соответ-

ствующим образом фазовое действие делителя, можно таким образом реализовать процесс прохождения микрочастицы, чтобы выполнялась процедура преобразования Адамара (см. выражение (1.103)).

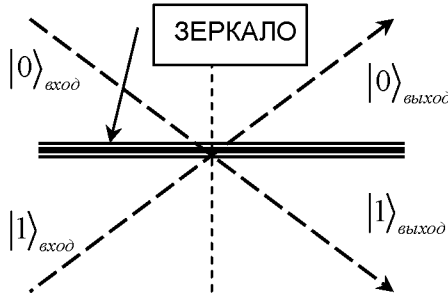


Рис.1.28. Схематическое отображение делителя 50/50

Для случая одной частицы, состояние на входе кубита, реализующего делитель оптического луча «50/50», можно представить в виде

$$|Q\rangle_{\text{вход}} = \alpha \cdot |0\rangle_{\text{вход-сверху}} + \beta \cdot |1\rangle_{\text{вход-снизу}}, \quad (1.106)$$

где α и β – соответственно амплитуда вероятности обнаружения частицы, падающей на делитель сверху и снизу.

В результате действия делителя получается конечное состояние кубита, отображаемое в виде

$$|Q\rangle_{\text{выход}} = H|Q\rangle_{\text{вход}} = \frac{1}{\sqrt{2}} \cdot [(\alpha + \beta) \cdot |0\rangle_{\text{выход-сверху}} + (\alpha - \beta) \cdot |1\rangle_{\text{выход-снизу}}]. \quad (1.107)$$

В соответствии с выражением (1.107) можно сформулировать ряд практически важных выводов:

- амплитуда вероятности обнаружить прошедший (выходящий) фотон в верхнем полупространстве составляет $(\alpha + \beta)$;
- амплитуда вероятности обнаружить прошедший (выходящий) фотон в нижнем полупространстве составляет $(\alpha - \beta)$;
- если амплитуды удовлетворяют условиям $\alpha = 0$ или $\beta = 0$, то вероятности обнаружения выходящей микрочастицы в верхнем или нижнем полупространстве будут одинаковы;

- если амплитуды удовлетворяют условию $\alpha = \beta$, то *вероятность обнаружения выходящей микрочастицы в верхнем полупространстве становится равной единице*, т.е. при этом условии происходит снятие начальной неопределенности состояния кубита и оно становится строго однозначно определенным.

Можно использовать последовательность с двумя одинаковыми делителями 50/50 (интерферометр Маха-Цандера, рис. 1.29).

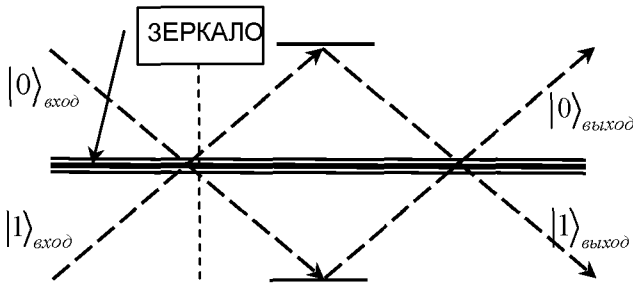


Рис. 1.29. Интерферометр Маха-Цандера

В случае двойного применения преобразования Адамара суммарное действие системы на входные сигналы сводится к тождественному преобразованию:

$$|Q\rangle_{\text{выход}} = HH|Q\rangle_{\text{вход}} = |Q\rangle_{\text{вход}}. \quad (1.108)$$

Интересно отметить, что когда на интерферометр Маха-Цандера подается всего один входной сигнал (например: верхний канал – $\alpha = 1$, а нижний входной канал – «пустой»), то:

- на выходе системы микрочастица *всегда* будет обнаружена наверху (реализация детерминированного состояния кубита);
- между делителями 50/50 микрочастица с одинаковой вероятностью может быть обнаружена как сверху, так и снизу (состояние кубита максимально неопределенное).

Базовый элемент на основе вариации фазы входного сигнала

Следующим базовым квантовым логическим элементом является сочетание преобразователя Адамара и фазовращателя (рис. 1.30).

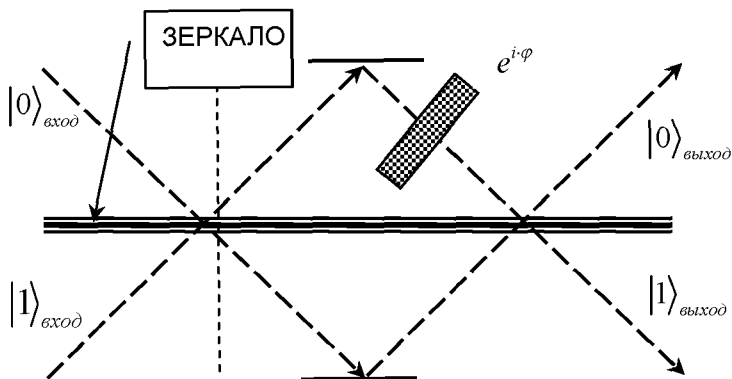


Рис.1.30. Интерферометр Маха-Цандера с фазовращателем

В качестве фазовращателя (фотонов) может выступать стеклянная пластинка конечной толщины. Физическая суть фазовращателя состоит в том, что он производит изменение фазы у одного из пучков. В символическом представлении его действие можно отобразить в виде (на примере верхнего входного пучка)

$$\Phi |0\rangle = e^{i\cdot\varphi} \cdot |0\rangle, \Phi |1\rangle = |1\rangle. \quad (1.109)$$

Кубит по схеме, представленной на рис. 1.30, описывается согласно выражениям (1.108) и (1.109) в виде

$$|Q\rangle_{\text{выход}} = H\Phi\Phi |Q\rangle_{\text{вход}}. \quad (1.110)$$

Если на входе этого кубита присутствует один пучок ($\alpha = 1, \beta = 0, |Q\rangle_{\text{вход}} = |0\rangle$), то выражение (1.110) трансформируется к виду

$$H\Phi H |0\rangle = \frac{1}{2} \left[(e^{i\cdot\varphi} + 1) \cdot |0\rangle + (e^{i\cdot\varphi} - 1) \cdot |1\rangle \right]. \quad (1.111)$$

Согласно выражению (1.111) можно сделать ряд выводов относительно состояния кубита в зависимости от фазового угла φ :

- если фазовый угол $\varphi = 0$, то значение (состояние) кубита строго определено и равно логической позиции «0»;

- если фазовый угол $\varphi = \pi$, то значение кубита также строго определено и равно логической позиции «1»;
- фазовый угол φ может выступать в роли фактора, способного переключать состояние кубита между логическими позициями «0» и «1»;
- при промежуточных значениях фазового угла $0 < \varphi < \pi$ вероятность состояния «0» равна значению $\cos^2(\varphi/2)$, а вероятность состояния «1» – $\sin^2(\varphi/2)$.

Базовый элемент по обеспечению перепутывания состояний индивидуальных кубитов

Рассмотрим более сложную в аппаратурном плане схему, состоящую из двух рассмотренных выше кубитов с фазовращателем и нового элемента – источника микрочастиц S (рис. 1.31). Источник S синхронно испускает пару микрочастиц, обладающих противоположным импульсом, таким образом, что частица, летящая налево (условно №1), обнаруживается в верхнем пучке, а частица (условно №2), летящая направо, будет обнаружена в нижнем пучке. И наоборот, если частица №1 обнаружена в нижнем пучке, то частица №2 будет в верхнем пучке.

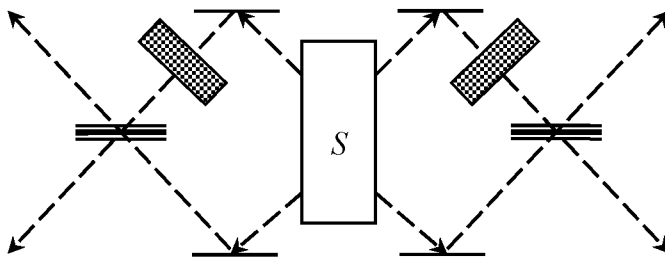


Рис. 1.31. Система, реализующая перепутывание квантовых состояний двух кубитов

В рамках принятого формализма кубитов можно считать, что эти микрочастицы несут противоположные значения логических переменных – «0» или «1». В используемом символьном представлении указанное двухчастичное состояние системы в общем случае можно отобразить в следующем виде:

$$\frac{1}{\sqrt{2}} \left[|0\rangle_1 |1\rangle_2 + e^{i\chi} \cdot |1\rangle_1 |0\rangle_2 \right], \quad (1.112)$$

где χ – разность фаз у испускаемых источником S микрочастиц, в дальнейшем принятая равной нулю.

Выражение (1.112) описывает такое специфическое состояние системы, которое получило название *перепутанного состояния*. Главной особенностью такого перепутанного состояния является то, что ни один из двух кубитов (справа и слева) не имеет изначально определенного значения. Но как только один из этих кубитов будет подвергнут процедуре измерения, то сразу «выяснится» определенное значение у другого кубита. Подобная «чувствительность» системы на внешнее возбуждение (т.е. измерения состояния одного из кубитов в системе) тем более неожиданно с классической точки зрения, если учесть тот факт, что в пространственном отношении оба кубита могут быть разнесены на очень большие расстояния.

Аналогичным образом можно создать перепутывание для произвольного количества кубитов. К примеру, частный случай перепутывание квантовых состояний трех кубитов (состояние Гринберга-Хорна-Цайлингера (ГХЦ)) можно описать с помощью выражения (при разности фаз $\chi = 0$)

$$\frac{1}{\sqrt{2}} \left[|0\rangle_1 |0\rangle_2 |0\rangle_3 + |1\rangle_1 |1\rangle_2 |1\rangle_3 \right], \quad (1.113)$$

И в этом случае, как только состояние одного из кубитов будет определено, два других кубита приобретут строго определенное значение.

С практической точки зрения представляется немаловажным, что существует достаточно много физико-технических систем, создающих перепутанные состояния. Это реализуется, в частности, при прохождении распада частиц со спином равным нулю на две микрочастицы со спином $1/2$, при условии сохранения внутреннего момента импульса системы в целом. В этом случае спины возникающих частиц имеют противоположные знаки и квантовое состояние такой системы можно представить в виде

$$\psi_{1-2} = \frac{1}{\sqrt{2}} \left[|\uparrow\rangle_1 |\downarrow\rangle_2 + |\downarrow\rangle_1 |\uparrow\rangle_2 \right], \quad (1.114)$$

где с помощью стрелок указаны квантовые состояния, соответствующие различным направлениям спина у микрочастиц.

1.3.2. Экспериментальные методы реализации квантовых компьютеров

а) Проблематика экспериментальной реализации квантовых компьютеров

В предыдущем разделе были рассмотрены общие теоретические аспекты, которые могут быть положены в основу практической реализации квантовых компьютеров. Согласно вышеизложенным материалам следует, что для реализации хотя бы одного логического элемента и вычислительной среды в целом необходимо:

- чтобы имелась система сильно взаимодействующих квантовых систем (образующих совокупность кубитов и элементов перепутывания квантовых состояний), выполненных на основе физических структур атомного масштаба;

- каждый индивидуальный кубит представлял собой автономную двухуровневую квантовую систему, в том смысле что обладал бы двумя устойчивыми и различимыми квантовыми состояниями;

- используемые для построения кубитов атомные структуры были бы практически абсолютно изолированы от неблагоприятного возмущающего воздействия факторов внешней среды.

В настоящем разделе представлен ряд экспериментальных методов (как возможных направлений на пути создания квантовых компьютеров), демонстрирующих принципиальные и технические аспекты реализации вычислительных сред на основе квантово-механических принципов. В современной экспериментальной практике наиболее широкое распространение получили методы, которые, несмотря на ряд пересечений в объектных и физических аспектах, можно с некоторой долей условности отнести к принципиально отличимым классам:

- квантовая электродинамика атомов в резонаторах (метод КЭР);
- ядерный магнитный резонанс (метод ЯРМ);

- ионные ловушки линейного типа.

У всех представленных выше экспериментальных методов есть свои достоинства и недостатки, в части возможности перекрытия всего диапазона практических аспектов, возникающих при создании квантовых вычислительных сред.

Эксперименты по КЭР являются прекрасным способом реализации (демонстрации) основных квантовых логических операций как таковых, однако осуществление большого массива операций вызывает существенные трудности принципиального и технического характера. В рамках методологии КЭР индивидуальные атомы (или ионы в магнитной ловушке), расположенные внутри оптического резонатора, рассматриваются в качестве двухуровневой системы, которая взаимодействует с внешним квантовым гармоническим осциллятором, выступающим в качестве внешнего управляющего фактора.

Решение проблемы увеличения масштабов вычислительных операций представляется более обещающим при использовании метода линейных ионных ловушек. В этом методе основная проблема заключается в создании сильно охлажденной цепочки ионов в ловушке линейного типа. Эту цепочку атомов можно рассматривать в качестве набора (регистра) индивидуальных кубитов, внешнее управление которым обеспечивается с помощью сфокусированного лазерного пучка. Решая успешно масштабный фактор, этот метод испытывает определенные ограничения в плане расширения номенклатуры элементарных логических операций.

Метод ЯРМ использует квантовые переходы между подуровнями атомных ядер в результате их взаимодействия с внешним магнитным полем. Этим методом продемонстрирована возможность выполнения небольшой последовательности элементарных квантовых (логических) операций на кубитах, в качестве которых выступают ядерные спины внутри отдельных молекул. Чрезвычайное разнообразие известных в настоящее время молекулярных систем расширяет практические возможности этого метода реализации вычислительных процессов.

Существует ряд альтернативных подходов к созданию квантовых вычислительных сред на основе твердотельных устройств (аналогов современных интегральных схем). Предпочтительность этих подходов

с точки зрения функциональных возможностей твердотельных устройств не вызывает сомнений. Однако весьма скромные практические результаты и стремление разработчиков как правило, реализовать полные аналоги существующих систем (основанных на классических принципах) до наноразмерного масштаба расходятся с проблематикой задач настоящего раздела.

Ниже будут продемонстрированы основные принципы создания квантовых вычислительных сред на примере метода КЭР.

б) Эксперименты по созданию вычислительных сред методом КЭР

Метод КЭР исследовался как в оптическом, так и микроволновом диапазоне. Основные физико-технические принципы проведения экспериментов в этих диапазонах достаточно близки. И в том, и в другом случае атом (или ион в электромагнитной или электростатической ловушке) представляет собой двухуровневую квантовую систему, взаимодействующую с внешним квантовым гармоническим осциллятором. В первом случае энергетическое состояние двухуровневого атома связывается с параметрами электромагнитной волны в резонаторе. Во втором случае – различные внутренние энергетические состояния индивидуального иона связываются с колебательными степенями свободы иона, находящегося в электромагнитной ловушке.

Поскольку в обоих случаях имеет место взаимодействие индивидуальных микрочастиц с электромагнитным полем, это взаимодействие может быть представлено гамильтонианом H одинакового вида (в упрощенной постановке, гамильтониан, для консервативных физических систем представляет собой полную энергию системы, то есть в классическом представлении сумму кинетической и потенциальной энергий системы) [9]:

$$H = -\frac{\hbar \cdot \Omega}{2 \cdot \pi} (a^- \cdot \sigma^+ + a^+ \cdot \sigma^-), \quad (1.115)$$

где a^+ и a^- – операторы рождения (+) и гибели (-) квантового осциллятора в системе; Ω – постоянная связи; σ^+ и σ^- – весовые операторы, повышающие или понижающие вероятность образования двухуровневой энергетической системы.

По своей физической сути, гамильтониан (1.115) отражает процессы генерации и рекомбинации фотонов (для случая атомной системы) или фононов (для случая ионной системы), благодаря которым происходит изменение начального энергетического состояния исходных микрочастиц (атомов или ионов). Несмотря на существенные различия в экспериментальном плане, атомно-резонаторные и ионно-ловушечные методы описываются в рамках единых теоретических представлений и моделей. И в том, и другом случае предполагается, что в качестве необходимых ключевых моментов экспериментальной практики выступает следующий круг факторов:

- обеспечение *точного равенства собственных частот* гармонического осциллятора, выступающего в роли инициатора энергетического возмущения и квантового перехода в рассматриваемой двухуровневой атомной (ионной) системе;
- реализация *режима сильной связи*, при котором взаимодействие, представленное выражением (1.115), выступает в качестве доминирующего процесса над всеми остальными возможными паразитными процессами релаксации системы и, в частности, такими как спонтанное излучение, фотон-фононное затухание и фазовая декогерентность, обусловленная влиянием теплового шума;
- достаточная, с точки зрения реализации элементарной вычислительной процедуры, *продолжительность метастабильного состояния* атомной (ионной) системы, находящейся в возбужденном квантовом состоянии;
- высокая степень *низкотемпературной* (порядка 0.6 К) стабилизации рабочей зоны, в которой происходит квантовый переход;
- обеспечение *высокой добротности резонатора* ($10^8 - 10^9$), ответственного за реализацию энергетического взаимодействия между внешним гармоническим осциллятором и атомной системой, испытывающей квантовый переход.

Суть метода КЭР состоит в том, что единичные атомы поставляются в резонатор, в котором они выступают в роли своеобразных антенн, поглощающих энергию единичного кванта электромагнитного излучения. В качестве таких атомов выступают циркулярные ридберовские

атомы [10], имеющие достаточно долгоживущие энергетические состояния, с значениями главных квантовых чисел n_g порядка 50 (этому состоянию можно поставить в соответствие символьное отображение $|g\rangle$). Под воздействием кванта электромагнитного излучения указанный атом переходит в новое квантовое состояние $|e\rangle$ со значением $n_e = 51$, в результате чего его энергия возрастет на величину $\Delta W_{g \rightarrow e}$, равную (в рамках модели «микрочастица в потенциальном ящике», см. выражение (1.23)):

$$\Delta W_{g \rightarrow e} = \frac{h^2}{8 \cdot m \cdot L^2} \cdot (n_e^2 - n_g^2). \quad (1.116)$$

Устойчивая работа рассматриваемой системы возможна лишь в том случае, если выполняется очевидное условие:

$$\Delta W_{g \rightarrow e} \gg W_T, \quad (1.117)$$

где $W_T \approx k \cdot T$ – среднее значение тепловой энергии атома, выступающей в роли возмущающего дестабилизирующего фактора.

Из выражений (1.116) и (1.117) следует условие для допустимых (критических) температурных режимов $T_{критич}$ проведения экспериментов по методу КЭР:

$$T_{критич} \leq \frac{\Delta W_{g \rightarrow e}}{k} = \frac{h^2}{8 \cdot m \cdot k \cdot L^2} \cdot (n_e^2 - n_g^2). \quad (1.118)$$

Оценка величины $T_{критич}$ (при $L = 10^{-10}$ м, $n_g = 50$, $n_e = 51$, $m = 10^{-26}$ кг), проведенная с помощью выражения (1.118), приводит к значениям температуры порядка 50 К (именно по этой причине в экспериментальной практике метода КЭР, как правило, используются криостаты с температурой 0,6 К (температура жидкого гелия)).

Общая схема экспериментальной установки, реализующей метод КЭР в миллиметровом диапазоне длин волн, представлена на рис. 1.32 [9].

Типичная совокупность этапов эксперимента, проводимого по методу КЭР, состоит из последовательности процессов прохождения од-

ного или двух атомов, разделенных между собой точно определенным (фиксированным) интервалом времени. Путь движения атомов – от источника микрочастиц O через блок резонатора C к блокам детектирования квантовых состояний указанных атомов – D_e и D_g .

Указанный процесс прохождения атомов повторяется многократно, с периодом порядка 10^{-3} с, продолжительность которого намного превышает время затухания внешних электромагнитных колебаний в резонаторе C . Таким путем достигается такое состояние системы, что при запуске очередной партии атомов резонатор C «постоянно» находится в невозмущенном состоянии. Невозмущенное стартовое состояние резонатора C является необходимым условием, реализующим индивидуальную вычислительную операцию при эксперименте по методу КЭР.

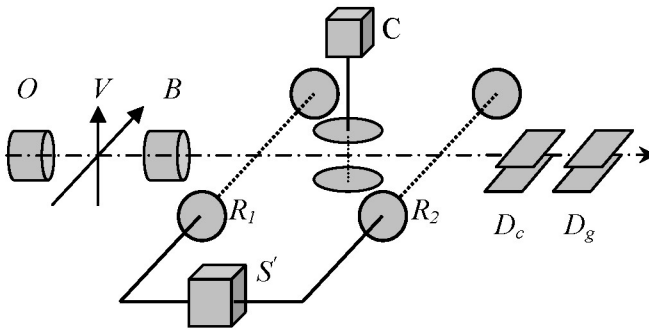


Рис.1.32. Схема установки в эксперименте с атомом в резонаторе

До входа в резонатор C атомы проходят через вспомогательный резонатор R_1 , в котором микроволновый импульс, поступающий от источника S_{1-2} , производит перемешивание квантовых состояний атома $|g\rangle$ и $|e\rangle$. После прохождения резонатора C производится вторичное перемешивание состояний $|g\rangle$ и $|e\rangle$ во вспомогательном резонаторе R_2 .

Атомы, испущенные источником микрочастиц O , отбираются по скоростям в зоне V . Отселектированные по скоростям в зоне V атомы

приготавливаются в блоке B в одном из квантовых состояний $|g\rangle$ или $|e\rangle$ при помощи последовательности лазерных импульсов и адиабатических переходов в радиочастотном диапазоне частот.

Резонатор C настроен в резонанс с атомным переходом $|e\rangle \rightarrow |g\rangle$, т.е. его резонансная частота $\nu_{рез}$ выбирается из условия:

$$\nu_{рез} = \frac{\Delta W_{g \rightarrow e}}{h} = \frac{h}{8 \cdot m \cdot L^2} \cdot (n_e^2 - n_g^2). \quad (1.119)$$

В наиболее простой постановке эксперимента по методу КЭР определяются вероятности квантового перехода $|e\rangle \rightarrow |g\rangle$ при размещении индивидуального атома в резонаторе C (при этом вспомогательные резонаторы R_1 и R_2 не задействованы). Диагностика состояния атома производится для различных времен взаимодействия τ между атомом и электромагнитным полем резонатора. Значение τ может варьироваться либо изменением скорости прохождения атомами активной зоны резонатора C , либо настройкой частоты атомного перехода в резонанс с модой электромагнитного поля за некоторую часть пролетного времени. При этом процессе обмена энергией между частицей и полем обратимая эволюция квантовых переходов $|e\rangle \leftrightarrow |g\rangle$ между базовыми квантовыми состояниями $|g\rangle$ и $|e\rangle$ сопровождается процессами *испускания и поглощения единичных фотонов*. Попутно следует отметить тот немаловажный факт, что последнее обстоятельство служит прямым доказательством возможности квантования электромагнитного поля.

Помимо доказательства квантования энергии электромагнитного поля, эксперимент, представленный выше, свидетельствует о том, что резонансное взаимодействие квантового состояния атома с модой электромагнитного поля превалирует над процессами релаксации (диссипации), обусловленными тепловыми флуктуационными факторами. В этой связи результирующий эффект, т.е. перепутывание состояний в системе «атом – поле», может быть использовано для управляемого

манипулирования процессом квантового перепутывания, составляющего основу квантовых вычислительных операций.

Фактор перепутывания квантовых состояний между атомом и электромагнитным полем позволяет реализовать ряд базовых устройств, предназначенных для целей создания квантовых вычислительных систем, и прежде всего элемента памяти на основе одного кубита. Такая память может быть создана при записи информации первым атомом и ее чтении с помощью второго атома, движущегося вслед первому атому. Первый атом, находящийся в квантовом состоянии $|e\rangle$, попадает в пустой резонатор C и находится там в течение времени взаимодействия τ , удовлетворяющего условию $\nu_{рез} \cdot \tau = \pi$. За это время взаимодействия атом переходит в квантовое состояние $|g\rangle$, оставляя однофотонное состояние в резонаторе C . Второй атом, входящий в резонатор C в квантовом состоянии $|g\rangle$, поглощает это однофотонное состояние и при отсутствии процессов спонтанного распада покидает резонатор C в квантовом состоянии $|e\rangle$. В этом эксперименте кубит, реализованный (в качестве носителя информации) на основе однофотонного поля, в своих конечных состояниях распределяется между квантовыми состояниями двух атомов. В промежуточном состоянии поле в резонаторе C является квантовой суперпозицией вакуумного и однофотонного состояний. Рассмотренный тип процессов (систем) является базовым при реализации логических элементов методом КЭР.

Для выполнения операций, представляющих интерес для квантовых вычислений, необходимо манипулировать со значительным количеством кубитов (порядка нескольких десятков). Это условие существенно ограничивает возможность практического использования метода КЭР, поскольку серьезные трудности в проблематике декогерентности начнутся в рамках этого метода уже при последовательности всего лишь из трех атомов. В этой связи эксперименты по методу КЭР следует рассматривать более в качестве эффективных подходов при исследовании вопросов декогерентности, диагностики квантовых состояний и пере-

путывания квантовых состояний, нежели в качестве физико-технической базы построения в обозримой перспективе квантовых вычислительных сред.

1.4. Теоретические и практические основы спинтроники

1.4.1. Спин микрочастицы: определение и форма представления

Спином микрочастицы называется собственный *момент импульса* элементарных частиц, имеющий квантовую природу и не связанный с перемещением частицы как целого. Спином называют также собственный момент импульса более сложных объектов, таких как атомное ядро или атом в целом. В этом случае спин определяется как векторная сумма, вычисленная по правилам сложения моментов, спинов элементарных частиц (протонов, электронов), образующих систему, и орбитальных моментов этих частиц, обусловленных их движением внутри системы. Спин измеряется в единицах $\hbar/2$ (приведенная постоянная Планка) и равен

$$s = \frac{\hbar}{2} \cdot J, \quad (1.120)$$

где J – характерное для каждого сорта частиц целое (в том числе нулевое) или полуцелое положительное число, определяемое как *спиновое квантовое число*, которое зачастую называют просто спином микрочастицы.

В общем случае любая микрочастица может обладать двумя видами углового (вращательного) момента: орбитальным и спином. В отличие от орбитального углового момента, который порождается движением микрочастицы в пространстве, спин не связан с движением в пространстве. Спин микрочастицы – исключительно квантовая характеристика, которую нельзя объяснить в рамках классической механики. Если представлять микрочастицу (например, электрон) как вращающийся шарик, а спин как момент, связанный с этим вращением, то оказывается, что поперечная скорость движения оболочки частицы должна быть выше скорости света, что недопустимо с позиции релятивистских представлений.

Будучи одним из проявлений углового момента, спин в квантовой механике описывается векторным оператором спина \hat{S} . Однако в отличие от орбитального углового момента, оператор спина не выражается через традиционные классические переменные, иными словами, спин – исключительно квантовая величина. Следствием этого является тот факт, что сам спин и его проекции на какую-либо пространственную ось может принимать не только целые, но и полуцелые значения (табл. 1.2).

Т а б л и ц а 1.2. *Характерные значения спина для различных микрочастиц (в единицах $\hbar/2$)*

Значение спина	Тип микрочастицы	Характерные представители среди микрочастиц
0	скалярные частицы	π или К-мезоны, атомы и ядра ^4He
1/2	спинорные частицы	электрон, протон, нейтрон, атомы и ядра ^3He
1	векторные частицы	фотон, векторные мезоны
3/2	спин-векторные частицы	Δ -изобары
2	тензорные частицы	гравитон, тензорные мезоны

Несмотря на тот факт, что спин микрочастицы никоим образом не связан с реальным вращением частицы вокруг собственной оси, наличие спина тем не менее порождает определённый магнитный момент ($\hat{\mu}$) у электрически заряженной микрочастицы. Последнее обстоятельство приводит к дополнительному (по сравнению с классической электродинамикой) взаимодействию микрочастицы с внешним магнитным полем. Отношение величины магнитного момента к величине спина микрочастицы называется *гиромагнитным отношением* и в отличие от орбитального углового момента оно не равно численному значению μ_0 ($\mu_0 = 1.25 \cdot 10^{-6}$ Гн/м):

$$\frac{\hat{\mu}}{\hat{S}} = g \cdot \mu_0, \quad (1.121)$$

где g – постоянная величина, называемая « g -фактором», характерная для каждого типа элементарных микрочастиц.

Ансамбли микрочастиц, имеющих различные значения спина, подчиняются различным статистикам (функциям распределения по энергиям). Вследствие того что все элементарные частицы одного и того же сорта тождественны (неразличимы) между собой, волновая функция системы из нескольких одинаковых микрочастиц может быть либо *симметричной*, либо *антисимметричной* относительно перестановки местами двух любых частиц. В первом случае говорят, что микрочастицы подчиняются статистике «Бозе-Эйнштейна» («бозоны»), во втором случае – статистике «Ферми-Дирака» («фермионы»). Оказывается, именно значение спина микрочастицы свидетельствует о том, каковы будут эти свойства относительно симметрии. Сформулированный Паули принцип о связи значения спина с типом статистики устанавливает следующее правило:

- микрочастицы с целочисленным спином ($s = 0, 1, 2, \dots$) являются бозонами (типичный представитель – фотон);
- микрочастицы частицы, обладающие полуцелым значением спина ($s = 1/2, 3/2, \dots$), являются фермионами (типичный представитель – электрон).

В практическом плане наиболее широкая постановка вопроса по использованию спиновых представлений в информационно-измерительной технике связана с фермионами, и в частности – электронами.

1.4.2. Этапы становления и перспективы развития спинтроники

Принцип действия практически всех известных в настоящее время промышленных электронных компонентов и приборов основан на манипулировании в той или иной форме токами электрически заряженных частиц, и в частности электронов или ионов. Отсюда и произошли привычные в настоящее время представления – электроника, микроэлектроника и нанoeлектроника.

В последние годы наблюдается настоящий прорыв в область нанотехнологий, позволяющих уменьшить характеристический размер отдельных электронных компонентов до уровня, сопоставимого с размерами атомов или с де Бройлевской длиной волны электрона. И в том, и в другом случае это обстоятельство приводит к необходимости учета

ряда специфических квантово-механических законов микромира и позволяет рассматривать совершенно новые практические возможности в направлении создания устройств наноинженерии.

Одна из этих возможностей непосредственно обусловлена тем, что электрон – это не просто заряженная микрочастица, а частица, обладающая такой специфической квантовой характеристикой, как спин – внутренний момент количества движения и связанный с ним магнитный момент. При этом весьма примечательно, что электрон может находиться в двух достаточно устойчивых спиновых квантовых состояниях (которые условно классифицируются как «спин вверх» и «спин вниз»), что напрямую связано с возможностью создания устройств релейного типа (в том числе на базе представлений о «кубитах»). Последнее значит, что спиновые системы самой природой предназначены для кодирования информации. Одно состояние может кодировать как условное «1», а второе – «0». В этом случае перепутанное состояние спиновых кубитов (с элементарными состояниями: $|\uparrow\rangle \equiv |\text{спин} \mapsto \text{вверх}\rangle$; $|\downarrow\rangle \equiv |\text{спин} \mapsto \text{вниз}\rangle$) некогерентной смеси микрочастиц можно представить в виде матрицы плотности [9]:

$$\rho = f \cdot (|\uparrow\rangle\langle\uparrow|) + (1-f) \cdot (|\downarrow\rangle\langle\downarrow|), \quad (1.122)$$

где f – доля общего числа микрочастиц.

Отбирая через селектор состояний (например, анизотропной магнитной системы Штерна-Герлаха), можно подготовить из исходной смеси (1.122) «чистые» состояния типа

$$\rho' = 1 \cdot (|\uparrow\rangle\langle\uparrow|) \quad (1.123)$$

Управление спиновыми состояниями электронов позволяет создать сверхмалые логические элементы, и – в перспективе – компьютерные компоненты с огромным быстродействием, малым энергопотреблением и большой информационной емкостью (даже на базе классических методологических и алгоритмических подходов).

Зарождение спинтроники как нового научно-технического направления принято датировать 1998 годом, когда в совместном заявлении

фирмы Bell Labs и Йельского университета была сформулирована задача создания устройств, в которых информацию хранили бы атомы вещества, а биты кодировались бы электронными спинами. (По менее распространенной версии начало спинтроники (1988 г.) связывают с открытием эффекта гигантского магнитосопротивления (ГМС) [11]).

На первых этапах развития спинтроники речь шла о квази-квантовых подходах использования физических представлений о спине в технических системах. Но у спинтроники есть и более интересная в части возможности реализации вычислительных сред перспектива, связанная с тем, что она способна эффективно реализовать идеологию параллельных квантовых вычислений. Первые успехи в этом направлении были получены уже в 2000 г., когда увенчалось успехом создание прообраза первого твердотельного квантового компьютера на основе кристалла кремния [12]. На этом пути удалось совместить существующие планерные технологии производства интегральных схем и идеологию квантовых вычислений. Квантовые биты реализовывались с помощью спинов ядер фосфора, внедренного в изотопно-чистый кремний, и управлялись посредством комбинации статического магнитного, статического электрического и радиочастотного электромагнитного полей. Несмотря на достаточно скромные результаты (это была всего лишь система на базе двух кубитов), полученные твердотельные структуры продемонстрировали принципиальную возможность построения спинового квантового компьютера.

Практически в этот же период удалось создать спин-поляризованный электрический ток в полупроводниковых слоистых структурах [13]. Для этого было применено редко используемое свойство материалов с так называемыми «квантовыми карманами». Под последним понимается наличие асимметрии в кристаллической решетке, которая приводит к тому, что электроны с противоположными спинами приобретают противонаправленные ненулевые скорости. Такие электроны в достаточном количестве испускаются под действием лазерного излучения с круговой поляризацией, падающего на поверхность "сандвича" из полупроводников различных типов. Меняя направление поляризации света, удалось добиться обращения спинового тока, а значит, открытый

метод может быть использован для создания логических переключателей для вычислительных систем.

Дальнейшее развитие работ было проведено в направлении создания спинового тока непосредственно между двумя полупроводниками p - и n -типов. При этом было обнаружено, что если упорядочить спины группы электронов и затем с помощью электрического поля направить их на смежный полупроводник, то спиновое упорядочивание будет сохраняться достаточно долго даже при комнатной температуре. Образуется своего рода спиновый резервуар, причем полный переносимый таким образом спин может обладать свойством как резервуара, так и смежного полупроводникового слоя (спинового переноса). С помощью внешнего электрического поля оказывается возможным переключение между этими двумя режимами. Важность этого открытия заключается в том, что управление спинами осуществлено с помощью электрического, а не магнитного поля, что очень важно именно для создания нового типа устройств вычислительной техники [14].

Следует отметить, что промышленная практика уже десять лет использует достижения спинтроники в виде жестких компьютерных дисков и прецизионных сенсоров магнитного поля, в которых использован эффект гигантского магнетосопротивления.

Ниже представлен ряд примеров реализации устройств спинтроники, который позволяет составить общее представление о конструктивно-технологических особенностях элементов, работающих на базе спиновых представлений.

1.4.3. Магнитно-полупроводниковые спиновые нанотранзисторы

Учитывая высокий уровень развития современной интегральной полупроводниковой электроники, практическое развитие спинтроники идет по пути интеграции магнитных систем и элементов конструкций полупроводниковой микроэлектроники. Этому способствует сравнительно легкое управление поведением спинами электронов в полупроводниках. Имеющийся в настоящее время теоретический задел позволяет создать два новых класса гибридных устройств наноэлектроники:

- на основе магнитных полупроводников (т.е. гибридная структура типа «ферромагнетик/полупроводник»);

- спин-электронные нанотранзисторы на основе наноразмерных гетероструктур с элементами магнитного управления.

Широкие перспективы использования наноразмерных гетероструктур обусловлены тем, что электронные спины полупроводника можно использовать в качестве детектора, реагирующего на изменения магнитного состояния в ферромагнетике. Так, при инъекции сквозь контакт ферромагнетика и полупроводника электроны полупроводника приобретают неравновесный спин, содержащий информацию о спине электронов в ферромагнетике. Для определения спиновой ориентации электронов в полупроводнике можно использовать как оптические, так и электрические методы детектирования.

Магнитные полупроводники

В традиционных элементах интегральной электроники металлические материалы используются исключительно для создания элементов электрической коммутации. В данном случае речь идет о металлических материалах, которые используются в сфере их магнитных свойств. Существенным недостатком этого подхода (обусловленным омическим характером электропроводности металлов) является невозможность реализации эффекта усиления входных сигналов по току или напряжению. Оптимальным решением этой проблемы является разработка новых материалов, которые сочетали бы свойства ферромагнетиков и полупроводников – «магнитные полупроводники». В такой постановке вопроса задача труднодостижима. Однако при ее реализации появляются большие практические перспективы, поскольку этот тип материалов, с одной стороны, выступал бы в качестве источников спин-поляризованных электронов, а с другой – легко интегрировался бы с традиционными полупроводниковыми устройствами интегральной электроники. Следует отметить, что зонная диаграмма магнитного полупроводника отличается от двухзонной (зоны: валентная, проводимости) структуры обычных полупроводников, металлов и диэлектриков наличием особой, третьей зоны, которая образуется d - и f -электронами атомов переходных или редкоземельных элементов. Идеальный с практической точки зрения магнитный полупроводник должен иметь тем-

пературу Кюри (температура, при которой ферромагнетик теряет свои свойства) выше комнатной температуры и допускать создание «*p-n*» переходов в пределах монокристаллической структуры.

Ввиду сложности создания магнитных полупроводников как таковых, в настоящее время привлекают внимание так называемые разбавленные магнитные полупроводники, т.е. сплавы, в которых отдельные атомы традиционного полупроводника замещаются (в ограниченном количестве) атомами с магнитными свойствами, например Mn^{2+} . Наибольшая температура Кюри $T_c = 110$ К (низкая для широкого практического использования температура) достигнута на сегодняшний день в магнитном твердом растворе Ga-Mn-As с *p*-типом проводимости. Этот материал был использован в качестве спинового инжектора (включая и нулевое внешнее магнитное поле) в электролюминесцентном диоде с немагнитной квантовой ямой In-Ga-As/Ga-As [15].

Основной фронт работ по данному направлению состоит в поиске новых ферромагнитных полупроводников (типа – твердый раствор) с более высокой температурой Кюри, которые могли бы быть использованы в качестве спиновых инжекторов при температурах порядка комнатной. Наиболее интересные результаты в этом направлении представлены в работе [16], в которой сообщается о наблюдении ферромагнетизма при $T_c = 320$ К в полупроводнике со структурой халькопирита $Cd_{1-x}Mn_xGeP_2$.

Спиновый биполярный транзистор

Развитие работ по спинтронике способствовало переходу от двухконтактных спин-электронных устройств к трехконтактным системам, состоящим из двух ферромагнитных тонкопленочных слоев (проявляющим эффект гигантского магнитосопротивления), разделенных парамагнитным слоем [17]. Схематическое изображение такого биполярного транзистора представлено на рис. 1.33.

Если говорить на языке биполярных транзисторов, то транзистор Джонсона состоит из базы (парамагнетик), эмиттера и коллектора (ферромагнетики). Если на коллектор подается потенциал, в цепи эмиттер–база происходит накопление электронов с ориентацией спинов

вверх-вниз. Ток коллектора теперь будет зависеть от того, параллелен ли его магнитный момент или антипараллелен намагниченности эмиттера.

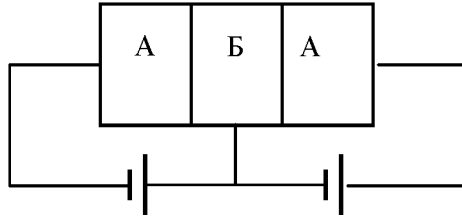


Рис.1.33. Схема спинового биполярного транзистора Джонсона (А – ферромагнитный слой; Б – диамагнитный слой)

Ферромагнитный эмиттер в данном случае играет роль поляризатора для накапливающихся спинов. Для изменения потенциала в цепи эмиттер–база необходимо приложить внешнее магнитное поле, которое «переключит» вектор магнитного момента либо коллектора, либо эмиттера на противоположное направление.

Специалисты предсказывают широкое использование разработок спинтроники в сфере создания квантового компьютера. Выражается надежда, что следующим этапом развития спинтроники станут устройства, в которых информация будет передаваться не посредством спинов электронов, а с использованием сложных кубитовых пар. Например, многоконтактные спиновые устройства, которые могут быть основаны на потоках запутанных состояний кубитов. Практическая реализация подобных квантовых устройств может быть основана на базе спин-электронных транзисторных устройств.

Глава 2. СПЕЦИФИКА МЕХАНИЗМОВ ЭЛЕКТРИЧЕСКОЙ ПРОВОДИМОСТИ НАНОРАЗМЕРНЫХ СТРУКТУР

2.1. Элементы теории электропроводности конденсированных сред

2.1.1. Проблематика и актуальность процессов транспорта носителей заряда в наноразмерных структурах

Значительный фронт современных научных и прикладных работ в направлении нанотехнологий прямым или косвенным образом связаны с проблематикой создания перспективных высокопроизводительных (быстродействие, объем памяти, массово-габаритные параметры) вычислительных систем, обладающих чрезвычайно высокой степенью интеграции элементов (*степень интеграции элементов* – количество элементов конструкций, содержащихся в единице объема изделия). В табл. 2.1 представлены сведения о реальной и прогнозируемой динамике повышения степени интеграции элементов в вычислительных системах.

Таблица 2.1. Динамика уменьшения размеров элементов конструкций электронных устройств в вычислительных системах

Год выпуска изделия	Степень интеграции Элемент / мм ³	Минимальный размер элемента, мкм	Допуск на изготовление, мкм
1960	10 ⁻² (биполярный транзистор)	3·10 ³	> 10
1970	0.1 – 1.0 (микросборка элементов)	10 ³	5 – 10
1980	10 ³ (ИС на биполярных структурах)	10 ²	2 – 5
1990	10 ⁶ (ИС на МОП структурах)	2 – 10	0.2 – 1.0
2000	10 ⁹ (СБИС – 1 ^е поколение)	0.5 – 1.0	0.05
2010	10 ¹² (СБИС – 2 ^е поколение)	10 ⁻¹	10 ⁻²
2020-2030	10 ¹⁴ – 10 ¹⁵ (прогноз)	10 ⁻²	10 ⁻³
Технология атомной сборки	10 ¹⁹ – 10 ²⁰ (без учета элементов коммутации)	10 ⁻⁴	Теоретический предел атомных технологий

Примечание: ИС – «интегральная схема»; СБИС – «сверхбольшая интегральная схема»; МОП – структура типа «металл – окисел – полупроводник».

Согласно данным табл. 2.1, в этом научно-техническом направлении можно говорить о вхождении в область наноразмерных масштабов уже с 2000 г., по крайней мере в части обеспечения необходимых допусков на изготовление индивидуальных элементов конструкций вычислительных систем.

Столь малые размеры элементов конструкций современных и перспективных устройств электронной техники накладывают определенную специфику на процессы электропроводности в конструкционных материалах, применяемых при создании этих устройств. В частности, представляется весьма проблематичным использование в физико-технической практике традиционных и широко используемых теоретических моделей и подходов, основанных на макроскопических представлениях физики сплошных сред (термодинамика, статистическая и молекулярная физика).

Такое пристальное внимание к вопросам электропроводности указанных конструкционных материалов обусловлено их ключевой ролью при рассмотрении проблематики обеспечения функционального назначения изделий электронной техники. В общем случае при анализе электрической проводимости σ конструкционных материалов (в рамках макроскопического подхода) принято оперировать *уравнением электропроводности* вида

$$\sigma = \sum_i q_i \cdot n_i \cdot \mu_i, \quad (2.1)$$

где i – индекс, характеризующий вклад в общую электропроводимость материала носителей зарядов $i^{\text{го}}$ типа; n_i – концентрация носителей зарядов $i^{\text{го}}$ типа; μ_i – подвижность носителей заряда $i^{\text{го}}$ типа; q_i – модуль значения электрического заряда для индивидуального носителя заряда $i^{\text{го}}$ типа.

В случае макроскопических систем установление типа и величины q_i индивидуальных носителей заряда, участвующих в процессе формирования электрического тока, не вызывает особых затруднений, поскольку они, как правило, известны для большинства используемых на практике конструкционных материалов. Для случая наноразмерных (хотя бы по одной пространственной координате) элементов конструк-

ции, выполненных на основе материалов с электронным типом проводимости, этот аспект может существенно усложниться. Это связано со спецификой рассмотренных выше зонных энергетических диаграмм наноструктур, обусловленной наличием размерных ограничений на процессы движения микрочастиц и порой недостаточным (для реализации процесса векторизации энергии поступательного движения с помощью электрического поля) количеством свободных энергетических состояний. Вопрос о том, какие именно электроны из их общего количества в материале в этом случае можно отнести к носителям заряда, будет рассмотрен ниже. Этот важный для практики создания электронных устройств аспект самым непосредственным образом связан и с вопросами установления концентрации носителей заряда n_i .

Более прозрачной представляется взаимосвязь подвижности носителей заряда μ_i с пространственными ограничениями наноразмерных структур. Очевидным следствием ограничительных факторов является появление ярко выраженной *анизотропии*, т.е. зависимости свойств конструкционных материалов от рассматриваемого пространственного направления. Следует отметить, что фактор анизотропии наблюдается и у макроскопических материалов, однако в этих средах он проявляется, как правило, в весьма незначительных масштабах (за исключением узкого круга конструкционных материалов специального назначения: сегнетоэлектрики; ферромагнетики; композиты).

В разд. 1 показано определяющее влияние вида потенциальной функции электронов (т.е. конфигурация барьеров на пути следования микрочастиц) в атомных и регулярных (кристаллических) структурах на характер движения последних, которая (функция) формирует специфику процессов квантования элементарных энергетических состояний электронов. В рассматриваемых наноразмерных структурах потенциальная функция электронов может кардинально отличаться для различных пространственных направлений (в особенности для структур типа (1D) и (2D)), что, собственно, и составляет их главное отличие от макроскопических кристаллических структур. С учетом вышеизложенного фактора резкой анизотропии потенциальной функции электронов

в наноразмерных структурах представляет интерес рассмотрение двух специфических вариантов процесса переноса носителей заряда:

- вдоль совокупности дискретных атомных потенциальных барьеров, т.е. в направлении отсутствия ограничений размерного характера;
- поперек совокупности дискретных атомных потенциальных барьеров, т.е. в направлении наличия ограничений размерного характера.

Первый, более простой вариант соответствует случаю практически свободного движения электрона (при большой длине цепочки атомов образующих наноразмерную структуру в направлении движения), при реализации которого волновой вектор k пробегает непрерывный ряд значений в диапазоне величин $\{(2\cdot\pi)/L - (2\cdot\pi)/a\}$ (см. разд. 1.2.2, пункты д; е), где L – длина атомной цепочки. В этом варианте движение индивидуального электрона происходит путем последовательного чередования этапов свободного пролета на *длине свободного пробега* по пути следования электрона (обратить внимание на факт перекрытия волновых функций электрона согласно рис. 1.9) с этапами ее неупругого взаимодействия с атомами среды (электропроводимость по механизму «*пролетного типа*»). Этот тип движения аналогичен поведению индивидуального электрона в совершенных макроскопических кристаллических структурах, в которых процессы рассеяния векторизованного движения электронов происходят на дефектах точечного (вакансии, атомы внедрения и замещения) или линейного (краевые и винтовые дислокации) типа и на *фононах* (совокупности взаимосвязанных элементарных колебательных движений атомов, образующих кристаллическую структуру).

Второй, более сложный для анализа вариант соответствует случаю жесткой локализации местонахождения электрона в ограниченной области пространства, которому можно поставить в соответствие дискретный набор волновых векторов k , т.е. дискретный спектр энергетических состояний (см. разд. 1.2.2, пункт г). В этом случае движение индивидуального электрона происходит в виде последовательности

дискретных «прыжков» («переходов»), которые могут быть реализованы путем:

- преодоления максимальной высоты потенциального барьера при реализации элементарного акта движения, для чего индивидуальным электронам необходимо обладать энергией большей, чем высота барьера (электропроводимость по кативационному механизму «прыжкового типа»);
- своеобразного «просачивания» индивидуального электрона сквозь потенциальный барьер, исключая необходимость преодоления этого барьера по механизму прыжкового (см. предыдущий пункт) типа (электропроводимость по механизму «туннельного типа»).

Приведенная классификация в определенной мере, в общих чертах, прослеживает основополагающие принципы и методы анализа процессов электронного и ионного транспорта в наноразмерных кристаллических структурах.

2.1.2. Элементы теории подвижности носителей заряда

а) Подвижность носителей заряда в атомно-молекулярных системах

В отсутствие внешнего электрического поля микрочастицы, находящиеся в атомно-молекулярных системах, совершают хаотическое (броуновское) движение, обусловленное наличием у них тепловой энергии $(k \cdot T)/2$ на каждую степень свободы движению [1]. Наложение на систему внешнего электрического поля приводит к тому, что микрочастицы, обладающие электрическим зарядом, приобретают на пути своего движения дополнительную энергетическую составляющую, имеющую векторизованный характер. Именно благодаря этому фактору векторизации движения микрочастиц имеет место появление электрического тока в макроскопических атомно-молекулярных системах.

В качестве количественной меры векторизации движения микрочастицы в электрическом поле общепринято использовать специальный параметр μ – *подвижность носителя заряда*, который определяется в виде

$$\mu = \frac{d\nu}{dE}, \quad (2.2)$$

где $d\nu$ – изменение *средней скорости* движения носителя заряда при изменении напряженности электрического поля на величину dE .

По своей физической сути численное значение подвижности носителя заряда характеризует среднюю скорость векторизованного движения микрочастицы в электрическом поле со значением напряженности, равным 1 В/м. В общем случае параметр μ является нетривиальной функцией от напряженности электрического поля. Однако для подавляющего числа практических случаев, в которых, как правило, выполняется условие $E < 10^3$ В/м, его принято считать постоянной величиной.

Практическая целесообразность введения параметра μ состоит в том, что с его помощью можно определить среднюю скорость направленного движения носителя заряда при любом (с учетом оговоренного выше условия) значении напряженности электрического поля E с помощью выражения

$$\nu(E) = \mu \cdot E. \quad (2.3)$$

Интересно оценить масштабный фактор процесса векторизации движения микрочастицы в электрическом поле. Для этого проведем сравнительный анализ броуновской и векторизованной составляющих скоростей движения микрочастицы. Для упрощения анализа примем, что векторизованная составляющая скорости движения микрочастицы соответствует механизму электропроводности *пролетного типа* при длине свободного пробега микрочастицы $\lambda = 10^{-8}$ м (типичное, по порядку величины, значение для монокристаллических структур металлического типа). Оценка величины векторизованной ν_E и броуновской (тепловой) ν_T составляющих скорости движения микрочастицы производится на основе очевидных энергетических соотношений (использован подход классической физики):

$$W_k(E) = \frac{m \cdot \nu_E^2}{2} = q \cdot \lambda \cdot E; \quad (2.4)$$

$$W_k(T) = \frac{m \cdot v_T^2}{2} = \frac{3 \cdot k \cdot T}{2}, \quad (2.5)$$

где $W_k(E)$, $W_k(T)$ – значения кинетической энергии движения микро-частицы соответственно для случая векторизованной и броуновской форм ее движения; q – величина электрического заряда носителя заряда; k – постоянная Больцмана. В табл. 2.2 приведены данные сравнительного анализа параметров векторизованной и броуновской (тепловой) форм движения микрочастицы.

Т а б л и ц а 2.2. *Зависимость отношения скоростей $(v_E)/(v_T)$ и кинетической энергии $w_k(E)$ от напряженности электрического поля E ($T = 300 \text{ К}$ ($3kT/2 = 0.039 \text{ эВ}$); $\lambda = 10^{-8} \text{ м}$)*

Параметр	Напряженность электрического поля E , В/м					
	10	10^2	10^3	10^4	10^5	10^6
$W_k(E)$, эВ	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
$(v_E)/(v_T)$	$\approx 10^{-6}$	$\approx 10^{-5}$	$\approx 10^{-4}$	$\approx 10^{-3}$	$\approx 10^{-2}$	$\approx 10^{-1}$

Из данных табл. 2.2 следует, что для большинства практических случаев ($E < 10^5 \text{ В/м}$; $T = 300 \text{ К}$) векторизованная составляющая скорости движения микрочастиц не превышает одного процента от значения броуновской составляющей скорости. В этой связи зачастую принято считать, что процесс токопрохождения не нарушает равновесную функцию распределения микрочастиц по энергиям (скоростям), а приводит лишь к появлению весьма незначительного дрейфа всего ансамбля микрочастиц в пространстве в направлении воздействия на носители заряда внешнего электрического поля.

В общем случае подвижность носителей заряда в материалах является достаточно сложной функцией от температуры и концентрации дефектов (точечного, линейного и поверхностного типов), имеющих в кристаллической структуре. Характерный вид этой функции для случая материалов металлического и полупроводникового типа представлен на рис. 2.1.

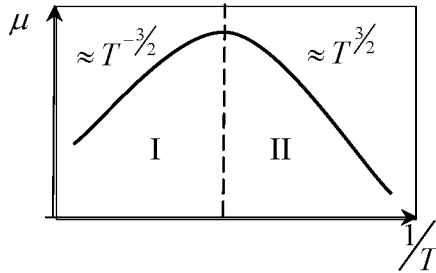


Рис. 2.1. Температурная зависимость подвижности носителей заряда для случая материалов металлического и полупроводникового типа

В температурной зависимости подвижности носителей заряда $\mu(T)$ достаточно четко просматриваются два характерных участка:

- I – область *высоких температур*, в которой определяющую роль в процессах рассеяния векторизованного типа движения электронов проводимости играют колебательные процессы атомов кристаллической решетки, интенсивность которых (фононов) возрастает с повышением температуры;
- II – область *низких температур*, в которой определяющую роль в процессах рассеяния векторизованного типа движения электронов играют процессы неупругого взаимодействия электронов проводимости с различными дефектами (точечного и линейного типа) и ионизированными атомами (как правило, примесной природы – обладающих небольшим значением энергии (потенциала) ионизации атомов) кристаллической структуры.

Следует отметить, что из всех типов точечных дефектов (вакансии, атомы внедрения и замещения), которые имеются в кристаллических структурах, наибольшее рассеивающее воздействие оказывают ионизированные примеси. Столь сильное влияние на процессы рассеивания со стороны ионизированных примесей обусловлено тем обстоятельством, что активная (с точки зрения возможности реализации неблагоприятного воздействия на векторизованное движение электрона, принимающего участие в процессе образования электрического тока проводимости) область электрического поля точечного заряда имеет значительную пространственную протяженность (*радиус экранирования* электриче-

ского поля – более 10 ангстрем). Отсюда следует, что электрическое возмущение параметров состояния среды, обусловленное присутствием всего лишь одного зарядового центра атомного масштаба, может быть распространено на элемент объема кристалла, содержащего более 10^3 атомов, образующих рассматриваемую кристаллическую структуру.

В случае макроскопических сред (объектов) негативное проявление рассмотренного выше момента в определенной мере нивелируется (с позиции критериев обеспечения функционального назначения материала или устройства) наличием в системе весьма значительного количества «нужных» атомов. В случае же наноразмерных объектов столь значительная величина радиуса экранирования может привести к губительным последствиям уже при наличии единичных зарядовых центров (ионов) примесного характера. Последнее обстоятельство налагает весьма строгие требования к технологическим аспектам производства наноразмерных устройств электронной техники.

б) Подвижность носителей заряда при механизме движения пролетного типа

Как уже отмечалось выше, пролетная модель подвижности носителей заряда может быть применима для проводящих сред, для которых уместно ввести в рассмотрение понятие длины свободного пробега носителей заряда λ . Типичными представителями таких сред являются:

- газовые среды, в которых в качестве основных носителей заряда выступают электроны и ионы различного зарядового состояния;
- совершенные (с точки зрения фактора регулярности атомной потенциальной диаграммы) монокристаллические структуры металлического и полупроводникового типа, в которых в качестве основных носителей заряда выступают электроны, находящиеся в зоне проводимости и «дырки» – в валентной зоне (см. рис. 1.10).

В свете рассматриваемой проблематики электропроводности кристаллических наноразмерных структур наибольший интерес представляет второй случай. В этих конденсированных средах длина свободного пробега носителей заряда может достигать величин порядка 100 – 300 ангстрем. Из этого следует тот факт, что носители заряда этих сред могут практически беспрепятственно преодолевать указанное выше зна-

чительное расстояние между очередными этапами рассеяния векторизованного движения носителя заряда на дефектах кристаллической структуры и фононах (т.е. колебательных движениях отдельных атомов, образующих кристаллическую структуру).

Проведем анализ характерных моментов модели подвижности носителей заряда, перемещающихся в проводящих средах по механизму (модели) пролетного типа на основе общих принципов и подходов, принятых в классической физике.

В целях упрощения анализа будем считать, что напряженность электрического поля E достаточно мала с точки зрения выполнения условия $v_T \gg v_E$ (см. табл. 2.2), где v_T и v_E – соответственно скорости движения носителя заряда хаотического (теплого) и векторизованного (дрейфового, обусловленного наличием внешнего электрического поля) типа. Считается, что носитель заряда (с величиной электрического заряда, равной q) совершает хаотическое движение в проводящей среде со скоростью v_T , на которое накладывается векторизованная составляющая движения со средней скоростью v_E (рис. 2.2.).

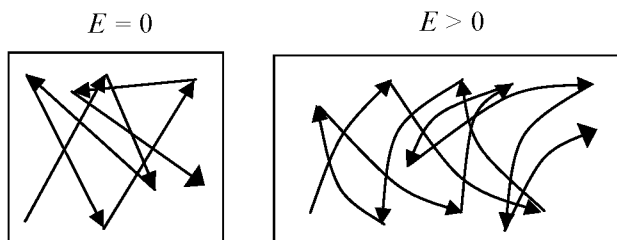


Рис 2.2. Схематическое отображение броуновского ($E = 0$) и векторно-броуновского ($E > 0$) движения носителя заряда

В общем случае векторизованная составляющая движения v_E может быть определена в виде

$$v_E = a / (2 \cdot \tau), \quad (2.6)$$

где $a = (E \cdot q) / m$ – ускорение, испытываемое носителем заряда при значении напряженности электрического поля, равном E ; m – эффективная масса носителя заряда; τ – средняя продолжительность движе-

ния носителя заряда между очередными соударениями (рассеянием) с атомами среды.

Значение средней продолжительности движения носителя заряда τ (между соударениями) можно определить исходя из выражения

$$\tau = \frac{\lambda}{\overline{v_T + v_E}} \approx \frac{\lambda}{v_T}, \quad (2.7)$$

при выводе которого было использовано сформулированное выше допущение о малости векторизованной скорости движения носителя заряда относительно хаотической составляющей движения.

Из выражений (2.6) и (2.7) нетрудно определить, что векторизованная составляющая движения носителя заряда v_E определяется выражением

$$v_E = \left(\frac{\lambda \cdot q}{2 \cdot m \cdot v_T} \right) \cdot E. \quad (2.8)$$

С учетом выражения (2.3), для подвижности носителей заряда при механизме движения пролетного типа имеем

$$\mu_{\text{прол}} = \frac{\lambda \cdot q}{2 \cdot m \cdot v_T}. \quad (2.9)$$

Выражение (2.9) с учетом того обстоятельства, что скорость хаотического движения $v_T = \sqrt{(3 \cdot k \cdot T)/m}$, можно представить в более удобном для расчетной практики виде:

$$\mu_{\text{прол}} = \frac{\lambda \cdot q}{2 \cdot \sqrt{3 \cdot k \cdot m \cdot T}}. \quad (2.10)$$

Определим область применимости выражения (2.10) исходя из принятого ранее допущения: $v_T \gg v_E$. Для этого определим область значений критической напряженности электрического поля $E_{кр}$, при которых это допущение выполняется. Принятое допущение эквивалентно выполнению условия $W_T \gg W_E$, где W_T и W_E – кинетическая энергия соответственно для хаотического и векторизованного типа движения

носителей заряда. Полагая $W_T = (3 \cdot k \cdot T)/2$ и $W_E = q \cdot \lambda \cdot E$, искомую напряженность электрического поля $E_{кр}$ определяем согласно выражению

$$E_{кр} = \frac{3 \cdot k \cdot T}{2 \cdot \lambda \cdot q}. \quad (2.11)$$

В табл. 2.3 представлена зависимость критической напряженности электрического поля $E_{кр}$ (по порядку величины) от длины свободного пробега носителей заряда λ , основанная на выражении (2.11).

Т а б л и ц а 2.3. *Зависимость критической напряженности электрического поля $E_{кр}$ от длины свободного пробега λ носителя заряда*

λ , м	10^{-10} (1 ангстрем)	10^{-9} (10 ангстрем)	10^{-8} (10^2 ангстрем)	10^{-7} (10^3 ангстрем)
$E_{кр}$, В / м	10^8	10^7	10^6	10^5

Согласно представленным данным, пролетная модель подвижности носителей заряда может быть использована в практических приложениях для случая макроскопических монокристаллических структур с электронным типом проводимости, при значениях напряженностей внешнего электрического поля не более 10^5 Вм⁻¹.

Использование пролетной модели подвижности для случая наноразмерных кристаллических структур в большинстве случаев представляется проблематичным. Причиной этого является не столько недостаточность величины λ относительно размеров структуры в направлении пространственных ограничений, сколько сильное негативное влияние поверхности раздела сред (ввиду появления нарушений в периодичности (регулярности) атомной потенциальной диаграммы) на характеристики материала, ответственные за процессы рассеяния векторизованной составляющей движения носителей заряда. По той же причине с большой осторожностью эту модель можно использовать на практике и при проведении качественного (а уж тем более – количественного) анализа процессов токопрохождения в пространственных направлениях, в которых отсутствуют размерные ограничения кристаллических структур.

в) Подвижность носителей заряда при механизме движения прыжкового типа

При механизме движения прыжкового типа, в отличие от рассмотренного выше пролетного типа, элементарный акт перемещения носителя заряда имеет атомарные масштабы. Сам процесс движения носителя заряда можно представить в виде последовательности прыжков (переходов) из одного положения равновесия в другое, пространственное расположение которых совпадает с расположением минимумов на функции потенциальной энергии носителя заряда. Носитель заряда может изменить свое местоположение (совершить прыжок) только лишь в том случае, если его кинетическая энергия в форме колебательного движения превосходит некоторое пороговое значение W_a , получившее название *энергии активации процесса перехода* [2].

Схематическое отображение потенциальной функции кристаллической структуры и диффузионного механизма пространственного перемещения носителя заряда в условиях отсутствия внешнего электрического поля представлено на рис. 2.3.

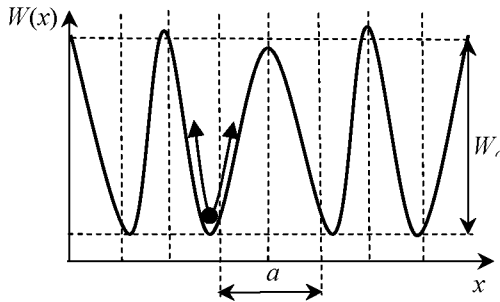


Рис. 2.3. Потенциальная диаграмма регулярной структуры в условиях отсутствия внешнего электрического поля

Для дальнейшего рассмотрения модели энергетической диаграммы, представленной на рис. 2.3, уместно привести следующие комментарии:

- черным кружком, находящимся на дне потенциальной ямы, условно обозначен элементарный носитель заряда, который нахо-

дится (некоторое время) в положении энергетического равновесия;

- стрелками, исходящими от носителя заряда вправо и влево, условно отображен фактор колебательного движения носителя заряда, инициированного с помощью энергии микрочастицы термического характера (считается, что система в целом находится в состоянии термодинамического равновесия);
- W_a (энергия активации) – энергия, необходимая для обеспечения процесса перехода (по механизму термической активации) носителя заряда из одного положения равновесия в другое (соседнее) положение равновесия (равновесные состояния соответствуют минимальным значениям потенциальной энергии на диаграмме);
- a (параметр решетки) – расстояние между соседними атомами кристаллической структуры, в которой инициируется векторизованное движение носителя заряда по пространственному направлению, соответствующему геометрической координате $\{x\}$ (рассматривается одномерное приближение);
- частичное отклонение местоположения по оси x минимумов функции потенциальной диаграммы от регулярного расположения (отображены с помощью вертикальных пунктирных линий, отстоящих друг от друга на расстоянии, равном параметру решетки a) обусловлено проявлением влияния тепловых колебательных движений атомов кристаллической решетки;
- частичное расхождение положения максимумов функции потенциальной диаграммы от регулярного расположения (отображены с помощью верхней горизонтальной пунктирной линии) обусловлено вышеизложенным обстоятельством (зависимость максимумов потенциальной энергии атомов от их взаимного расположения представлена на рис. 1.7 и 1.9).

Характерной особенностью представленной на рис. 2.3 потенциальной диаграммы является то обстоятельство, что она *симметрична по отношению к возможности реализации прыжка носителя заряда как вправо, так и влево*, относительно начального местоположения носителя заряда. Из этого следует, что *вероятности* реализации прыжка впра-

во и влево будут равны, что исключает возможность проявления движения носителя заряда в векторизованной форме. Для определения указанных вероятностей реализации прыжков можно воспользоваться статистикой Максвелла-Больцмана (в этом случае речь идет о носителях заряда, обладающих целочисленными значениями спина (*собственный момент импульса элементарных частиц*)). В этом случае для вероятностей индивидуальных прыжков носителя заряда имеем [3]

$$P_{\text{право}}(E=0) = P_{\text{лево}}(E=0) = \exp\left(-\frac{W_a}{k \cdot T}\right), \quad (2.12)$$

где E – напряженность внешнего электрического поля; $P_{\text{право}}(E=0)$ и $P_{\text{лево}}(E=0)$ – вероятности прыжка носителя заряда соответственно вправо и влево относительно начального расположения.

Таким образом, ввиду равенства вероятностей переходов можно заключить, что при отсутствии внешнего электрического поля носитель заряда будет совершать хаотическое движение (в форме индивидуальных прыжков), что исключает возможность появления электрического тока.

В условиях наличия внешнего электрического поля потенциальная диаграмма кристаллической структуры претерпевает принципиальное изменение (появление наклона), которое отображено на рис. 2.4 (тепловое колебательное движение атомов кристаллической структуры исключено из рассмотрения с целью повышения наглядности при отображении фактора сдвига максимумов потенциальной энергии). Фрагмент $q \cdot U$ (где U — разность потенциалов, приложенная к проводнику от внешнего источника) характеризует степень наклона потенциальной диаграммы под воздействием внешнего электрического поля.

В электрическом поле рассматриваемая регулярная кристаллическая структура теряет симметрию относительно возможности перемещения (прыжка) носителя заряда вправо и влево. В частности, для прыжка по направлению вправо потребуется несколько меньшая энергия (на величину $(-q \cdot a \cdot E/2)$ относительно начального значения энергии активации W_a), чем для прыжка по направлению влево

$(+q \cdot a \cdot E/2)$. Следовательно, прыжки по направлению вправо будут происходить несколько чаще, чем прыжки по направлению влево, а значит, появляется возможность реализации векторизованной формы движения носителя заряда в рассматриваемой регулярной структуре.

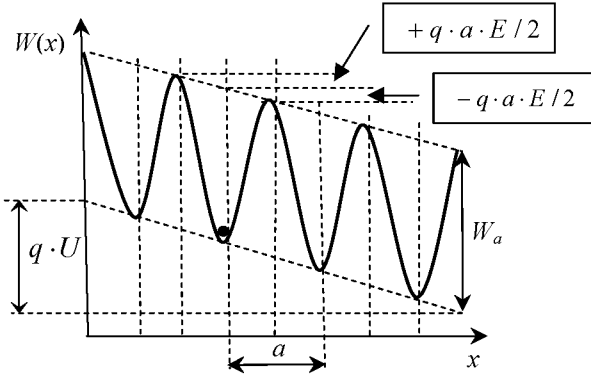


Рис. 2.4. Потенциальная диаграмма регулярной структуры в условиях внешнего электрического поля

Проведем оценку величины разности вероятностей реализации прыжков носителя заряда по направлениям вправо и влево. Для этого обратим внимание на тот факт, что энергетические барьеры, препятствующие прыжку носителя заряда по направлениям вправо и влево, стали разными:

$$P_{\text{право}}(E \neq 0) = \exp \left(- \frac{(W_a - \frac{q \cdot a \cdot E}{2})}{k \cdot T} \right); \quad (2.13a)$$

$$P_{\text{лево}}(E \neq 0) = \exp \left(- \frac{(W_a + \frac{q \cdot a \cdot E}{2})}{k \cdot T} \right), \quad (2.13б)$$

где E – напряженность внешнего электрического поля; q – величина электрического заряда носителя заряда.

Зная вероятности прыжков $P_{\text{право}}(E \neq 0)$ и $P_{\text{лево}}(E \neq 0)$, можно определить частоту переходов носителя заряда по направлениям вправо и влево:

$$v_{\text{право}}(E \neq 0) = v_0 \cdot \exp\left(-\frac{(W_a - q \cdot a \cdot E)}{2 \cdot k \cdot T}\right); \quad (2.14a)$$

$$v_{\text{лево}}(E \neq 0) = v_0 \cdot \exp\left(-\frac{(W_a + q \cdot a \cdot E)}{2 \cdot k \cdot T}\right), \quad (2.14б)$$

где v_0 – частота тепловых колебаний носителя заряда в потенциальной яме (параметр v_0 по порядку величины равен: 10^{15} Гц – электроны; $10^{10} - 10^{12}$ Гц – атомы; $10^6 - 10^8$ Гц – молекулы).

(В выражениях (2.14a) и (2.14б) использовано то представление, что каждый носитель заряда каждую секунду совершает v_0 попыток преодолеть потенциальный барьер, из которых удачными являются те из них, при которых он обладает необходимой для реализации перехода кинетической энергией).

Из выражений (2.14a) и (2.14б) следует, что $v_{\text{право}}(E \neq 0) > v_{\text{лево}}(E \neq 0)$, и эту разницу частот перехода (прыжков) $\Delta v = v_{\text{право}}(E \neq 0) - v_{\text{лево}}(E \neq 0)$ можно определить в виде

$$\Delta v = v_0 \cdot \exp\left(-\frac{W_a}{k \cdot T}\right) \cdot \left[\exp\left(\frac{q \cdot a \cdot E}{2 \cdot k \cdot T}\right) - \exp\left(-\frac{q \cdot a \cdot E}{2 \cdot k \cdot T}\right) \right]. \quad (2.15)$$

В целях упрощения анализа выражения (2.15) примем допущение, что выполняется условие $(q \cdot a \cdot E) \ll (2 \cdot k \cdot T)$. В этом случае с помощью представления аналитических функций в виде степенного ряда можно получить выражения:

$$\exp\left(\frac{q \cdot a \cdot E}{2 \cdot k \cdot T}\right) \approx 1 + \frac{q \cdot a \cdot E}{2 \cdot k \cdot T} ; \quad (2.16)$$

$$\exp\left(-\frac{q \cdot a \cdot E}{2 \cdot k \cdot T}\right) \approx 1 - \frac{q \cdot a \cdot E}{2 \cdot k \cdot T} .$$

Подставляя выражение (2.16) в выражение (2.15) для определения разности частот перехода носителя заряда, получим

$$\Delta \nu = \left[\nu_0 \cdot \left[\exp\left(-\frac{W_a}{k \cdot T}\right) \right] \cdot \left(\frac{q \cdot a}{k \cdot T}\right) \right] \cdot E. \quad (2.17)$$

С помощью выражения (2.17) можно определить среднюю скорость движения носителя заряда в электрическом поле:

$$\nu(E) = a \cdot \Delta \nu = \left[a \cdot \nu_0 \cdot \left[\exp\left(-\frac{W_a}{k \cdot T}\right) \right] \cdot \left(\frac{q \cdot a}{k \cdot T}\right) \right] \cdot E. \quad (2.18)$$

Согласно определению понятия подвижности носителя заряда (см. выражение(2.3)), среднюю скорость движения носителя заряда в рамках прыжкового механизма перемещения микрочастиц можно определить в виде

$$\nu(E) = \mu_{\text{прыжк}} \cdot E. \quad (2.19)$$

Сравнивая выражения (2.18) и (2.19), для прыжковой подвижности носителя заряда $\mu_{\text{прыжк}}$ имеем

$$\mu_{\text{прыжк}} = \frac{\nu_0 \cdot q \cdot a^2}{k \cdot T} \cdot \exp\left(-\frac{W_a}{k \cdot T}\right). \quad (2.20)$$

При выводе выражения (2.20) было использовано допущение, что $(q \cdot a \cdot E) \ll (2 \cdot k \cdot T)$. Исходя из этого обстоятельства, определим область применимости выражения (2.20). Для этого определим область значений критической напряженности электрического поля $E_{\text{кр}}$ согласно выражению

$$E_{\text{кр}} = \frac{2 \cdot k \cdot T}{q \cdot a}. \quad (2.21)$$

В табл. 2.4 представлена зависимость критической напряженности электрического поля $E_{кр}$ от температуры среды T , в которой происходит движение носителя заряда по прыжковому механизму (использовано выражение (2.21)).

Т а б л и ц а 2.4. *Зависимость критической напряженности электрического поля $E_{кр}$ от температуры проводящей среды ($q = 1.6 \cdot 10^{19}$ Кл; $a = 10^{-10}$ м)*

T, K	3	30	300	600	900	1200
$E_{кр}, В / м$	$5 \cdot 10^6$	$5 \cdot 10^7$	$5 \cdot 10^8$	10^9	$1.5 \cdot 10^9$	$2 \cdot 10^9$

Согласно представленным данным, прыжковая модель подвижности носителей заряда может быть использована для случая макроскопических монокристаллических структур (с ионным механизмом электропроводности) при значениях напряженностей внешнего электрического поля не более 10^6 В·м⁻¹. Это условие выполняется для большинства практических приложений, связанных с созданием элементов конструкций устройств электронной техники.

з) Механизм движения носителей заряда туннельного типа

При механизме движения туннельного типа элементарный акт перемещения носителя заряда в определенной мере сходен с рассмотренным выше механизмом прыжкового типа, но протяженность элементарного прыжка имеет пространственную протяженность более чем атомного масштаба. Сам процесс движения носителя заряда можно представить в виде своеобразного «просачивания» носителя заряда сквозь потенциальный барьер, минуя необходимость его преодоления по рассмотренному выше активационному механизму. Иными словами, для этого механизма движения свойственно то обстоятельство, что носителю заряда не требуется для реализации процесса перехода обладать энергией, большей чем максимальная высота потенциального барьера.

Следует отметить тот факт, что туннельный принцип преодоления носителем заряда потенциального барьера, стоящего на пути движения микрочастицы, имеет исключительно квантовую природу. Согласно основным принципам классической физики, микрочастица может находиться лишь в тех областях пространства, в которых ее потенциальная

энергия не превышает полную энергию. Это следует из того обстоятельства, что кинетическая энергия частицы не может быть отрицательной, т.к. в этом случае импульс частицы становится мнимой величиной. Исходя из этого обстоятельства следует, что если две области пространства разделены потенциальным барьером конечной протяженности (таким, что выполняется условие $W_{\text{потенц}} > W_{\text{полной}}$), то просачивание частицы сквозь него в рамках представлений классической теории оказывается невозможным.

С физической точки зрения туннельный эффект можно объяснить как своеобразное проявление принципа неопределенностей В. Гейзенберга (см. выражение 1.21). Согласно этому принципу следует, что при ограничении местоположения квантовой частицы по некоторой координате (например x), т.е. увеличении степени определённости ее нахождения по указанной координате x , её импульс p_x становится все менее определённым. Неопределённость импульса Δp (что приводит также к появлению неопределенности кинетической энергии ΔW_x) может добавить частице недостающую для преодоления барьера часть энергии. Таким образом, можно сделать вывод о том, что квантовая микрочастица с *некоторой вероятностью* может проникнуть через потенциальный барьер, обладая энергией, значительно меньшей чем высота преодолеваемого потенциального барьера W_a .

Рассмотрим в одномерном приближении процесс прохождения микрочастицы через потенциальный барьер на основании вероятностных волновых представлений, представленных в разд. 1.1. Допустим, что микрочастица движется вдоль оси x в области пространства, в которой потенциальная диаграмма претерпевает изменения согласно рис. 2.5. Характерной особенностью потенциальной диаграммы является следующее обстоятельство:

- в области (1), простирающейся по координате x от $-\infty$ до 0, потенциальная энергия рассматриваемой системы $U(x) = 0$;
- в области (2), простирающейся по координате x от 0 до $+\infty$, потенциальная энергия рассматриваемой системы $U(x) = U_0$.

В общем случае, процесс движения квантовой микрочастицы описывается одномерным амплитудным уравнением Шредингера (см. разд. 1):

$$\frac{d^2\phi(x)}{dx^2} + \frac{8 \cdot \pi^2 \cdot m}{h^2} \cdot (W - U(x)) \cdot \phi(x) = 0, \quad (2.22)$$

которое для рассматриваемых областей пространства (1) и (2) трансформируется в амплитудные уравнения вида

$$\frac{d^2\phi_1(x)}{dx^2} + k_1^2 \cdot \phi_1(x) = 0; \quad (2.23)$$

$$\frac{d^2\phi_2(x)}{dx^2} + k_2^2 \phi_2(x) = 0, \quad (2.24)$$

где $k_1 = \frac{2\pi}{h} \cdot \sqrt{2 \cdot m \cdot W}$; $k_2 = \frac{2\pi}{h} \cdot \sqrt{2 \cdot m \cdot (W - U_0)}$; W – кинетическая энергия движения микрочастицы в рассматриваемом силовом поле.

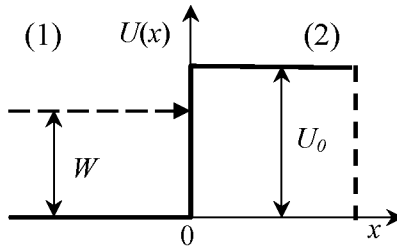


Рис.2.5. Схематическое отображение потенциальной диаграммы бесконечно протяженного энергетического барьера

Согласно материалам разд. 1, общие решения уравнений (2.23) и (2.24) имеют следующий вид:

$$\phi_1(x) = A_1 \cdot \exp(i \cdot k_1 \cdot x) + B_1 \cdot \exp(-i \cdot k_1 \cdot x), \text{ при } x < 0; \quad (2.25)$$

$$\phi_2(x) = A_2 \cdot \exp(i \cdot k_2 \cdot x) + B_2 \cdot \exp(-i \cdot k_2 \cdot x), \text{ при } x > 0. \quad (2.26)$$

Фрагмент $A_1 \cdot \exp(i \cdot k_1 \cdot x)$ характеризует волну де Бройля, распространяющуюся в области (1) в положительном направлении оси x , где

A_1 – амплитуда волны вероятности местоположения микрочастицы. Соответственно, фрагмент $B_1 \cdot \exp(-i \cdot k_1 \cdot x)$ характеризует волну вероятности, распространяющуюся в области (1) в отрицательном направлении оси x , где B_1 – амплитуда волны, отраженной от потенциального барьера.

Поскольку квадрат амплитуды волны де Бройля равен вероятности нахождения микрочастицы в определенной точке пространства, то отношение

$$R = \frac{|B_1|^2}{|A_1|^2} \quad (2.27)$$

характеризует отражательную способность барьера, а параметр R получил название *коэффициента отражения* движения микрочастицы от потенциального барьера (ранее, при рассмотрении движения частицы в потенциальной яме, по умолчанию считалось, что $R = 1$).

Фрагмент $A_2 \cdot \exp(i \cdot k_2 \cdot x)$ характеризует волну де Бройля, распространяющуюся в области (2) в положительном направлении оси x . Квадрат амплитуды этой волны характеризует вероятность проникновения микрочастицы в область пространства (2). Отношение вида

$$D = n \cdot \frac{|A_2|^2}{|A_1|^2} \quad (2.28)$$

характеризует проникающую способность барьера, а параметр D получил название *коэффициента прозрачности* движения микрочастицы. Параметр $n = \lambda_1 / \lambda_2$ получил название *показателя преломления* вероятностных волн де Бройля.

Фрагмент $B_2 \cdot \exp(-i \cdot k_2 \cdot x)$ характеризует волну вероятности, распространяющуюся в области (2) в отрицательном направлении оси x . Поскольку в области (2) отсутствует как таковая отражающая граница, а следовательно, нет отраженной волны, то следует принять, что амплитуда волны $B_2 = 0$.

Для *низких потенциальных барьеров*, для которых справедливо выполнение условия $W > U_0$, коэффициент отражения R равен [4]

$$R = \left(\frac{k_1 - k_2}{k_1 + k_2} \right). \quad (2.29)$$

Выражение (2.29) представляется весьма примечательным в том плане, что оно свидетельствует о принципиальном расхождении подходов классической и квантовой физики. Согласно классической физике, при выполнении условия $W > U_0$ коэффициент отражения R должен быть тождественно равен нулю. Согласно квантовой физике, при выполнении этого же условия коэффициент отражения $R > 0$, даже если W значительно превосходит U_0 . Для иллюстрации этого момента в табл. 2.5 приведены значения коэффициента отражения R для различных значений отношения $W/U_0 \geq 1$.

Таблица 2.5. Зависимость коэффициента отражения R от значения отношения w/U_0 [4]

W/U_0	1.0	1.11	1.25	2.0	10
R	1.0	0.27	0.146	0.03	0.0007

Для *высоких потенциальных барьеров*, для которых справедливо выполнение условия $W < U_0$, волновое число k_2 в выражении (2.26) является мнимым числом: $k_2 = i \cdot k$, где $k = \frac{2\pi}{h} \cdot \sqrt{2 \cdot m_e \cdot (U_0 - W)}$ – действительное число. В этом случае для волновой функции в области (2) имеем (обратить внимание на тот факт, что волновая функция уже не является ПМВ, как это имело место выше) выражение: $\phi_2(x) = A_2 \cdot \exp(-k \cdot x)$.

Для потенциального барьера конечной толщины d (в области (2) является задняя стенка, отображенная на рис. 2.5 в виде утолщенной вертикальной пунктирной линии) вероятность обнаружения микрочастицы (прошедшей сквозь переднюю стенку барьера) у задней стенки барьера равна

$$P_d = |\phi_2|^2 = A_2^2 \cdot \exp\left(-\frac{4 \cdot \pi \cdot d}{h} \cdot \sqrt{2 \cdot m_e \cdot (U_0 - W)}\right). \quad (2.30)$$

Отличие вероятности P_d от нуля свидетельствует о возможности прохождения (просачивания) микрочастицы сквозь потенциальный барьер конечной толщины d даже при выполнении условия $W < U_0$.

Этот неожиданный (для классической физики) квантовый феномен получил название *туннельного эффекта*.

Рассмотренный выше коэффициент прозрачности D для случая потенциального барьера конечной толщины равен [4]

$$D = D_0 \cdot \exp\left(-\frac{4 \cdot \pi \cdot d}{h} \cdot \sqrt{2 \cdot m_e \cdot (U_0 - W)}\right), \quad (2.31)$$

где D_0 – постоянный коэффициент, близкий по порядку величины единице.

Выражение (2.31) может быть обобщено и для потенциальных барьеров произвольной геометрической конфигурации:

$$D = D_0 \cdot \exp\left(-\frac{4 \cdot \pi \cdot d}{h} \cdot \int_{x_1}^{x_2} \sqrt{2 \cdot m_e \cdot (U_0 - W)} \cdot dx\right). \quad (2.32)$$

В табл. 2.6 в качестве примера, характеризующего масштабы проявления туннельного механизма преодоления потенциального барьера, представлены значения коэффициента прозрачности D (рассчитанные согласно выражению (2.31)), соответствующие различным значениям толщины барьера d и разности между значениями максимума потенциальной энергии (см. рис. 2.5) и кинетической энергии поступательного движения микрочастицы – $(U_0 - W)$.

Таблица 2.6. *Зависимость коэффициента прозрачности D потенциального барьера от толщины барьера d и разности энергий $(U_0 - W)$*

$D, \cdot 10^{-10}$ м	1.0	1.5	2.0	5.0	10
$D(U_0 - W = 0.5 \text{ эВ})$	0.49	0.34	0.24	0.027	$7 \cdot 10^{-4}$
$D(U_0 - W = 1.0 \text{ эВ})$	0.36	0.22	0.13	$6 \cdot 10^{-3}$	$4 \cdot 10^{-5}$
$D(U_0 - W = 2.5 \text{ эВ})$	0.20	0.089	0.04	$3 \cdot 10^{-4}$	$1 \cdot 10^{-7}$
$D(U_0 - W = 5.0 \text{ эВ})$	0.10	0.032	0.01	$1 \cdot 10^{-5}$	$1 \cdot 10^{-10}$

Исходя из данных табл. 2.6, можно сделать ряд выводов, представляющих практический интерес:

- потенциальные барьеры, расположенные между индивидуальными атомами кристаллической структуры и имеющие, как правило, протяженность порядка 10^{-10} м, обладают достаточно высокой прозрачностью при прохождении через них квантовых микрочастиц;
- коэффициент прозрачности D очень сильно уменьшается при увеличении протяженности потенциального барьера d и, в частности, может быть принят равным нулю при значениях d , превышающих 10^{-9} м.

Еще один примечательный фактор квантового подхода (расходящийся с классическим подходом) состоит в том, что при туннельном механизме прохождения сквозь потенциальный барьер энергия микрочастицы не претерпевает изменения, т.е. *квантовая микрочастица покидает потенциальный барьер с тем же значением энергии, с какой она изначально входила в него.*

2.1.3. Концентрация носителей заряда в проводниках электронного типа

а) Механизм электропроводности кристаллических структур с позиции зонных энергетических диаграмм

В общем случае в конструкционных материалах с электронным типом проводимости *общая концентрация электронов* превышает величину $10^{22} - 10^{23}$ электронов/см³ (в том числе и в диэлектрических материалах). Если все эти электроны будут задействованы в процессе электропроводности, то электрическая проводимость всех типов материалов будет достигать чрезвычайно высоких значений, что не наблюдается в реальной научно-технической практике.

В этой связи возникает вопрос, какие из оговоренных выше электронов следует считать непосредственными участниками процесса электрической проводимости? Ответ на этот вопрос позволит сформировать механизмы электропроводности кристаллических структур, обладающих электронным типом проводимости. Этот вопрос тем более важен, если учесть то обстоятельство, что без ответа на него невозмож-

но поставить на научную основу методы проектирования современных электронных устройств.

В настоящем разделе анализ вышеизложенного вопроса представлен на качественном уровне рассмотрения и в дальнейшем его результаты будут перекрыты более корректным количественным анализом, основанном на использовании статистики заполнения электронами энергетических уровней зонной диаграммы.

Рассмотренные выше (см. разд. 1, рис. 1.10) зонные энергетические диаграммы кристаллических структур позволяют ответить на этот вопрос. Для определения (на качественном уровне анализа) понятия о концентрации носителей заряда в кристаллических структурах, обладающих электронным типом проводимости, обратимся к более детальному рассмотрению зонной энергетической диаграммы собственного полупроводника (т.е. в отсутствие легирующей примеси донорного или акцепторного типа).

На рис. 2.6 представлен характерный вид энергетической диаграммы собственного полупроводника при температуре окружающей среды $T \approx 0$ К и в условиях отсутствия внешнего электрического поля ($E = 0$ В/м).

Для уяснения физической сути рассматриваемого вопроса прокомментируем некоторые позиции рис. 2.6:

- точками отображены электроны (не более двух – согласно принципу Паули), расположенные на индивидуальном энергетическом уровне;
- векторами, исходящими от отдельных электронов, отображены направление и величина скорости поступательного движения электрона, находящегося на рассматриваемом энергетическом уровне;
- W_c – значение энергии, соответствующее наиболее низко расположенному энергетическому уровню в зоне проводимости собственного полупроводника;
- W_v – значение энергии, соответствующее наиболее высоко расположенному энергетическому уровню в валентной зоне собственного полупроводника;
- ΔW – ширина запрещенной зоны собственного полупроводника;

- посредством равенства модулей векторов движения электронов, находящихся на конкретном энергетическом уровне, отражен тот факт, что движение электронов по направлениям вправо и влево компенсируется в части переноса электрического заряда (из этого следует, что в целом не будет иметь место перенос заряда в веществе, а следовательно, и не будет макроскопического электрического тока);
- посредством неравенства модулей векторов движения электронов, находящихся на разных энергетических уровнях, отражено то обстоятельство, что движение ансамбля электронов имеет хаотический характер, обусловленный влиянием тепловых факторов.

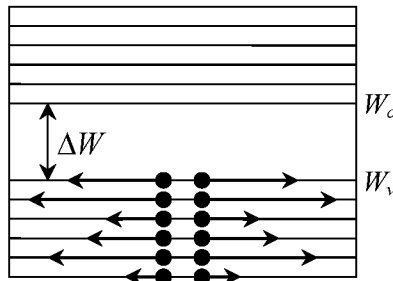


Рис. 2.6. Зонная энергетическая диаграмма собственного полупроводника при температуре $T \approx 0$ К

Согласно модельному представлению (см. рис. 2.6) состояния электрона в кристаллической структуре необходимо отметить следующие моменты:

- каждый электрон, находящийся на энергетическом уровне, расположенном в разрешенной энергетической зоне полупроводника, обладает кинетической энергией, реализуемой в форме хаотического *поступательного движения* в пространственных пределах кристаллической структуры;
- движение электронов имеет симметричный характер, с той точки зрения, что интенсивности движения ансамблей электронов проводимости по направлениям «вправо» и «влево» являются одинаковыми, т.е. вносящими эквивалентный вклад в процесс переноса носителей заряда в материале;

- средняя кинетическая энергия теплового движения электронов (которая равна по порядку величины $W_T = kT = 0,025$ Эв, при $T = 300$ К) считается настолько малой по сравнению с шириной запрещенной зоны ΔW (более 1 Эв), что можно исключить из рассмотрения процессы перехода электронов из валентной зоны в зону проводимости, стимулированные тепловыми флуктуациями.

Поместим полупроводник в электрическое поле с напряженностью E . В этом случае на каждый электрон, расположенный в валентной зоне, будет действовать механическая сила $F = q \cdot E$, под воздействием которой он, следуя логике представлений классической физики, должен получить ускорение $a = (q \cdot E)/m$. Следствием подобного рода ускоренного движения электрона в электрическом поле будет нарушение вышеупомянутой симметрии макроскопического переноса носителей заряда по направлениям «вправо» и «влево» (к примеру, все стрелки на рис. 2.6 по направлению «вправо» должны удлиниться, а по направлению «влево» – укоротиться).

Однако возможна ли такая ситуация в рассматриваемой системе? Ведь если электроны изменят скорость своего движения, то они тем самым изменят свою кинетическую, а следовательно, и полную энергию (она отложена по оси ординат зонной энергетической диаграммы). В этом случае они уже не могут оставаться на прежних энергетических уровнях и должны перейти на другие *свободные энергетические уровни* (с большей или меньшей энергией, в зависимости от того, увеличилась или уменьшилась скорость движения электрона под воздействием электрического поля). Фактор отсутствия свободных энергетических уровней в валентной зоне *в принципе исключает возможность изменения скорости* находящегося в ней электрона и в этой связи возникает парадоксальная ситуация, при которой электрон «игнорирует» факт наличия внешнего электрического поля (принято считать, что при отсутствии свободных энергетических уровней эффективная масса электрона стремится к бесконечности).

Ситуация меняется коренным образом, если часть электронов перемещается из валентной зоны в зону проводимости полупроводника. Такое перемещение электронов вполне возможно, если часть электро-

нов валентной зоны получит дополнительную энергию, достаточную для преодоления запрещенной зоны полупроводника (в рассматриваемом случае запрещенная зона выступает в качестве потенциального барьера, имеющего высоту ΔW (см. рис. 2.6)). В этом случае *в собственном полупроводнике появятся две, частично заполненные электронами энергетические зоны (валентная и проводимости)*. Сам факт частичного заполнения энергетических зон электронами открывает возможность изменения энергетического состояния этих электронов под воздействием внешнего электрического поля.

Это обстоятельство нашло свое отражение на рис. 2.7, иллюстрирующем механизм процессов ускорения (и замедления) движения электрона, находящегося в пределах отдельной энергетической зоны.

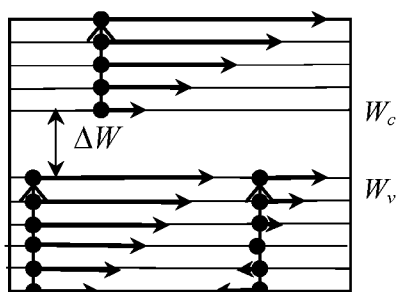


Рис. 2.7. Схематическое отображение эволюции энергетического состояния электронов при их движении во внешнем электрическом поле

Обращаясь к рис. 2.7, следует заострить внимание на трех принципиальных моментах, отражающих специфику процесса электронной проводимости в кристаллических структурах полупроводникового и металлического типа:

- обеспечение процесса электропроводности сопряжено с необходимостью создания благоприятных условий, при которых становится возможным изменение энергетического состояния электронов (наличие незаполненных энергетических уровней в рассматриваемой энергетической зоне);
- реализация процесса электрической проводимости происходит путем последовательности прыжков электронов с одного энергетического уровня на другой, расположенный выше или ниже исходного

уровня (расстояние между соседними уровнями, по порядку величины, равно 10^{-23} Эв, поэтому такие прыжки по энергетической шкале сравнительно легко осуществимы);

- с помощью фактора неравенства изменения скоростей при осуществлении единичного прыжка *условно* отображено расхождение величин подвижностей носителей заряда в зоне проводимости и в валентной зоне (в зоне проводимости величина подвижности носителей заряда несколько больше).

На основании вышеизложенных моментов можно сформулировать итоговый вывод: *электропроводность кристаллических тел с электронным типом проводимости возможна лишь в том случае, когда реализуется частичное заполнение электронами энергетических уровней в разрешенной энергетической зоне.*

С этой точки зрения в качестве носителей заряда следует рассматривать не все электроны, которые содержатся в кристаллической структуре, а лишь часть из них, а именно:

А) для собственных полупроводников это:

- совокупность электронов, появившихся в зоне проводимости в результате их перемещения туда из валентной зоны;
- совокупность электронов в валентной зоне, численно равная (по порядку величины) количеству незаполненных энергетических уровней (дырки) в этой зоне, появившихся в результате перехода электронов в зону проводимости.

Б) для легированных (примесного типа) полупроводников это:

- совокупность электронов (полупроводники «*n*-типа»), находящихся в зоне проводимости в результате их перемещения туда из донорного уровня (фактически узкой энергетической зоны);
- совокупность электронов (полупроводники «*p*-типа») в валентной зоне, численно равная (по порядку величины) количеству уровней в акцепторной зоне (ввиду ее узости часто называемой – уровнем), на которые переходят электроны валентной зоны.

В материалах металлического типа ситуация незаполненной энергетической зоны возникает автоматически, поскольку у них имеет место пересечение (смыкание) зоны проводимости и валентной зоны (т.е. за-

прешенная зона отсутствует как таковая). Именно поэтому в этих материалах не наблюдается недостатка в носителях заряда, что, собственно, и обуславливает достаточно высокие показатели электропроводности конструкционных материалов металлического типа.

В материалах полупроводникового типа ситуация незаполненной энергетической зоны создается искусственно, путем [5]:

- интенсивного энергетического воздействия (теплого, светового, радиационного или электрического (случай сильного поля) характера) на электроны, расположенные в валентной зоне, с целью их перемещения в зону проводимости;
- диффузионного (термического) или радиационного (ионная имплантация) внедрения в кристаллическую решетку материала совокупности специфических инородных атомов (легирование материала), образующих в запрещенной зоне полупроводника дополнительную разрешенную энергетическую зону вблизи дна зоны проводимости (изначально заполненную электронами; полупроводник «*n*-типа») или вблизи потолка валентной зоны (изначально свободную от электронов; полупроводник «*p*-типа»).

Рассмотрим еще одну точку зрения на возможный механизм токопрохождения, не приводящую к необходимости наличия незаполненных энергетических зон (при всей ее ошибочности, она, тем не менее, полезна в той части, что позволяет высветить ряд полезных количественных представлений и аргументов в пользу рассмотренной выше позиции). В принципе, можно допустить гипотетический механизм «обменного» способа перемещения, при котором одна группа носителей заряда будет перескакивать вниз, на соседние энергетические уровни (уменьшая тем самым скорость своего движения под воздействием электрического поля), на смену ушедших *синхронно* с ними вверх (увеличивая тем самым скорость своего движения) другой группы носителей заряда. В этом случае не возникает необходимости в наличии свободных уровней в энергетической зоне. Рассмотрим вопрос об оговоренной выше синхронности процесса обмена местоположениями электронов с конструктивной точки зрения. Примем, что длина свободного пробега электронов в материале равна $\lambda = 10^{-8}$ м (типичное зна-

чение). В этом случае значение дополнительной энергии, приобретаемой электроном от электрического поля с напряженностью $E = 10^2$ В/м, составит $W_E = q \cdot \lambda \cdot E = 10^{-5}$ Эв. Учитывая тот факт, что расстояние между соседними энергетическими уровнями равно (по порядку величины) $\Delta W = 10^{-23}$ Эв, получим, что для обеспечения длины свободного пробега $\lambda = 10^{-8}$ Эв электрону необходимо осуществить 10^{18} (!) синхронизированных обменных перемещений (процессов). Столь грандиозное количество необходимых обменных процессов ставит под сомнение правомерность использования гипотезы об «обменном» характере механизма электропроводности для случая кристаллических структур.

б) Функция плотности и количества квантовых состояний

Определим число индивидуальных квантовых состояний, которые может занимать микрочастица, обладающая энергией в интервале значений $W - (W+dW)$.

Рассмотрение этого вопроса удобно начать с позиции представлений классической физики, в которой энергетическое состояние микрочастиц *однозначно* определяется заданием ее трех координат $\{x, y, z\}$ и трех проекций импульса $\{p_x, p_y, p_z\}$. Действительно, задание координат микрочастицы позволяет однозначно определить ее потенциальную энергию W_n , а проекций импульса – ее кинетическую энергию W_k , а следовательно, и полную энергию микрочастицы $W = W_n + W_k$. Именно в связи с последним обстоятельством принято отображать энергетическое состояние микрочастицы в виде местоположения *фазовой* точки в воображаемом шестимерном пространстве $\{x, y, z, p_x, p_y, p_z\}$, получившем название *фазового пространства* состояния системы. Дальнейшая эволюция системы может быть отображена в этом фазовом пространстве посредством перемещения указанной фазовой точки по однозначно определенной *фазовой траектории*.

Введем в рассмотрение понятие об элементарном объеме фазового пространства состояния микрочастицы:

$$\Delta\Gamma = \Delta\Gamma_V \cdot \Delta\Gamma_P = dx \cdot dy \cdot dz \cdot dp_x \cdot dp_y \cdot dp_z, \quad (2.33)$$

где $\Delta\Gamma_V = dx \cdot dy \cdot dz$ – элемент объема в координатном подпространстве; $\Delta\Gamma_P = dp_x \cdot dp_y \cdot dp_z$ – элемент объема в импульсном подпространстве.

Поскольку в классической физике не накладывается никаких ограничений на допустимые энергетические состояния микрочастицы, то элементы объемов $\Delta\Gamma_V$ и $\Delta\Gamma_P$, а следовательно и $\Delta\Gamma$, могут быть бесконечно малыми величинами. Что, собственно, и отражает фактор однозначности определения состояния микрочастицы.

Для упрощения дальнейшего анализа остановимся на случае свободных частиц, для которого характерно отсутствие внешних потенциальных (силовых) полей. В этом случае энергетическое состояние микрочастицы становится не зависимым от местоположения микрочастицы в физическом объеме V (однако предполагается, что заведомо известно о нахождении микрочастицы в пределах этого объема). Исходя из сделанного допущения, можно принять, что элементарный фазовый объем $\Delta\Gamma_V = V$ и в связи с этим выражение (2.33) можно преобразовать к виду

$$\Delta\Gamma = V \cdot \Delta\Gamma_P = V \cdot dp_x \cdot dp_y \cdot dp_z. \quad (2.34)$$

Следуя материалам разд. 1, движение микрочастиц описывается в рамках волновых (вероятностных) представлений, согласно которым их состояние подчиняется принципу неопределенностей Гейзенберга (см. выражение (1.25)). Принципиальное отличие квантово-механического подхода от классического состоит в том, что в этом случае состояние микрочастицы отображается в фазовом объеме не в виде точки (классический подход), а посредством конечной области фазового пространства объемом h^3 . С учетом этого обстоятельства элементарный фазовый объем $\Delta\Gamma$ можно представить в виде

$$\Delta\Gamma = \Delta\Gamma_V \cdot \Delta\Gamma_P = V \cdot \Delta\Gamma_P = h^3. \quad (2.35)$$

Из выражения (2.35) можно определить объем элементарного фазового импульсного подпространства $\Delta\Gamma_P$, приходящегося на одну микрочастицу:

$$\Delta\Gamma_p = \frac{h^3}{V}. \quad (2.36)$$

Процедура деления общего фазового пространства на индивидуальные фазовые ячейки (объемом h^3/V или в общем случае – h^3), определяющие возможные квантовые состояния микрочастицы, называется *квантованием фазового пространства*. На рис. 2.8 приведено схематическое представление импульсного фазового подпространства энергетического состояния микрочастицы с учетом фактора сферической симметрии при отображении импульса микрочастицы в части выполнения соотношения:

$$p^2 = p_x^2 + p_y^2 + p_z^2.$$

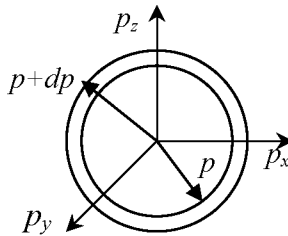


Рис. 2.8. Импульсное фазовое подпространство энергетического состояния микрочастицы

Определим число квантовых состояний, которые может иметь микрочастица, имеющая энергию, заключенную в интервале значений $W - (W+dW)$. Для этого выделим в фазовом пространстве (см. рис. 2.8) две сферические поверхности, соответствующие значениям импульса p (энергии W) и $(p+dp)$ (энергии $(W+dW)$). Между указанными сферическими поверхностями заключен шаровой слой объемом $\Delta\Gamma_c = 4 \cdot \pi \cdot p^2 \cdot dp$. Общее количество квантовых состояний dG_c (т.е. фазовых ячеек объемом $\Delta\Gamma_p$, см. выражение (2.36)), заключенных в этом шаровом слое, будет равно

$$dG_c = \frac{\Delta\Gamma_c}{\Delta\Gamma_p} = \frac{4 \cdot \pi \cdot V}{h^3} \cdot p^2 \cdot dp. \quad (2.37)$$

Введем в рассмотрение функцию плотности квантовых состояний $g(p)$ в импульсном пространстве, такую, чтобы выполнялось условие

$$dG_c = g(p) \cdot dp = \frac{4 \cdot \pi \cdot V}{h^3} \cdot p^2 \cdot dp. \quad (2.38)$$

Для перехода от параметра p к параметру W воспользуемся очевидными выражениями (с учетом того обстоятельства, что микрочастица является свободной, а следовательно, ее кинетическая энергия равна полной энергии): $W = (p^2)/(2m)$; $dW = \left(\frac{p}{m}\right) \cdot dp$. Переходя к энергетическому представлению функции плотности квантовых состояний, имеем

$$g(W) \cdot dW = \frac{2 \cdot \pi \cdot V}{h^3} \cdot (2 \cdot m)^{3/2} \cdot W^{1/2} \cdot dW. \quad (2.39)$$

Из выражения (2.39) следует, что *плотность квантовых состояний* $g(W)$, выражающая число состояний микрочастицы, приходящееся на единичный интервал энергий, определяется выражением

$$g(W) = \frac{2 \cdot \pi \cdot V}{h^3} \cdot (2 \cdot m)^{3/2} \cdot W^{1/2}. \quad (2.40)$$

Интегрируя выражение (2.40) в пределах от W_1 до W_2 ($W_1 < W_2$), можно получить количество квантовых состояний $G(W_1, W_2)$ электронов, содержащихся в указанном интервале энергий:

$$G(W_1, W_2) = \frac{8 \cdot \pi \cdot V}{3 \cdot h^3} \cdot (2 \cdot m)^{3/2} \cdot \left(W_2^{3/2} - W_1^{3/2} \right). \quad (2.41)$$

Из выражения (2.40) следует, что с ростом массы микрочастицы плотность квантовых состояний $g(W)$ возрастает. Именно с учетом этого обстоятельства *считается, что макроскопические тела имеют практически непрерывный (а не дискретный, как это имеет место для случая единичного атома или наноразмерных структур) спектр энергетических состояний и в их поведении, как правило, не наблюдается*

проявлений квантово-механического характера. Именно благодаря фактору этой непрерывности спектра энергетических состояний, макроскопические объекты становятся предметом рассмотрения методами классической физики. Этот переход от случая дискретного спектра энергетических состояний (объект – единичный атом) к случаю непрерывного спектра (объект – макроскопическое тело объемом в 1 см^3 , с количеством атомов порядка 10^{23}) условно отображен на рис. 2.9. На рисунке отражен акцент на величину масштабного фактора расхождения количества квантовых состояний единичного атома и макроскопического объекта, приходящихся на определенный (достаточно малый) энергетический интервал.

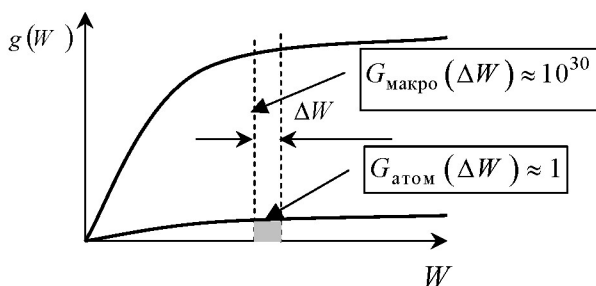


Рис.2.9. Плотность и количество квантовых состояний для атомных и макроскопических объектов

В случае рассмотрения электронов в качестве микрочастиц каждой фазовой ячейке отвечает не одно, а два состояния, отличающихся друг от друга направлением спина (принцип Паули). С учетом этого для плотности квантовых состояний электронов имеем

$$g(W) = \frac{4 \cdot \pi \cdot V}{h^3} \cdot (2 \cdot m)^{3/2} \cdot W^{1/2}.$$

Весьма важным и удобным для практических целей является то обстоятельство, что вид функции плотности квантовых состояний $g(W)$ (см. выражение (2.40)) является *одинаковым* для всех типов микрочастиц (электроны, фотоны, атомы). Однако эта функция может отслеживать специфику ограничений на вид энергетического спектра разрешенных состояний кристаллических структур.

В частности, для случая собственного полупроводника, зонная диаграмма которого содержит запрещенную зону, функция плотности квантовых состояний $g(W)$ имеет вид, представленный на рис. 2.10.

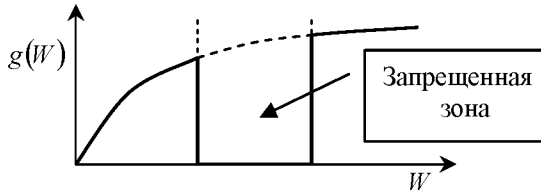


Рис. 2.10. Функция плотности квантовых состояний $g(W)$ для случая собственного полупроводника

Для общности рассмотрения и в силу их большой практической значимости на рис. 2.11 представлены функции плотности квантовых состояний $g(W)$ для случая примесных полупроводников донорного и акцепторного типа.

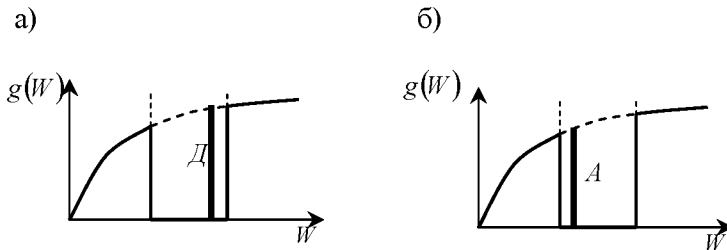


Рис. 2.11. Функции плотности квантовых состояний $g(W)$ для случая донорного (а) и акцепторного (б) полупроводника

На рис. 2.11 использованы следующие условные обозначения: А – акцепторный уровень; Д – донорный уровень (*под понятием «уровень» следует понимать очень узкую энергетическую зону с количеством энергетических уровней, соизмеримым с количеством внесенных в полупроводник атомов легирующей примеси*).

*в) Вероятность заполнения квантовых состояний
в кристаллических структурах*

В предыдущем разделе рассматривался вопрос о плотности и количестве квантовых состояний, рассматриваемых в качестве своеобразных энергетических «квартир», которые заселяются микрочастицами. В настоящем разделе рассматривается вопрос о вероятностях заселения этих «квартир» в зависимости от энергии квантового состояния. Общим физическим принципом, используемым при анализе данного вопроса, является «*принцип минимальной энергии*», согласно которому считается, что ансамбль микрочастиц в целом выбирает из всей совокупности энергетических состояний такое состояние (или группу состояний), при котором его энергия будет иметь минимальное значение. Особенно наглядно этот принцип проявляется при температуре системы $T = 0$ К, при которой микрочастицы занимают только нижние (малоэнергетические) квантовые состояния.

С учетом специфики своего поведения в ансамбле все микрочастицы принято подразделять на два класса:

- *фермионы* – которые проявляют ярко выраженное стремление к «*уединению*», т.е. не могут находиться в данном квантовом состоянии, если оно уже занято другой микрочастицей (к ним относятся электроны, протоны и другие микрочастицы, обладающие полуцелым спином: $(\frac{h}{4\pi}), (\frac{3h}{4\pi}), \dots$);

- *бозоны* – которые проявляют ярко выраженное стремление к «*объединению*» своего положения по энергетической шкале, т.е. могут заселять конкретное квантовое состояние в неограниченном количестве (к ним относятся фотоны и другие микрочастицы, обладающие целочисленным спином: $0, (\frac{h}{2\pi}), (\frac{h}{\pi}), \dots$).

При рассмотрении вероятностных характеристик фактора заселенности энергетических состояний (при $T > 0$ К) большое значение имеет величина отношения количества микрочастиц N , имеющих в системе, к количеству квантовых состояний в системе G . С этой точки зрения ансамбли микрочастиц подразделяются на два типа:

- $\frac{N}{G} \ll 1$ – невырожденный ансамбль, в котором у микрочастиц

есть определенная возможность собственного «выбора» квантового состояния («квартиры»), а следовательно, *вероятность заселения конкретного квантового состояния становится достаточно малой величиной;*

- $\frac{N}{G} \approx 1$ (при этом предполагается, что $N \leq G$) – *вырожденный*

ансамбль, в котором у микрочастиц весьма ограничена возможность собственного «выбора» квантового состояния, а следовательно, *вероятность заселения конкретного квантового состояния приобретает значения, близкие к единице.*

Деление ансамблей микрочастиц на вырожденные и невырожденные случаи может содержать элемент некоторой неоднозначности. Например, ансамбль электронов в кристаллических структурах (металлы, полупроводники) является вырожденным при низких температурах и невырожденным при высоких температурах.

В соответствии с приведенной выше классификацией ансамблей микрочастиц проводится и классификация физических статистик (подробнее смотри в источнике [3]):

- статистика, отражающая поведение невырожденных ансамблей микрочастиц, называется *классической статистикой*, или *статистикой Максвелла-Больцмана*;

- статистика, отражающая поведение вырожденных ансамблей микрочастиц, называется *квантовой статистикой*, или *статистикой Ферми-Дирака*.

Следует отметить, что квантовая статистика применима только к квантовым системам, в то время как в классической статистике могут рассматриваться как классические, так и квантовые системы. Последнее обстоятельство находит свое наглядное проявление для случая ансамбля электронов, находящегося в кристаллической структуре (металлы, полупроводники), который, в зависимости от температуры, может менять степень своего вырождения (подробнее условие снятия вырождения рассмотрено ниже).

Введем в рассмотрение понятие о функции распределения частиц по энергии $f(W)$. Физический смысл указанной функции состоит в том, что произведение вида $[f(W) \cdot dW]$ равно вероятности заполнения микрочастицами квантовых состояний с энергиями, заключенными в интервале значений $W - (W + dW)$. Отметим, что речь идет о вероятности заполнения квантовых состояний, если они действительно содержатся в интервале энергий dW . Последнее наблюдается не всегда. Например, если интервал энергий dW расположен в запрещенной зоне собственного полупроводника, то определенная вероятность нахождения электронов в этом интервале энергий теоретически существует, а фактически электронов в этом интервале быть не может.

В классической (Максвелла-Больцмана) статистике функция распределения микрочастиц по энергиям $f_{MB}(W)$ имеет вид

$$f_{MB}(W) = e^{\frac{\mu}{kT}} \cdot e^{-\frac{W}{kT}}, \quad (2.42)$$

где k – постоянная Больцмана; μ – химический потенциал.

Физический смысл химического потенциала системы состоит в том, что он отражает энергию, которую микрочастица вносит (выносит) в систему при ее вхождении (выходе) в состав системы. Для случая невырожденного ансамбля микрочастиц химический потенциал равен

$$\mu = k \cdot T \cdot \ln \left[\frac{N}{V} \cdot \left(\frac{h^2}{2 \cdot \pi \cdot m \cdot k \cdot T} \right)^{\frac{3}{2}} \right]. \quad (2.43)$$

Подставив (2.43) в выражение (2.42), получим

$$f_{MB}(W) = \frac{N}{V} \cdot \left(\frac{h^2}{2 \cdot \pi \cdot m \cdot k \cdot T} \right)^{\frac{3}{2}} \cdot e^{-\frac{W}{kT}}. \quad (2.44)$$

Согласно выражению (2.44) функция распределения частиц по энергиям $f_{MB}(W)$ является монотонно убывающей функцией от энергии микрочастиц, имеющей максимальное значение $e^{\frac{\mu}{kT}}$ при значении энергии микрочастицы $W = 0$.

В квантовой (Ферми-Дирака) статистике функция распределения микрочастиц по энергиям $f_{\Phi Д}(W)$ имеет вид

$$f_{\Phi Д}(W) = \left(e^{\frac{W - \mu}{k \cdot T}} + 1 \right)^{-1}. \quad (2.45)$$

В выражении (2.45) символом μ , как и для рассмотренного выше случая, обозначен химический потенциал системы, который для вырожденного газа принято называть *энергией (или уровнем) Ферми* (W_F). Применительно к кристаллическим структурам металлического и полупроводникового типа энергия Ферми имеет вполне определенный физический смысл, а именно – *работы выхода электронов* из указанных материалов. Следует отметить, что энергия Ферми и, как следствие, работа выхода электронов в несобственных полупроводниках зависит от величины концентрации и типа (донорная, акцепторная) легирующей примеси. В частности, в полупроводниках донорного типа энергия Ферми меньше, чем в собственных полупроводниках, а в полупроводниках акцепторного типа – больше. Этот фактор положен в основу проектирования полупроводниковых приборов (структур диодного и транзисторного типа) и современных интегральных схем.

Характерный вид функции распределения частиц по энергиям $f_{\Phi Д}(W)$ легко просматривается при обращении к частному случаю: $T \approx 0$. Нетрудно видеть, что в этом случае выражение (2.45) можно заменить аналитическим представлением вида

$$f_{\Phi Д}(W) = \begin{cases} 1 & \text{при } W < \mu ; \\ 0 & \text{при } W > \mu ; \\ 1/2 & \text{при } W = \mu . \end{cases} \quad (2.45a)$$

Из выражения (2.45a) следует, что при выполнении условия $T \approx 0$ вероятность заполнения квантовых состояний равна единице при всех значениях энергии микрочастиц, не превосходящих энергию Ферми (случай полного вырождения энергетического состояния ансамбля микрочастиц).

Вид функций распределения частиц по энергиям $f_{MB}(W)$ и $f_{ФД}(W)$ представлен на рис. 2.12.

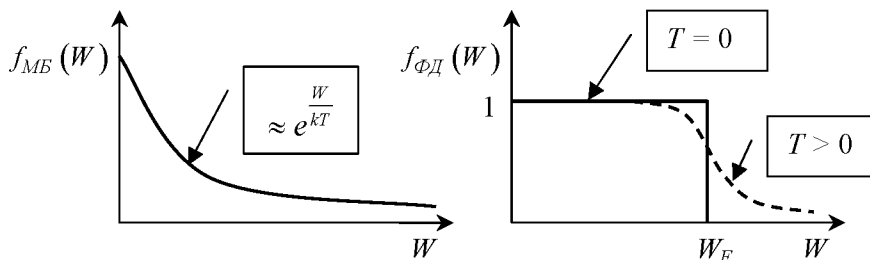


Рис. 2.12. Схематическое отображение функций распределения частиц по энергиям $f_{MB}(W)$ и $f_{ФД}(W)$

Проведем сопоставительный анализ функций распределения частиц по энергиям $f_{MB}(W)$ (невырожденный ансамбль микрочастиц) и $f_{ФД}(W)$ (вырожденный ансамбль микрочастиц). Нетрудно заметить, что при выполнении условия

$$e^{\frac{W-\mu}{kT}} \gg 1 \quad (2.46)$$

обе функции распределения частиц по энергиям $f_{MB}(W)$ и $f_{ФД}(W)$ приобретают тождественный вид. Отсюда следует, что выполнение условия $(W - \mu) \gg kT$ приводит к снятию вырождения по распределению микрочастиц по энергиям. Это обстоятельство представляет определенное удобство в практическом плане, поскольку из него следует, что при больших значениях энергий микрочастиц W можно использовать (ввиду устранения фактора вырожденности ансамбля микрочастиц) статистику Максвелла-Больцмана.

Относительно фактора вырождения ансамбля микрочастиц в кристаллических структурах можно сделать ряд общих заключений:

- в металлах, ввиду чрезвычайно высокой (порядка 10^{28} м^{-3}) концентрации свободных электронов, т.е. способных перемещаться на значительные расстояния, ансамбль электронов всегда находится в вырожденном состоянии и поэтому его поведение подчиняется квантовой статистике Ферми-Дирака;

- в собственных и слаболегированных полупроводниках концентрации свободных электронов значительно ниже ($10^{18} - 10^{24} \text{ м}^{-3}$ – в зависимости от содержания легирующей примеси донорного или акцепторного типа) относительно металлов, поэтому ансамбль электронов, как правило, находится в невырожденном состоянии и его поведение подчиняется классической статистике Максвелла-Больцмана.

Характерные моменты в различии поведения идеального невырожденного (статистика Максвелла-Больцмана) и вырожденного (статистика Ферми-Дирака) ансамблей микрочастиц представлены в табл. 2.7.

Т а б л и ц а 2.7. *Средняя кинетическая энергия и среднеквадратичная скорость движения в вырожденном и невырожденном ансамбле микрочастиц [4]*

Параметр		Тип ансамбля микрочастиц	
		невырожденный	вырожденный
W	$T = 0 \text{ К}$	0	$\left(\frac{3}{2}\right) \cdot W_F$
	$T > 0 \text{ К}$	$\frac{3}{2} \cdot k \cdot T$	$\frac{3}{5} \cdot W_F \cdot \left[1 + \frac{5 \cdot \pi^2}{12} \cdot \left(\frac{k \cdot T}{W_F} \right)^2 \right] \approx \frac{3}{5} \cdot W_F$
$v_{\text{кс}}$	$T = 0 \text{ К}$	0	$\approx \sqrt{\frac{W_F}{m}}$
	$T > 0 \text{ К}$	$\sqrt{\frac{3 \cdot k \cdot T}{m}}$	

Примечание: типичные значения энергии Ферми (W_F) приведены в приложении № 2 (см. пункт «Г»).

На основании данных, приведенных в табл. 2.7, можно сделать ряд выводов, представляющих практический интерес:

- в невырожденном ансамбле микрочастиц средняя энергия и среднеквадратичная скорость микрочастиц являются функциями температуры и при абсолютном нуле обращаются в нуль;
- в вырожденном ансамбле микрочастиц средняя энергия и среднеквадратичная скорость микрочастиц уже при абсолютном нуле имеют огромные значения (по сравнению с тепловой составляющей) и практически не зависят от температуры, что свидетельствует о нетермической природе их происхождения.

г) *Полная функция распределения ансамбля микрочастиц по энергиям*

В приведенных выше разделах рассмотрены функции плотности и вероятностей заполнения микрочастицами квантовых состояний, но ни в коей мере не рассмотрен вопрос о фактическом заселении микрочастицами индивидуальных энергетических состояний зонной диаграммы. Представляется очевидным, что количество микрочастиц dN , находящихся на энергетических уровнях, расположенных в пределах энергетического интервала dW , будет пропорционально:

- протяженности указанного энергетического интервала dW на зонной энергетической диаграмме (очевидно, что расположенного только в разрешенных зонах);
- значению функции плотности квантовых состояний $g(W)$ в указанном энергетическом интервале dW ;
- вероятности заполнения индивидуальных квантовых состояний $f_{MB}(W)$ или $f_{ФД}(W)$ (в зависимости от типа ансамбля микрочастиц: невырожденный или вырожденный):

$$dN = f_{MB(илиФД)}(W) \cdot g(W) \cdot dW . \quad (2.47)$$

Функция вида $N(W) = dN/dW$ получила название *полной функции распределения микрочастиц по энергиям*

$$N(W) = \frac{dN}{dW} = f_{MB(илиФД)}(W) \cdot g(W) \quad (2.48)$$

и характеризует фактическую плотность распределения микрочастиц по энергиям.

В классической (Максвелла-Больцмана) невырожденной статистике для полной функции распределения микрочастиц по энергиям $N_{MB}(W)$ на основании выражений (2.40) и (2.42), имеем

$$N_{MB}(W) = \frac{2 \cdot N}{\sqrt{\pi} \cdot (k \cdot T)^3} \cdot \sqrt{W} \cdot e^{-\frac{W}{k \cdot T}} . \quad (2.49)$$

Следует отметить, что на вид полной функции распределения микрочастиц по энергиям $N_{MB}(W)$ оказывают влияние два конкурирующих фактора:

- фактор возрастания $N_{MB}(W)$ с увеличением W , обусловленный проявлением влияния функции плотности квантовых состояний в системе $g(W)$;
- фактор убывания $N_{MB}(W)$ с увеличением W , обусловленный проявлением влияния функции вероятности заполнения квантовых состояний в системе.

Практическая значимость полной функции распределения микрочастиц по энергиям $N_{MB}(W)$ состоит в том, что с ее помощью можно определить количество микрочастиц $N(W_1, W_2)$, энергия которых находится в определенном энергетическом интервале ($W_1 - W_2$):

$$N(W_1, W_2) = \int_{W_1}^{W_2} N_{MB}(W) \cdot dW = \frac{2 \cdot N}{\sqrt{\pi} \cdot (k \cdot T)^3} \int_{W_1}^{W_2} \sqrt{W} \cdot e^{-\frac{W}{k \cdot T}} \cdot dW. \quad (2.50)$$

Графическое отображение полной функции распределения микрочастиц по энергиям $N_{MB}(W)$ представлено на рис. 2.13.

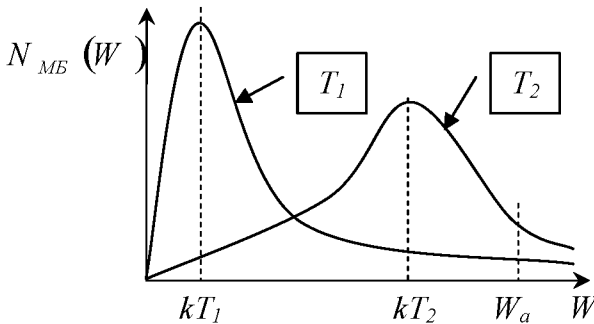


Рис. 2.13. Полная функция распределения микрочастиц по энергиям для невырожденного ансамбля частиц ($T_2 > T_1$)

Относительно некоторых фрагментов рис. 2.13 уместно привести ряд замечаний, имеющих практическую значимость:

- численное значение полной площади под кривой распределения по энергиям $N_{MB}(W)$ представляет собой общее количество частиц N_0 , входящих в состав рассматриваемого ансамбля микрочастиц (это следует из того очевидного факта, что интеграл (2.50) при значениях $W_1 = 0$ и $W_2 = \infty$ охватывает всю совокупность микрочастиц в ансамбле);

- положение максимума полной функции распределения по энергиям $N_{MB}(W)$ по шкале энергий соответствует значению средней энергии микрочастиц в ансамбле, равной по порядку величины $\approx kT$;

- повышение температуры ансамбля микрочастиц (например, от температуры T_1 до температуры T_2) приводит к сдвигу максимума полной функции распределения микрочастиц $N_{MB}(W)$ в направлении больших значений энергии (при сохранении общей площади поверхности под кривой распределения $N_{MB}(W)$);

- с повышением температуры полная функция распределения микрочастиц по энергиям $N_{MB}(W)$ приобретает все более «размытый» характер, из чего следует, что микрочастицы в этом случае имеют более высокую степень свободы при выборе своего энергетического состояния;

- площадь поверхности, расположенной правее значения энергии W_a (энергия активации) и простирающейся вплоть до значений $W \rightarrow \infty$, численно равна количеству микрочастиц, способных реализовать (по своим энергетическим возможностям) определенный элементарный физико-химический процесс.

В прикладных задачах довольно часто встречаются случаи, когда энергия активации элементарных физико-химических процессов W_a существенно превышает среднюю тепловую энергию микрочастиц в ансамбле – kT . Это следует из того, что на практике параметр W_a , как правило, соизмерим со значением 1 эВ, в то время как $kT = 0,025$ эВ (при $T = 300^\circ \text{C}$). Такой уровень расхождения ($W_a / kT = 40$) позволяет считать определяющим фактором функцию $f_{MB}(W)$ (полагая при

этом, что функция $g(W) = \text{const}$, что способствует существенному упрощению выражения (2.49). В этом случае для проведения оценочных расчетов концентрации активных (в энергетическом плане) микрочастиц вместо выражения (2.50) можно использовать более простое аналитическое выражение:

$$N_{MB}(W_1, W_2) = C \cdot \int_{W_1}^{W_2} e^{-\frac{W}{k \cdot T}} \cdot dW, \quad (2.51)$$

где C – постоянная величина.

Наибольший практический интерес представляет случай, когда $W_1 = W_a$, а $W_2 \rightarrow \infty$. Для этого случая на основании выражения (2.51), можно определить концентрацию активных микрочастиц $N_{MB}(W > W_a)$, энергия которых больше W_a , а следовательно, эти микрочастицы способны реализовать рассматриваемый элементарный физико-химический процесс посредством механизма термической активации:

$$N_{MB}(W > W_a) = N_0 \cdot e^{-\frac{W_a}{k \cdot T}}, \quad (2.52)$$

где N_0 – полная концентрация микрочастиц в ансамбле.

Некоторое представление о масштабе величин параметра $N_{MB}(W > W_a)$ можно составить на основании данных, приведенных в табл. 2.8.

Таблица 2.8. *Зависимость относительной концентрации активных микрочастиц от энергии активации W_a и температуры T*

$\frac{N_{MB}(W \geq W_a)}{N_0}$	W_a эВ	Температура, К				
		50	100	300	600	900
0.5		$\approx 10^{-51}$	$9.36 \cdot 10^{-14}$	$2.06 \cdot 10^{-9}$	$4.54 \cdot 10^{-5}$	$1.27 \cdot 10^{-3}$
1.0		0	$\approx 10^{-51}$	$\approx 10^{-17}$	$2.06 \cdot 10^{-9}$	$1.62 \cdot 10^{-6}$
1.5		0	0	$\approx 10^{-26}$	$9.36 \cdot 10^{-14}$	$2.06 \cdot 10^{-9}$

В квантовой (Ферми-Дирака) вырожденной статистике для полной функции распределения электронов (с учетом их спиновых состояний) по энергиям $N_{ФД}(W)$ на основании выражений (2.40) и (2.45) имеем

$$N_{\text{ФД}}(W) = \frac{4 \cdot \pi \cdot V}{h^3} \cdot (2 \cdot m)^{3/2} \cdot \sqrt{W} \cdot \left(e^{\frac{W - W_F}{k \cdot T}} + 1 \right)^{-1}. \quad (2.53)$$

При очень низких температурах ($T \approx 0$) вероятность заполнения квантовых состояний практически равна единице для совокупности малознергетических уровней и в этой связи выражение (2.53) приобретает более простой вид:

$$N_{\text{ФД}}(W) = \frac{4 \cdot \pi \cdot V}{h^3} \cdot (2 \cdot m)^{3/2} \cdot \sqrt{W}. \quad (2.54)$$

Интегрируя выражение (2.54) в энергетическом интервале $\{0 - W_F\}$, определяем общее количество электронов, имеющих энергию в указанном энергетическом интервале:

$$N_{\text{ФД}}(0, W_F) = \frac{8 \cdot \pi \cdot V}{3 \cdot h^3} \cdot (2 \cdot m)^{3/2} \cdot W_F. \quad (2.55)$$

Из выражения (2.55) можно определить энергию Ферми W_F :

$$W_F = \frac{\hbar^2}{2 \cdot m} \cdot \left(\frac{3 \cdot n}{8 \cdot \pi} \right)^{2/3}, \quad (2.56)$$

где $n = N(0, W_F)/V$ – концентрация электронного газа в системе.

Любопытно сравнить родственные понятия химического потенциала μ (невырожденная статистика, см. выражение (2.43)) и энергии Ферми W_F (вырожденная статистика, см. выражение (2.56)). Оба эти понятия имеют одинаковый физический смысл, а именно – *это энергия, которую микрочастица вносит (выносит) в систему при ее вхождении (выходе) в состав системы*. Однако в первом случае параметр μ является функцией температуры, а во втором – параметр W_F не является функцией температуры. Последнее обстоятельство следует из того, что энергия Ферми, характеризующая поступательное движение электрона в системе, не является энергией теплового движения и имеет исключительно квантовую природу.

Зная значение энергии Ферми W_F , можно определить *максимальную скорость поступательного движения* электронов v_F , находящихся на верхнем энергетическом уровне в разрешенной зоне:

$$v_F = \sqrt{\frac{2 \cdot W_F}{m}}, \quad (2.57)$$

и температуру Ферми T_F (условная величина, которая показывает, при какой температуре невырожденный газ имел бы энергию теплового движения, равную энергии Ферми):

$$T_F = \frac{W_F}{k}. \quad (2.58)$$

В табл. 2.9 представлены значения энергии Ферми W_F , температуры Ферми T_F и скорости поступательного движения v_F для некоторых металлов.

Таблица 2.9. *Квантово-механические характеристики некоторых металлов [4]*

Металл	W_F , эВ	$T_F \cdot 10^4$	$v_F \cdot 10^6$ мс ⁻¹ ; (v_F/v_T)
Li	4.7	5.5	1.3 (13.7)
Na	3.1	3.7	1.1 (11.1)
Cu	7.1	8.2	1.6 (16.9)
Ag	5.5	6.4	1.4 (14.8)

Примечание: Значения v_T соответствуют температуре $T = 300$ К.

Из данных табл. 2.9 следует, что при температуре $T = 300$ К скорость поступательного движения микрочастицы v_F превосходит более чем на порядок скорость теплового движения v_T . Очевидно, что величина этого расхождения между скоростями v_F и v_T будет возрастать с уменьшением температуры.

д) Концентрация носителей заряда и электропроводность в кристаллических структурах металлического и полупроводникового типа

Анализ вопроса о концентрации носителей заряда в кристаллических структурах с электронным типом проводимости проводится на основе вышеизложенных материалов (см. пункты настоящего раздела: а, б, в, г). С методической точки зрения целесообразно начать рассмотрение этого вопроса с материалов металлического типа.

В материалах металлического типа зонная энергетическая диаграмма не имеет запрещенной зоны, а следовательно, в них зона проводимости непосредственно примыкает к валентной зоне. При температуре $T = 0$ К зона проводимости металла полностью свободна от электронов, а валентная зона, напротив, полностью заполнена. С повышением температуры металла ($T > 0$ К) часть электронов, находящихся в *верхней части* валентной зоны, подвергаются тепловому возбуждению (т.е. электроны увеличивают свою энергию) и под действием этого фактора переходят на более высоко расположенные энергетические уровни, находящиеся в *нижней части* зоны проводимости. Поскольку тепловому возбуждению подвергаются не все электроны валентной зоны, а только часть из них (которые и являются носителями заряда при реализации процесса электропроводимости), то вопрос об их количестве приобретает исключительно практический интерес.

Энергетическому воздействию подвергаются электроны, находящиеся на энергетических уровнях валентной зоны в интервале энергий протяженностью kT , непосредственно примыкающем к уровню Ферми. Электроны, находящиеся на более глубоких энергетических уровнях валентной зоны, практически не подвергаются тепловому возбуждению. Проведем оценку количества этих возбужденных (т.е. перешедших в зону проводимости) электронов. Примем, что в валентной зоне изначально находится N электронов, расположенных на $N/2$ энергетических уровнях (на каждом уровне может находиться, согласно принципу Паули, не более двух электронов). Протяженность валентной зоны определяется интервалом энергий $\{0 - W_F\}$. В целях упрощения примем допущение, что энергетические уровни расположены в валентной зоне на одинаковом расстоянии друг от друга, равном ΔW_i :

$$\Delta W_i = \frac{W_F}{(N/2)} = \frac{2 \cdot W_F}{N} \approx 10^{-23} \text{ эВ.} \quad (2.59)$$

(Столь малое значение интервала ΔW_i позволяет считать, что энергетические зоны являются не дискретными, а непрерывными.)

Как отмечалось выше, тепловому возбуждению подвергаются электроны в энергетическом интервале с протяженностью kT , расположен-

ном ниже уровня Ферми W_F . В этом энергетическом интервале валентной зоны изначально находится количество энергетических уровней ΔN_{kT} , равное

$$\Delta N_{kT} = \frac{k \cdot T}{\Delta W_i} = \frac{k \cdot T \cdot N}{2 \cdot W_F}. \quad (2.60)$$

На этих ΔN_{kT} энергетических уровнях валентной зоны располагаются ΔN_0 термически возбужденных электронов, количество которых можно оценить с помощью выражения (принимая во внимание принцип Паули):

$$\Delta N_0 = \frac{k \cdot T \cdot N}{W_F}. \quad (2.61)$$

Термически возбужденные электроны валентной зоны переходят на нижние уровни зоны проводимости, которые по определению расположены выше уровня Ферми. Эти перешедшие электроны находятся в состоянии *динамического равновесия* с оставшимися в валентной зоне дырками. Последнее означает, что термически возбужденные электроны, перешедшие в зону проводимости, могут потерять свою избыточную (тепловую) энергию путем ее передачи атомам кристаллической решетки и вернуться в валентную зону. В этой связи их равновесное количество будет меньше общего количества возбужденных электронов в валентной зоне ΔN_0 . Полагая, что за уровень Ферми переходит не более половины возбужденных электронов ΔN_0 , для оценки количества носителей заряда ΔN в проводниках металлического типа получим

$$\Delta N \approx \frac{k \cdot T \cdot N}{2 \cdot W_F}. \quad (2.62)$$

В табл. 2.10 представлена зависимость относительной концентрации носителей заряда $\Delta N/N$ в металлах от температуры, полученная на основании выражения (2.62).

Т а б л и ц а 2.10. *Зависимость $\Delta N / N$ от температуры в проводниках металлического типа (при значении энергии Ферми $W_F = 5$ эВ)*

T, K	10	50	100	150	200	250
$\Delta N / N$	$8 \cdot 10^{-5}$	$4 \cdot 10^{-4}$	$8 \cdot 10^{-4}$	$1.2 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	$2 \cdot 10^{-3}$
T, K	300	350	400	450	500	600
$\Delta N / N$	$2.4 \cdot 10^{-3}$	$2.8 \cdot 10^{-3}$	$3.2 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$	$4 \cdot 10^{-3}$	$4.8 \cdot 10^{-3}$
T, K	800	1000	1200	1400	1600	2000
$\Delta N / N$	$6.4 \cdot 10^{-3}$	$8 \cdot 10^{-3}$	$9.6 \cdot 10^{-3}$	$1.1 \cdot 10^{-2}$	$1.3 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$

Согласно данным табл. 2.10 можно сформулировать ряд важных, с практической точки зрения, выводов:

- вплоть до температур 2000 К, относительное количество электронов, перешедших выше уровня Ферми, не превышает 2%, что позволяет считать электронный газ в металле вырожденным и принять точку зрения, что его распределение мало отличается от распределения при абсолютном нуле;

- в рабочем диапазоне функционирования технических устройств (250 – 400 К) относительное количество свободных электронов (т.е. находящихся выше уровня Ферми) составляет по порядку величины 0,1 %;

- несмотря на малые величины относительного количества свободных электронов $\Delta N / N$, абсолютные значения концентрации носителей заряда в металлах могут достигать весьма больших значений (порядка $10^{20} - 10^{21} \text{ см}^{-3}$).

С учетом вышеизложенного можно считать, что концентрация носителей заряда в проводниках металлического типа весьма слабо зависит от температуры, а температурная зависимость электропроводности этих конструкционных материалов всецело определяется вариацией подвижности носителей заряда (см. выражение (2.1)).

В собственных полупроводниках зонная энергетическая диаграмма имеет запрещенную зону, а следовательно, у них нижний уровень зоны проводимости отстоит от верхнего уровня валентной зоны на расстоянии, равном ширине запрещенной зоны $\Delta W_{зз}$. При температуре $T = 0$ К зона проводимости полупроводника полностью свободна от электронов

(как и у металла), а валентная зона, напротив, полностью заполнена. С повышением температуры полупроводника ($T > 0$ К) часть электронов, находящихся в *верхней части* валентной зоны, подвергаются тепловому возбуждению и под действием этого фактора переходят на более высоко расположенные энергетические уровни, находящиеся в *нижней части* зоны проводимости. Однако в случае собственных полупроводников эти переходы связаны с необходимостью преодоления энергетического барьера ΔW_{33} , поэтому *количество перешедших электронов будет гораздо меньше*, чем в случае металлов (т.е. при отсутствии энергетического интервала ΔW_{33}).

В результате перехода электронов из валентной зоны в зону проводимости полупроводника на зонной диаграмме последнего образуется *две частично заполненные энергетические зоны* (в определенной мере аналогичные рассмотренному выше случаю металлов). В обеих этих частично заполненных энергетических зонах носителями заряда являются электроны, однако ввиду разной подвижности носителей заряда (в зоне проводимости она больше) механизмы проводимости в этих зонах получили различное название: электронный (в зоне проводимости) и дырочный (в валентной зоне). Последним отмечается тот факт, что концентрация носителей заряда соизмерима с концентрацией пустых мест (дырок) на энергетических уровнях валентной зоны. Представляется очевидным, что в собственных полупроводниках концентрации электронов в зоне проводимости n_i и дырок в валентной зоне p_i одинаковы (с учетом фактора их происхождения):

$$n_i = p_i, \quad (2.63)$$

где i – индекс, обозначающий, что рассматривается случай собственных полупроводников.

В собственных полупроводниках для энергетических уровней зоны проводимости, как правило, выполняется условие $(W - \mu) \gg kT$ ($\mu = W_F$), что приводит к снятию вырождения по распределению микрочастиц по энергиям (см. выражение (2.46)). Таким образом, электронный газ, находящийся в зоне проводимости, можно считать невырожденным и в этой связи использовать (ввиду устранения фактора

вырожденности ансамбля микрочастиц) статистику Максвелла-Больцмана. С учетом этого обстоятельства для концентрации электронов n_i (в зоне проводимости) и дырок p_i (в валентной зоне) имеем [4]

$$n_i = p_i = 2 \cdot \left(\frac{2 \cdot \pi \cdot k \cdot T \cdot \sqrt{m_n \cdot m_p}}{h^3} \right)^{3/2} \cdot e^{-\frac{W_i}{2 \cdot k \cdot T}}, \quad (2.64)$$

где m_n и m_p – эффективная масса соответственно для электрона (в зоне проводимости) и дырки (в валентной зоне); W_i – ширина запрещенной зоны собственного полупроводника.

При выводе выражения (2.64) использовано то обстоятельство, что температурное смещение уровня Ферми столь незначительно, что им можно пренебречь и считать $W_F \approx -(W_i/2)$. Для большинства собственных полупроводников это допущение является правомерным. Однако следует иметь в виду, что для полупроводников с малой шириной запрещенной зоны этим смещением нельзя пренебрегать уже при комнатной температуре (например, у полупроводника типа InSb, у которого $W_i = 0.17$ эВ).

Поскольку в собственном полупроводнике имеется два типа носителей заряда (электроны и дырки), то выражение для электропроводности (2.1) этого типа конструкционных материалов приобретает вид

$$\sigma_i = q \cdot n_i \cdot \mu_n + q \cdot n_p \cdot \mu_p = q \cdot n_i \cdot (\mu_n + \mu_p), \quad (2.65)$$

где q – заряд электрона (дырки); μ_n и μ_p – подвижность соответственно для электрона (в зоне проводимости) и дырки (в валентной зоне).

С учетом выражений (2.64) и (2.65) в первом приближении можно считать, что температурная зависимость электропроводности собственных полупроводников определяется выражением

$$\sigma_i = \sigma_0 \cdot e^{-\frac{W_i}{2 \cdot k \cdot T}}, \quad (2.66)$$

где σ_0 – постоянная, равная электропроводности собственного полупроводника при температуре $T \rightarrow \infty$.

Температурную зависимость электропроводности собственных полупроводников удобно представлять в полулогарифмических координатах вида $\lg \sigma_i = \varphi(1/T)$:

$$\ln \sigma_i = \ln \sigma_0 - \left(\frac{W_i}{2 \cdot k} \right) \cdot \frac{1}{T}. \quad (2.67)$$

Типичный вид температурной зависимости собственных полупроводников представлен на рис. 2.14.

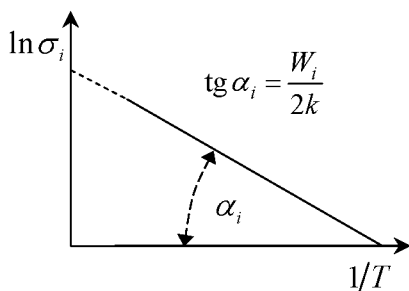


Рис.2.14. Температурная зависимость собственных полупроводников

С помощью экспериментальной зависимости, отображенной на рис. 2.14, можно определить основные характеристические параметры электропроводности собственных полупроводников W_i и σ_0 .

Из выражения (2.66) следует, что концентрация носителей заряда (выражение (2.64)), а следовательно, и электропроводность (выражение (2.67)) собственных полупроводников должны существенно зависеть от ширины запрещенной зоны полупроводника W_i и температуры T (табл. 2.11).

Т а б л и ц а 2.11. *Зависимость концентрации носителей заряда от параметров W_i и T*

W_i , эВ	Концентрация носителей заряда n_i , см ⁻³		
	0.5	1.0	1.5
$T = 100$ К	$8.7 \cdot 10^{-4}$	0	0
$T = 300$ К	$2.1 \cdot 10^{14}$	$4.2 \cdot 10^5$	$8.7 \cdot 10^{-4}$
$T = 600$ К	$4.5 \cdot 10^{18}$	$2.1 \cdot 10^{14}$	$9.4 \cdot 10^9$

Данные табл. 2.11 позволяют оценить масштабный фактор в изменении концентрации носителей заряда (и электропроводности) при вариациях ширины запрещенной зоны собственного полупроводника (W_i) и его температуры (T). В частности, следует отметить тот немаловажный факт, что при изменении ширины запрещенной зоны собственного полупроводника или его температуры всего в два раза концентрация носителей заряда может измениться на несколько порядков.

В примесных полупроводниках донорного типа зонная энергетическая диаграмма имеет в запрещенной зоне дополнительную узкую донорную зону, изначально заполненную электронами. Протяженность этой дополнительной донорной зоны ΔW_δ можно оценить согласно выражению

$$\Delta W_\delta \approx N_\delta \cdot \Delta W_1, \quad (2.68)$$

где N_δ – концентрация донорной примеси в собственном полупроводнике; ΔW_1 – среднее расстояние между соседними энергетическими уровнями в пределах зоны. Полагая $N_\delta = 10^{17} \text{ см}^{-3}$ (достаточно высокий уровень легирования) и $\Delta W_1 = 10^{-23} \text{ эВ}$, получаем $\Delta W_\delta = 10^{-5} \text{ эВ}$. Отсюда следует, что протяженность ΔW_δ в 10^5 раз меньше, чем протяженность зоны проводимости или валентной зоны, которые равны по порядку величины 1 эВ. Ввиду незначительной протяженности донорной зоны по энергетической шкале ее принято называть *донорным уровнем*. Другой характерной и важной для практических приложений особенностью донорного уровня является то обстоятельство, что он, как правило, располагается *в непосредственной близости от дна зоны проводимости* (на расстоянии порядка 0.05 – 0.1 эВ).

Дальнейшее рассмотрение температурной зависимости электропроводности донорных полупроводников проводится на основе схематического отображения зонной энергетической диаграммы, представленной на рис. 2.15.

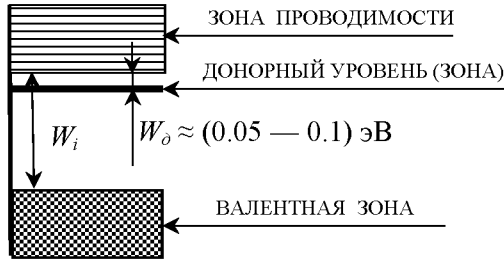


Рис.2.15. Энергетическая диаграмма донорного полупроводника

В области низких температур, т.е. при выполнении условия $W_i > kT$, переходами электронов из валентной зоны в зону проводимости можно пренебречь (ввиду того, что $W_i \approx 1 \text{ эВ} \gg W_d$) и считать, что все электроны в зоне проводимости появились там благодаря переходам электронов с донорного уровня. В этой связи в первом приближении можно принять, что равновесная концентрация электронов в зоне проводимости, попавших туда с донорного уровня, определяется выражением (сопоставить с выражением (2.64))

$$n_d = N_d \cdot e^{-\frac{W_d}{2 \cdot k \cdot T}} \quad (2.69)$$

В более корректной постановке равновесная концентрация электронов в зоне проводимости определяется более сложным выражением [5]

$$n_d = (N_d \cdot N_c)^{1/2} \cdot e^{-\frac{W_d}{2 \cdot k \cdot T}} = \sqrt{2 \cdot N_d} \cdot \left(\frac{2 \cdot \pi \cdot m_n \cdot k \cdot T}{h^2} \right)^{3/4} \cdot e^{-\frac{W_d}{2 \cdot k \cdot T}}, \quad (2.69a)$$

где N_c – эффективное число энергетических состояний в зоне проводимости собственного полупроводника.

В табл. 2.12 представлена зависимость равновесной относительной концентрации электронов в зоне проводимости n_d / N_d от фактора $W_d / (kT)$, полученная на основании выражения (2.69).

Т а б л и ц а 2.12. *Зависимость n_{δ} / N_{δ} от параметра $W_{\delta} / (kT)$*

$W_{\delta} / (kT)$	15	10	5	4	3	2	1
n_{δ} / N_{δ}	$5.5 \cdot 10^{-4}$	$6.7 \cdot 10^{-3}$	0.082	0.135	0.223	0.368	0.607

Из данных табл. 2.12 следует, что даже при относительно малых значениях температуры в зону проводимости полупроводника переходит с донорного уровня значительное количество электронов от числа имеющихся на данном уровне.

В области высоких температур переходами электронов из валентной зоны в зону проводимости в общем случае пренебрегать нельзя. Действительно, по мере повышения температуры подавляющее большинство электронов донорного уровня переходит в зону проводимости и наблюдается эффект *истощения примесного уровня*, с точки зрения его роли в качестве поставщика электронов. В этом случае концентрация электронов в зоне проводимости, обусловленная наличием донорной примеси, достигнет своего предельного значения:

$$n_{\delta} = N_{\delta}. \quad (2.70)$$

При более высоких температурах начинается процесс перехода электронов из валентной зоны в зону проводимости. Поскольку концентрация электронов в валентной зоне гораздо выше концентрации донорной примеси, вклад этих электронов в общую концентрацию носителей заряда полупроводника может быть существенным. С учетом этого дополнительного обстоятельства для концентрации носителей заряда в полупроводниках донорного типа имеем выражение вида

$$n = N_{\delta} + n_i = N_{\delta} + 2 \cdot \left(\frac{2 \cdot \pi \cdot k \cdot T \cdot \sqrt{m_n \cdot m_p}}{h^3} \right)^{3/2} \cdot e^{-\frac{W_i}{2 \cdot k \cdot T}}. \quad (2.71)$$

С учетом всех вышеизложенных моментов можно построить обобщенную температурную зависимость концентрации носителей заряда в полупроводниках донорного типа в виде, представленном на рис. 2.16.

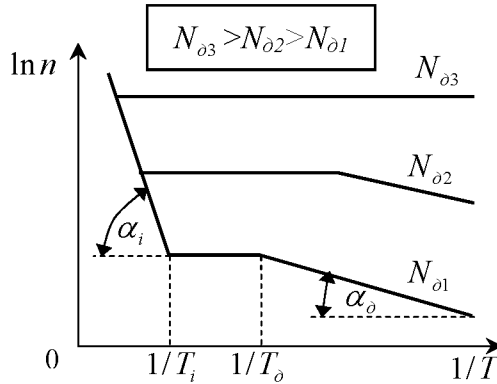


Рис. 2.16. Температурная зависимость концентрации носителей заряда в полупроводниках донорного типа

Для пояснения температурной зависимости концентрации носителей заряда в донорном полупроводнике уместно привести следующие комментарии:

- область, расположенная правее точки $1/T_o$ по оси абсцисс, характеризует вклад в концентрацию носителей заряда от донорного уровня (см. зависимость с концентрацией донорной примеси $N_{\delta 1}$);
- область, расположенная левее точки $1/T_i$ по оси абсцисс, характеризует вклад в концентрацию носителей заряда от валентной зоны полупроводника (см. зависимость с концентрацией донорной примеси $N_{\delta 1}$);
- горизонтальная прямая (см. зависимость с концентрацией донорной примеси $N_{\delta 1}$), расположенная между точками $1/T_i$ и $1/T_o$ по оси абсцисс, свидетельствует о наличии фактора истощения донорной примеси и недостаточном уровне нагрева полупроводника (для осуществления энергетической активизации процесса перехода электронов из валентной зоны в зону проводимости);
- углы наклона α_i и α_o прямолинейных участков зависимости характеризуют соответственно энергетические параметры W_i и W_o , определяемые согласно выражениям:

$$W_i = 2 \cdot k \cdot \operatorname{tg} \alpha_i, \quad (2.72a)$$

$$W_o = 2 \cdot k \cdot \operatorname{tg} \alpha_o; \quad (2.72b)$$

- температура перехода T_i механизма электропроводности от примесной к собственной проводимости тем выше, чем шире запрещенная зона полупроводника W_i и больше концентрация легирующей примеси N_o в собственном полупроводнике.

В примесных полупроводниках акцепторного типа зонная энергетическая диаграмма имеет в запрещенной зоне дополнительную узкую акцепторную зону (ее также принято называть «уровнем» – ввиду малой протяженности по энергетической шкале), изначально пустую, т.е. в ней отсутствуют электроны. Характерной и важной для практических приложений особенностью акцепторного уровня является то обстоятельство, что он, как правило, располагается в непосредственной близости от потолка валентной зоны (на расстоянии порядка 0.05 – 0.1 эВ).

Относительно этого случая легирования собственных полупроводников можно повторить все выкладки, изложенные для донорных полупроводников, с тем принципиальным отличием, что ситуация частично занятой энергетической зоны образуется в валентной зоне, в результате перехода части электронов валентной зоны на акцепторный уровень.

В области низких температур концентрация носителей заряда (дырок) p в валентной зоне полупроводника в первом приближении равна

$$p = N_a \cdot e^{-\frac{W_a}{kT}}, \quad (2.73)$$

где W_a – расстояние от потолка валентной зоны до акцепторного уровня; N_a – концентрация акцепторной примеси в полупроводнике.

В области высоких температур концентрация носителей заряда (дырок) p в валентной зоне полупроводника равна

$$p = N_a + n_i = N_a + 2 \cdot \left(\frac{2 \cdot \pi \cdot k \cdot T \cdot \sqrt{m_n \cdot m_p}}{h^3} \right)^{3/2} \cdot e^{-\frac{w_i}{2 \cdot k \cdot T}}. \quad (2.74)$$

Температурная зависимость концентрации носителей заряда (дырок) p в валентной зоне полупроводника отображается графиком, аналогичным графику, представленному на рис. 2.16 (с заменой соответствующих индексов: « n » на « p » и « d » на « a »). Более подробное рассмотрение вопросов электропроводности полупроводниковых материалов донорного и акцепторного типов в статических и динамических электрических полях представлено в литературных источниках [4,5].

2.2. Элементы теории электропроводности наноразмерных структур

2.2.1. Общие ограничения применимости статистических методов анализа для случая наноразмерных структур

а) Статистические флуктуации характеристических параметров макро- и микросистем

В вышеизложенных материалах широко использовались статистические методы анализа, свойственные макроскопическим системам, содержащим *чрезвычайно большое количество* микрочастиц. Из общих принципов статистической физики с неизбежностью вытекает возможность проявления в системе флуктуаций, в результате которых система может самопроизвольно перейти из наиболее вероятного равновесного состояния (с *максимальной энтропией*) в одно из менее вероятных состояний. При этом ход этого флуктуационного процесса будет обратным тому, при котором происходит возрастание энтропии системы.

На практике явление флуктуации параметров системы может наблюдаться в явном виде в двух полярных случаях:

- когда *размеры системы очень малы* (фактор, характерный для рассматриваемых наноразмерных структур); в этом случае явление флуктуации параметров будет наблюдаться во времени часто и масштаб флуктуации будет относительно велик;

- *размеры системы очень велики* (макроскопические системы), но фиксируются чрезвычайно малые вариации параметров системы (например: шумовая составляющая информационных сигналов в современных электронных устройствах).

Пусть имеется *замкнутая система*, находящаяся в состоянии статистического равновесия и имеющая в этом состоянии равновесную энтропию S_0 . Система считается замкнутой, если она не обменивается с окружающей средой энергией и веществом (в более широкой трактовке сюда же можно отнести и обмен информацией). Будем считать, что текущее состояние системы может быть охарактеризовано с помощью некоторого *внутреннего параметра* δ , который в состоянии термодинамического равновесия принимает значение δ_0 . В этом случае энтропия системы будет некоторой функцией параметра δ и можно принять $S = S(\delta)$. Вероятность того, что рассматриваемая замкнутая система находится в состоянии, характеризуемом значением параметра δ , находящимся в интервале значений $\{\delta - (\delta + \Delta\delta)\}$, можно определить на основании формулы Больцмана [6]:

$$d\omega = C \cdot e^{\frac{S(\delta) - S(\delta_0)}{k \cdot T}} \cdot d\delta, \quad (2.75)$$

где C – постоянная величина, удовлетворяющая условию нормировки вероятностей состояния системы: $\oint d\omega = 1$. Поскольку представление о полной замкнутости системы является весьма сильной абстракцией реальных систем, этот случай представляет довольно ограниченный практический интерес.

Более реалистичным представляется случай квазизамкнутой системы, являющейся некоторой малой подсистемой, погруженной в большой термостат с постоянной температурой. Принято считать, что флуктуации происходят только в рассматриваемой подсистеме (учитывая ее малый объем относительно термостата) и не затрагивают сам термостат. Ввиду имеющего место обмена энергией и микрочастицами между подсистемой и термостатом состояние подсистемы будет функ-

цией некоторого внешнего макроскопического параметра ζ . Предполагается, что изменение (флуктуация) макроскопического параметра ζ происходит столь медленно, что подсистема постоянно находится в состоянии текущего статистического равновесия. Поскольку подсистема вместе с термостатом составляют (по определению) рассмотренную выше замкнутую систему, то к ним в целом применимо выражение (2.75). Однако в этом случае суммарное изменение энтропии рассматриваемой объединенной системы можно представить в виде

$$\Delta S = \Delta S_0 + \Delta S', \quad (2.76)$$

где ΔS_0 и $\Delta S'$ – изменение термодинамической энтропии соответственно у термостата и подсистемы.

Опуская нетривиальные математические преобразования, для статистического распределения флуктуации макроскопического параметра ζ имеем [6]

$$d\omega = \frac{1}{\sqrt{2 \cdot \pi \cdot \Delta^2}} \cdot e^{-\frac{(\zeta - \zeta_0)^2}{2 \cdot \Delta^2}} \cdot d\zeta, \quad (2.77)$$

где Δ^2 – величина среднеквадратического отклонения внешнего параметра ζ , оказывающего влияние на состояние подсистемы и определяемого согласно выражению

$$\Delta^2 = \frac{\int_{-\infty}^{+\infty} (\zeta - \zeta_0)^2 \cdot e^{-\frac{W(\zeta_0)(\zeta - \zeta_0)^2}{2 \cdot k \cdot T}} \cdot d\zeta}{\int_{-\infty}^{+\infty} e^{-\frac{W(\zeta_0)(\zeta - \zeta_0)^2}{2 \cdot k \cdot T}} \cdot d\zeta}, \quad (2.78)$$

где $W(\zeta_0)$ – потенциальная энергия при перемещении системы в некотором воображаемом (а в ряде случаев реальном) поле сил.

Из выражения (2.77) следует, что вероятность флуктуации резко уменьшается с ростом ее абсолютной величины и *уменьшением параметра* Δ^2 . Последняя величина пропорциональна концентрации мик-

рочастиц и абсолютной температуре объединенной системы «термостат – подсистема». В этой связи можно ожидать, что интенсивность флуктуационных процессов будет уменьшаться с уменьшением температуры системы. Это важное обстоятельство составляет физическую основу большинства практических методов подавления шумовой составляющей информационных сигналов в современных устройствах электронной техники.

Важность параметра Δ^2 требует более пристального внимания к нему с позиции ожидаемых флуктуаций характеристических (в макроскопических системах – термодинамических, таких как концентрация, давление, химический потенциал и т.д.) параметров наноразмерных структур. Действительно, в этих структурах, физические объемы которых чрезвычайно малы, содержится весьма *малое количество атомов*, реализующих их функциональное назначение. Поэтому незначительные изменения в наноразмерных структурах количества и типа микрочастиц или их местоположения, происшедшие в результате проявления факторов флуктуационной природы, приводит к значительному нарушению исходного атомно-молекулярного порядка, необходимого для обеспечения требуемого функционального назначения (качества) изделия.

Здесь уместно отметить, что ключевой особенностью статистических закономерностей является их вероятностный характер, позволяющий предсказать степень ожидания (вероятность) наступления определенного события из совокупности возможных событий, составляющих полную группу. В этой связи достоверность вероятностных прогнозов будет тем выше, чем большее количество микрочастиц составляет статистический ансамбль. Для системы, состоящей из N микрочастиц, относительное отклонение наблюдаемого физического параметра A , являющегося функцией N , определяется выражением [6]

$$\frac{\Delta A}{A_0} = \frac{1}{\sqrt{N}}, \quad (2.79)$$

где ΔA – среднеквадратическое отклонение от математического ожидания A_0 .

В табл. 2.13 приведена зависимость $\Delta A/A_0 = \varphi(N)$ от числа индивидуальных атомов, входящих в состав физической системы. Эта зависимость представляет практический интерес с точки зрения возможности использования и установления границ применимости статистических методов при проведении анализа наноразмерных структур.

Согласно данным табл. 2.13 видно, что с увеличением числа микро-частиц физические параметры испытывают все меньшие относительные отклонения, и таким образом при очень больших значениях N эти параметры могут считаться постоянными величинами во времени.

Т а б л и ц а 2.13. Зависимость $\Delta A/A_0 = \varphi(N)$

Параметр	Физический объем системы, нм ³ (по порядку величины)							
	0.03	0.09	1.0	10	100	10 ³	10 ⁴	10 ⁹ (1 мкм ³)
N	10	10 ²	10 ³	10 ⁴	10 ⁵	10 ⁶	10 ⁹	10 ¹²
$\Delta A/A_0$	0.32	0.10	0.03	0.01	0.003	0.001	3·10 ⁻⁵	10 ⁻⁶

К примеру, оценка погрешности (обусловленной факторами флуктуационной природы) при определении концентрации газовых молекул в стандартной атмосфере (при нормальных условиях $N \approx 10^{19}$ см⁻³) приводит к величине порядка 10⁻¹¹ %. Измерение этой величины произвести практически невозможно, поскольку этот показатель намного превышает метрологические возможности современных измерительных средств. В этой связи можно утверждать, что для *макроскопических систем результаты статистического анализа имеют практически достоверный характер*. Однако эта ситуация изменяется радикальным образом, если наноразмерная структура содержит весьма незначительное количество частиц (см. столбцы со значениями N , заключенными в интервале величин 10 – 10⁴ микро-частиц).

В литературных источниках довольно часто встречаются сообщения о попытках создания наноразмерных конструкций, состоящих всего из 10 – 100 атомов (в ряде случаев речь идет об изделиях, состоящих

всего из одного атома!). В свете данных табл. 2.13, успешность реализации функционального назначения (качества) подобного рода устройств представляется довольно проблематичной, если речь идет о температурах эксплуатации этих изделий, превышающих несколько градусов Кельвина.

б) Статистическая энтропия наноразмерных структур

Характерной особенностью развития технических изделий как на современном этапе их создания, так и в историческом разрезе, является повышение их сложности в структурном и материальном плане. Эта сложность, в первую очередь, ассоциируется в конструкторско-технологической практике с необходимостью использования все более упорядоченных, с атомно-молекулярной точки зрения, материальных сред.

В максимальной мере фактор упорядоченности материальных сред проявляется в изделиях современной электронной техники, для конструктивных материалов которых допустимый уровень примесных включений (в настоящее время) не превышает 10^{-4} % (подробнее этот вопрос рассмотрен в разд. 3). Следует также отметить, что характерный объем элементов конструкций современных электронных приборов приближается к значению $\Delta V = 10^{-3}$ мкм³ (в промышленной практике), из чего следует, что эти элементы содержат в своем составе не более $\Delta N = 10^9$ атомов. В опытно-конструкторских разработках, основанных на использовании традиционной планарной технологии, эти показатели соответственно равны $\Delta V = 10^{-6}$ мкм³; $\Delta N = 10^6$. В настоящее время на этапе становления находится революционно новая технология атомной сборки изделий наноразмерного масштаба, охватывающих широкую отраслевую направленность (электроника, химия, биология, робототехника и т.д.). В этих наноразмерных элементах конструкций и изделий количество индивидуальных атомов исчисляется значениями $10^2 - 10^4$. В этой связи представляется очевидным, что проблематика обеспечения порядка в пространственном расположении и избирательности атомов определенного типа при создании изделий приобретает исключительный интерес.

В свете вышеизложенных моментов следует отметить, что наиболее близко к представлениям о мере порядка в атомно-молекулярных средах подошла современная статистическая физика. Именно ее подходы позволили сформировать конструктивные представления, необходимые для введения в рассмотрение меры упорядоченности физических сред. В качестве такой *меры упорядоченности* материальных сред выступает *статистическая энтропия* физических систем (понятие энтропии ранее было сформировано в рамках равновесной термодинамики, но ввиду ее формального (феноменологического) характера и применимости исключительно к макроскопическим системам оно представляется малопригодным для рассматриваемого круга вопросов).

Составить представление о физической сущности статистической энтропии можно на основе конкретного примера, в котором рассматривается достаточно простая физическая система – «идеальный газ». Пусть имеется замкнутый объем газа V с общим количеством одинаковых микрочастиц N , который перегороден проницаемой для микрочастиц мембраной на два одинаковых объема $V/2$. В принципе, эти два объема могут поделить между собой общее количество микрочастиц в любой пропорции (в том числе теоретически допустимо, но маловероятно, что все микрочастицы N соберутся в одном из указанных объемов).

Для определенности обозначим условно эти объемы соответственно индексами A и B и определим вероятность P_n события, состоящего в том, что n микрочастиц из N будет обнаружено в объеме A и $(N - n)$ микрочастиц – в объеме B . Общее количество вариантов $\omega(n, N - n)$, с помощью которых можно реализовать ситуацию $A(n)$ и $B(N - n)$, определяется выражением [7]

$$\omega(n, N - n) = \frac{N!}{n! \cdot (N - n)!} \quad (2.80)$$

В табл. 2.14 представлены значения функции $\omega(n, N - n)$, рассчитанные с помощью выражения (2.80) для случая $N = 20$.

Таблица 2.14. Значения функции $\omega(n, N - n)$, вероятности макросостояний P_i и частной энтропии макросостояний S_i для случая $N = 20$

$(i) - \text{№}$ макросо- стояния системы	Тип макросостояния системы		Количество вариантов $\omega(n, N - n)$	Вероят- ность P_i	Энтропия макросост. S_i , Дж/К
	$A(n)$	$B(N - n)$			
1	0	20	1	$\approx 10^{-6}$	0
2	1	19	20	$\approx 2 \cdot 10^{-5}$	$4.1 \cdot 10^{-23}$
3	2	18	190	$\approx 2 \cdot 10^{-4}$	$7.2 \cdot 10^{-23}$
4	3	17	$1.14 \cdot 10^3$	0.001	$9.7 \cdot 10^{-23}$
5	4	16	$4.85 \cdot 10^3$	0.005	$1.2 \cdot 10^{-22}$
6	5	15	$1.55 \cdot 10^4$	0.015	$1.3 \cdot 10^{-22}$
7	6	14	$3.88 \cdot 10^4$	0.037	$1.5 \cdot 10^{-22}$
8	7	13	$7.75 \cdot 10^4$	0.074	$1.6 \cdot 10^{-22}$
9	8	12	$1.26 \cdot 10^5$	0.120	$1.7 \cdot 10^{-22}$
10	9	11	$1.68 \cdot 10^5$	0.160	$1.8 \cdot 10^{-22}$
«11»	10	10	$1.85 \cdot 10^5$	0.176	$1.9 \cdot 10^{-22}$
Далее результаты повторяются в убывающем порядке					
-	-	-	$\sum = 105$	$\sum = 1.0$	$S = 2.0 \cdot 10^{-22}$

Для дальнейшего рассмотрения уместно привести ряд комментариев относительно отдельных позиций данных табл. 2.14:

- в первом столбце перечислены различные *макросостояния системы* (обозначенные индексом i), характеристика которых представлена в объединенном столбце № 2;
- в третьем столбце приведено полное количество *микросостояний системы* (т.е. произведение количества различных вариантов размещения n микрочастиц в объеме A на количество вариантов размещения $(N - n)$ микрочастиц в объеме B), реализующих определенное макросостояние системы $i^{\text{го}}$ типа;
- в четвертом столбце приведена *вероятность реализации макросостояния системы $i^{\text{го}}$ типа* (т.е. размещения n микрочастиц в объеме A и размещения $(N - n)$ микрочастиц в объеме B), определяемая согласно выражению

$$P_i = \frac{\omega_i(n, N-n)}{\sum_i \omega_i(n, N-n)}; \quad (2.81)$$

- в пятом столбце приведена *частная статистическая энтропия* макросостояния i^{20} типа (характеризующая меру неопределенности обнаружения макросостояния « $A(n) - B(N-n)$ »), определяемая зависимостью

$$S_i = k \cdot \ln \omega_i(n, N-n), \quad (2.82)$$

где k – постоянная Больцмана;

- в нижней строке представлены сводные данные по всем рассматриваемым статистическим параметрам макроскопической системы, и в частности, *полная статистическая энтропия* S макроскопической системы в целом (от суммарного количества вариантов размещения не взаимодействующих между собой микрочастиц, с учетом всех возможных частных реализаций при заполнении объемов $A(n)$ и $B(N-n)$, с общим количеством микрочастиц N):

$$S_i = k \cdot \ln \sum_i \omega_i(n, N-n); \quad (2.83)$$

- вероятность реализации макросостояний с номерами $i = 1, 2, 3, 19, 20, 21$ исчезающе мала уже при столь незначительных значениях $N = 20$, и эта тенденция будет неизмеримо усугубляться с возрастанием количества микрочастиц в системе N ;

- суммарная вероятность реализации ряда макросостояний, расположенных в центральной части таблицы с номерами $i = 9, 10, 11$ (равновесное), 12, 13, составляет 0.74, что свидетельствует о том обстоятельстве, что рассматриваемая макроскопическая система, как правило, чаще всего будет находиться в одном из макросостояний из этой указанной группы.

Следует отметить, что большие значения вероятностей обнаружения центральных макросостояний имеют место не потому, что природа «любит» эти равновесные макросостояния по каким-то дополнительным признакам, а по той простой причине, что эти макросостояния

реализуются с помощью *огромного количества микросостояний* (вариантов). При этом каждое микросостояние, независимо от его принадлежности к каждому конкретному макросостоянию, имеет одинаковую вероятность P_{micro} своего проявления. Для рассмотренного примера эта вероятность проявления индивидуального микросостояния P_{micro} имеет значение весьма близкое к нулю:

$$P_{micro} = \frac{1}{\sum_i \omega_i(n, N-n)} \approx 10^{-6}. \quad (2.84)$$

В реальной физико-технической практике оперируют с минимальным объемом физических сред порядка 10^{-9} мкм³ (содержащих в своем составе не 20 атомов, как это было в рассмотренном примере, а $N = 10^3$ атомов), и в этом случае, согласно выражениям (2.80) и (2.84), вероятность проявления индивидуального микросостояния P_{micro} можно считать практически равной нулю.

Довольно близкие значения вероятностей появления макросостояний с номерами № 9, 10, 11 (близких к равновесному состоянию) обусловлены достаточно малым количеством атомов в системе. С ростом количества атомов в системе это различие будет проявляться в более отчетливой форме. Это обстоятельство представлено в табл. 2.15, в которой приведены значения отношения вероятности равновесного макросостояния $P_{равнов} = P\left(\frac{N}{2}, \frac{N}{2}\right)$ (количество атомов в ящиках *A* и *B* одинаково) к вероятности макросостояния $P_{10\%} = P\left[\left(\frac{N}{2} - 0.1 \cdot \frac{N}{2}\right), \left(\frac{N}{2} + 0.1 \cdot \frac{N}{2}\right)\right]$ (т.е. при 10% отклонении от равновесного макросостояния) при различных значениях общего количества атомов в системе N .

Из данных табл. 2.15 следует, что при значениях параметра $N > 1000$ можно считать, что вероятность появления равновесного макросостояния $P_{равнов} = P\left(\frac{N}{2}, \frac{N}{2}\right)$ становится практически равной единице, а появление остальных макросостояний становится весьма маловероятным событием.

Таблица 2.15. *Зависимость отношения $P_{\text{равное}}/P_{10\%}$ от значения параметра N*

N	20	100	200	500	1000
$P_{\text{равное}}/P_{10\%}$	1.10	1.64	2.71	12.2	149

Интересно применить результаты, представленные в табл. 2.14 и 2.15, к процессу создания технических изделий (конструкций). Характерной особенностью этих изделий является ярко выраженный неравновесный характер физических сред, используемых для обеспечения функционального качества изделия.

Например, в диодной структуре, основанной на базе резкого ($p-n$) перехода, с одной стороны от границы раздела находится донорная примесь, а с другой – акцепторная примесь, что свидетельствует о ярко выраженном неравновесном состоянии структуры в целом. Этому неравновесному состоянию структуры, обеспечивающему эффект выпрямления электрических сигналов, можно, с некоторой долей условности, поставить в соответствие макросостояние №1 табл. 2.14. Представляется очевидным, что ожидать от «природы» спонтанного возникновения указанной диодной структуры, реализуемой с помощью всего лишь одного микросостояния, не представляется возможным ввиду чрезвычайно малой вероятности ее проявления. Именно по этой причине в производственно-технической практике приходится использовать различного рода технологии по преобразованию рабочих сред из равновесного состояния, характерного для сырьевых ресурсов, в требуемое неравновесное состояние, характерное для изделия (подробнее этот вопрос с учетом аспектов производственного характера рассмотрен в разд. 3).

В процессе эксплуатации изделия указанная начальная неравновесность состояния физических сред изделия является причиной активизации процессов диффузионного перехода примесей в соседние области, что в итоге приводит к перемешиванию донорной и акцепторной примесей. Конечным итогом этого процесса перемешивания примеси является образование однородной (равновесной) рабочей среды (с полной

потерей функционального назначения диодной структуры), которой можно поставить в соответствие равновесное макросостояние позиции №11 табл. 2.14, с учетом фактора голоvizны этого равновесного состояния, зависимость вероятности которого от количества микрочастиц в системе представлена в табл. 2.15.

2.2.2. Носители заряда для случая наноразмерных структур кристаллического типа

а) Специфика формализма зонных энергетических диаграмм для случая наноразмерных структур

В разд. 2.1.3 (пункт а) рассмотрен механизм изменения скорости электрона при его движении в кристаллических структурах электронного типа под воздействием внешнего электрического поля. Показано, что ключевым моментом при реализации векторизованной формы движения (а следовательно, и образования электрического тока), обеспечивающим возможность изменения скорости электрона под воздействием электрического поля, является фактор *частичного заполнения* электронами разрешенной энергетической зоны (проводимости и валентной). Указанное частичное заполнение энергетических зон позволяет практически непрерывное изменение энергии электрона при его нахождении в пределах рассматриваемой энергетической зоны (с учетом того обстоятельства, что расстояние между соседними энергетическими уровнями в зоне имеет очень малую величину и составляет по порядку 10^{-23} эВ). Именно в этой связи можно утверждать о квазинепрерывном характере динамики векторизованного движения электрона проводимости в веществе.

Для макроскопических кристаллических структур характерным фактором является чрезвычайно большое количество энергетических состояний в разрешенных энергетических зонах, а следовательно, и значительная их протяженность по энергетической шкале. В этой связи частичное заполнение энергетических зон электронами оставляет определенную возможность (благодаря наличию большого количества свободных энергетических уровней) для изменения энергетических

состояний указанных электронов под воздействием внешнего электрического поля, что и приводит в конечном счете, к реализации векторизованной формы движения электронов (т.е. электрического тока).

Распространим рассмотренное выше модельное представление зонного механизма электронной проводимости на случай наноразмерных структур кристаллического типа. Сразу же необходимо отметить, что дискретный (а не квазинепрерывный) характер распределения энергетических уровней, свойственный наноструктурам с ограничениями размерного характера, не позволяет в полной мере воспользоваться формализмом зонных энергетических диаграмм, свойственным макроскопическим структурам. Поэтому речь может идти только о ряде качественных аспектов зонных диаграмм наноструктур, представляющих практический интерес при рассмотрении вопроса электрической проводимости в наноразмерных кристаллических структурах.

В упрощенном (примитивном) виде зонное энергетическое представление возможных состояний наноразмерной структуры отображено на рис. 2.17 (на примере наноструктуры типа «квантовой пленки» – (2D)).

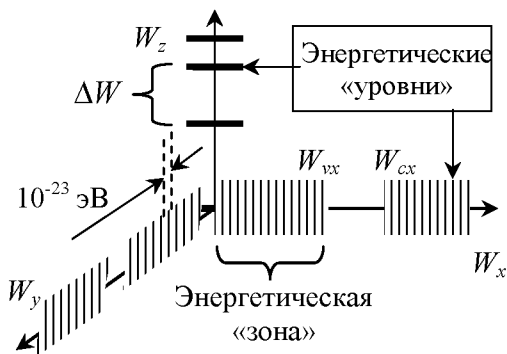


Рис. 2.17. Схематическое «координатное» представление зонной энергетической диаграммы наноразмерной кристаллической структуры

На рисунке каждому координатному направлению поставлен в соответствие фрагмент зонной энергетической диаграммы, ответственный за реализацию векторизованной формы движения по данному

координатному направлению. Количество энергетических уровней, содержащихся в каждой энергетической «зоне», соизмеримо с количеством атомов в линейной цепочке атомов по рассматриваемому координатному направлению. Жирными линиями условно отражены узкие зоны в координатном направлении с пространственным ограничением наноразмерной кристаллической структуры, характерным для квантовой точки.

Согласно схематическому представлению рис. 2.17, элементарный процесс векторизованного изменения скорости носителя заряда, сопровождаемый переходами электрона (под воздействием внешнего электрического поля) на соседние энергетические уровни, имеет:

- практически *безактивационный* характер ($W_{axy} \approx 10^{-22}$ эВ, где W_{axy} – энергия активации элементарного акта энергетического перехода на зонной диаграмме) для координатных направлений x и y ;
- ярко выраженный *активационный* характер при движении носителя заряда вдоль координатного направления z .

(Под понятием «*активационный*» подразумевается такой механизм протекания элементарного физико-химического процесса, при котором успешная реализация этого процесса становится возможной лишь в том случае, если энергия микрочастиц превышает некоторое пороговое значение (потенциальный барьер) W_{axy} , называемое *энергией активации* рассматриваемого физико-химического процесса.)

Интересно оценить масштабы значений вероятности реализации энергетических переходов, стимулированных влиянием внешнего электрического поля, для рассмотренных выше частных случаев. Полагая, что электронный газ в энергетических зонах является невырожденным (фактор частичного заполнения зон), воспользуемся классической статистикой Максвелла-Больцмана, согласно которой вероятности переходов (а следовательно, и изменения скорости носителя заряда) P_E можно в первом приближении оценить с помощью выражения

$$P_E = \exp\left(-\frac{W_a - q \cdot \lambda \cdot E}{k \cdot T}\right), \quad (2.85)$$

где λ – длина свободного пробега электрона в проводящей среде; q – электрический заряд носителя; E – напряженность электрического поля.

В выражении (2.85) находит свое отражение тот томент, что энергетическая составляющая от движения электрона в электрическом поле способствует снижению (преодолению) потенциального барьера W_a , стоящего на пути процесса изменения энергетического состояния электрона в рамках зонной диаграммы.

Для случаев координатных направлений x и y (т.е. в условиях полного отсутствия ограничений размерного характера) вероятность перехода $P_{Ex} = P_{Ey}$, ввиду чрезвычайной малости параметра W_{axy} (соответственно нет никакой необходимости принимать меры для уменьшения энергии активации W_{axy}), составит (при $T = 300$ K):

$$P_{Ex} = P_{Ey} \approx \exp\left(-\frac{W_{axy}}{k \cdot T}\right) = \exp\left(-\frac{10^{-22} \text{ эВ}}{0.025 \text{ эВ}}\right) \approx 1. \quad (2.86)$$

Именно исходя из полученного выше результата, можно утверждать о безактивационном характере механизма электропроводности (переходов) для случая координатных направлений, не имеющих ограничений размерного характера.

Проведем аналогичную оценку вероятностей переходов по координатному направлению z , имеющему ограничения размерного характера. В этом случае *значительные* изменения скорости могут иметь место при реализации перехода на соседний энергетический уровень, отстоящий от исходного уровня на значительном (по сравнению с W_{axy}) расстоянии ΔW (см. рис. 2.17). Вероятность подобного перехода P_{Ez} в соответствии с выражением (2.86) определяется в виде

$$P_{Ez} = \exp\left(-\frac{\Delta W - q \cdot \lambda \cdot E}{k \cdot T}\right). \quad (2.87)$$

В табл. 2.16 представлена зависимость вероятности перехода P_{Ez} от величины энергетического интервала ΔW и напряженности электрического поля E при длине свободного пробега микрочастиц $\lambda = 10^{-8}$ м (типичное значение для твердотельных кристаллических структур).

Т а б л и ц а 2.16. *Зависимость вероятности перехода P_{Ez} от величины энергетического интервала ΔW и напряженности электрического поля E ($\lambda = 10^{-8} \text{ м}$; $T = 300 \text{ К}$)*

ΔW , эВ	Напряженность электрического поля, В/м					
	10^3	10^4	10^5	10^6	10^7	10^8
0.01	0.671	0.673	0.698	1.0	1.0	1.0
0.05	0.135	0.136	0.141	0.202	1.0	1.0
0.10	0.018	0.018	0.019	0.027	1.0	1.0
0.50	$2.0 \cdot 10^{-9}$	$2.0 \cdot 10^{-9}$	$2.1 \cdot 10^{-9}$	$3.1 \cdot 10^{-9}$	$1.1 \cdot 10^{-7}$	1.0

Приведенные в табл. 2.16 данные свидетельствуют о том, что в направлениях с размерными ограничениями кристаллической структуры реализация векторизованной формы движения носителя заряда (по рассматриваемому прямому межзонному механизму процесса электропроводности):

- представляется достаточно маловероятным событием при напряженностях внешнего электрического поля $E < 10^5 - 10^6$ В/м (электрические режимы, характерные для обеспечения функционированного назначения современных полупроводниковых структур в устройствах микроэлектроники) и при значениях энергетического интервала между соседними уровнями $\Delta W > 0.1$ эВ;

- представляется весьма вероятной при напряженностях внешнего электрического поля $E > 10^6 - 10^7$ В/м (значение, сопоставимое с показателями электрической пробивной прочности для широкого круга конструкционных диэлектрических материалов).

Обращение к межзонному механизму процесса электропроводности обусловлено тем обстоятельством, что в направлениях с размерными ограничениями кристаллической структуры ширина энергетических зон крайне мала (этот факт условно отображен на рис. 2.17 – «жирные» штрихи на оси W_z вместо зон конечной ширины). Протяженность ΔW_z этих энергетических зон можно оценить согласно выражению

$$\Delta W_z = W_{axy} \cdot N_z \approx 10^{-22} \cdot N_z, \quad (2.88)$$

где N_z – количество атомов (в цепочке) в направлении с размерными ограничениями кристаллической структуры, эВ.

Для типичных наноразмерных кристаллических структур ($N_z < 10^3$) протяженность ΔW_z , согласно выражению (2.88), не превышает 10^{-19} эВ. Именно в таких незначительных масштабах следует ожидать проявление эффектов векторизации движения электрона, основанных на базе зонных механизмов токопрохождения.

Представленные выше материалы не преследуют цель строгого анализа вопроса, а лишь способствуют его пониманию на качественном уровне рассмотрения. Строгое рассмотрение энергетических состояний и динамики движения электронов в реальных кристаллических структурах предполагает использование более сложных модельных представлений, и в частности, о *зонах Бриллюэна* [9]. Зона Бриллюэна определяется как область в обратном пространстве кристаллической решетки, окружающей один из узлов обратной решетки и ограниченной совокупностью плоскостей, проходящих через середины векторов, соединяющих эту точку с ее ближайшими соседями в обратной решетке. Подобный подход особенно эффективен для кристаллических структур, в которых поверхности постоянной энергии (см. разд. 1.2.2, в) в формализации «*k*-пространства» сильно отличаются от сферической формы, характерной только для ряда простых кристаллических структур кубического типа. Детальное рассмотрение этого подхода требует достаточно высокой специальной физико-математической подготовки и в этой связи этот подход вынесен за рамки целевых задач настоящего пособия.

б) Анализ возможных скоростей движения носителя заряда в электрическом поле

В модели свободных электронов можно было говорить о взаимосвязи энергии и волнового вектора индивидуального электрона (см. разд. 1), не рассматривая ограничений, накладываемых фактором периодического потенциального поля, обусловленного наличием кристаллической структуры. В теории свободных электронов (т.е. при полном отсутствии потенциальных полей) эта взаимосвязь выглядит достаточно просто (дисперсионное уравнение, см. выражение (1.15)):

$$W_k = \frac{\hbar^2}{8 \cdot m \cdot \pi^2} \cdot k^2. \quad (2.89)$$

В рамках модели *свободных* электронов им разрешено иметь *любое значение энергии*, отвечающее некоторому вещественному значению волнового вектора k .

Наличие внутреннего периодического потенциального поля, собственного кристаллическим структурам, усложняет вид зависимости $W_k = \varphi(k)$, и в частности:

- приводит к появлению на дисперсионных зависимостях разрешенных и запрещенных энергетических зон,
- некоторым образом изменяет вид дисперсионной зависимости (см. рис. 1.13) от параболического вида (выражение (2.89)), в особенности на верхней границе индивидуальной разрешенной энергетической зоны.

В этом случае электроны, находящиеся в рамках кристаллической структуры, уже не могут рассматриваться в качестве свободных микрочастиц, с той точки зрения, что на энергию их движения распространяется ограничение (запрещенные энергетические зоны). При этом электроны могут считаться «свободными» в том смысле, что, обладая определенной кинетической энергией в пределах разрешенной зоны, они могут беспрепятственно двигаться по кристаллу сколь угодно долго (практически без рассеяния, в случае отсутствия точечных, линейных и поверхностных дефектов), пока не достигнут границы кристаллической структуры, т.е. нарушения периодичности потенциала.

Если все энергетические состояния в разрешенной зоне *заполнены электронами*, то принцип запрета Паули не допускает *никакого отклика* системы на внешнее электрическое поле (см. модельное представление, разд. 2.1.3,а). Последнее означает, что указанные электроны просто «игнорируют» наличие внешнего электрического поля и по этой причине не являются участниками процесса токопрохождения (квантово-механический феномен, не имеющий своего аналога в классической физике).

Для случая энергетической зоны, которая частично заполнена электронами, реакция на внешнее энергетическое воздействие, коим является и электрическое поле, затрагивает преимущественно те электроны, которые обладают максимальной энергией, т.е. занимают энергетиче-

ские состояния близкие к энергии Ферми (в целях упрощения, рассматриваются кристаллические структуры металлического типа). Последнее объясняется тем, что именно для этих электронов имеется большое количество свободных энергетических уровней в разрешенной зоне, которые могут быть задействованы в процесс изменения энергии (а следовательно, и скорости движения) электронов.

Рассмотрим поведение электронов, находящихся в кристаллической структуре, в условиях воздействия на них внешнего электрического поля. В целях упрощения анализа и не умаляя общности рассмотрения, примем, что в пустой разрешенной зоне находится всего лишь один электрон. Примем также, что этот электрон может, благодаря воздействию внешнего электрического поля, иметь векторизованную форму движения по кристаллической структуре, не испытывая при этом процессов рассеяния.

В одномерном приближении и при использовании допущения, что периодический потенциал имеет такой вид, что энергия электрона минимальна (W_{\min}) в центре разрешенной зоны (при $k = 0$) и максимальна (W_{\max}) на границах зоны (при значениях волнового вектора $k = \pm(\pi/a)$, где a – параметр решетки), зависимость $W(k)$ может быть отображена в виде, представленном на рис. 2.18 [8].

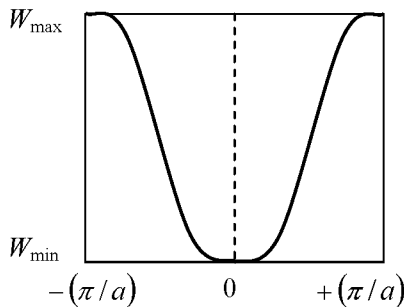


Рис.2.18. Зависимость энергии электрона в разрешенной энергетической зоне от волнового вектора

Следует обратить внимание на тот факт (см. рис. 2.18), что зависимость $W(k)$ весьма сильно *отличается от параболического вида* (см.

выражение (2.89)) на краях энергетической зоны, т.е. при значениях волнового вектора близких $k \pm (\pi/a)$. Это обстоятельство имеет большое значение для последующего рассмотрения.

На дне энергетической зоны (т.е. в области значений $k < |\pm(\pi/a)|$) зависимость $W(k)$ можно представить в виде

$$W(k) = W_{\min} + \frac{h^2}{8 \cdot m \cdot \pi^2} \cdot k^2. \quad (2.90)$$

Скорость движения электрона на этом энергетическом интервале *линейно зависит* от волнового вектора k и может быть определена на основании выражения (2.90) в виде [Блейкмор]

$$v(k) = \frac{2 \cdot \pi}{h} \cdot \frac{dW}{dk} = \frac{h \cdot k}{2 \cdot \pi \cdot m}. \quad (2.91)$$

Своего максимального значения скорость электрона достигает не на границе энергетической зоны (как это можно ожидать из модели движения свободного электрона, согласно выражению (2.89)), а при некотором промежуточном значении волнового вектора k , соответствующем середине зоны. Это следует из того, что вблизи потолка энергетической зоны зависимость $W(k)$ имеет вид

$$W(k) = W_{\max} - \frac{h^2}{8 \cdot m \cdot \pi^2} \cdot \left(k \pm \frac{\pi}{a} \right)^2, \quad (2.92)$$

а, соответственно, скорость электрона описывается выражением

$$v(k) = -\frac{h}{2 \cdot \pi \cdot m} \cdot \left(k \pm \frac{\pi}{a} \right). \quad (2.93)$$

Из выражения (2.93) следует, что по мере приближения волнового вектора k к граничным значениям $k = \pm(\pi/a)$ скорость движения электрона уменьшается и в пределе (т.е. при $k = \pm(\pi/a)$) обращается в ноль. Подобное («парадоксальное», с позиций представлений классической физики) поведение скорости электрона объясняется тем обстоятельством, что при значениях $k < \pm(\pi/a)$ состояние электрона описы-

вается *бегущей волной* де Бройля (при этом символы (\pm) характеризуют направление скорости движения), а при значениях $k = \pm(\pi/a)$ состояние электрона описывается *стоячей волной* [8].

В свете вышеизложенных моментов зависимость скорости движения электрона $v(k)$, находящегося в разрешенной энергетической зоне, от волнового вектора k может быть представлена в виде, приведенном на рис. 2.19.

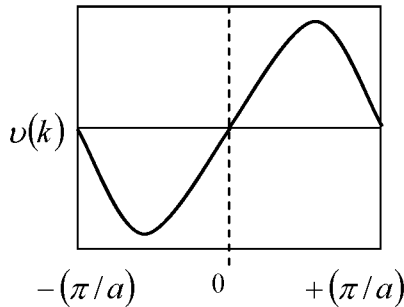


Рис.2.19. Зависимость скорости движения электрона в разрешенной энергетической зоне от волнового вектора

Рис. 2.19 демонстрирует следующие принципиальные моменты, важные в части их использования при анализе вопроса электропроводности кристаллических структур, обладающих электронным механизмом проводимости:

- максимальная скорость движения электрона соответствует не граничным значениям волнового вектора $k = \pm(\pi/a)$, а некоторому промежуточному значению волнового вектора, находящемуся в диапазоне значений: $(\pi/a) < k < 0$ (прямая волна); $0 < k < (\pi/a)$ (обратная волна);
- значениям волнового вектора, удовлетворяющим условию $dv(k)/dk = 0$, соответствует максимальная скорость движения электронов, энергия которых находится в пределах разрешенной энергетической зоны;
- характер поведения скорости движения электрона на участке от v_{\max} до границ энергетической зоны (по координатной шкале k)

не имеет классического аналога и, более того, принципиально расходится с известными представлениями классической физики.

Если внешнее электрическое поле приложено в течение длительного периода времени, то электрон будет двигаться по кристаллической структуре (с возрастанием волнового вектора k) до тех пор, пока не будет выполнено условие $k = \pm(\pi/a)$, при котором $v(\pm(\pi/a)) = 0$. Возможность дальнейшей эволюции скорости движения электрона зависит от того, насколько велико значение напряженности электрического поля.

В том случае, если напряженность внешнего электрического поля E будет слишком велика (конкретизация понятия «велика» дана ниже), то существует конечная вероятность события, что электрон осуществит переход (прыжок) через запрещенную зону в энергетическое состояние с волновым вектором $k = -(\pi/a)$, которое находится в следующей, более высоко расположенной на энергетической диаграмме разрешенной зоне. Вероятность такого (*туннельного*) перехода электрона через запрещенную зону достаточно велика, если выполняются условия [9]:

$$[\Delta W / (q \cdot E)] \leq \lambda; \quad [\Delta W / (q \cdot E)] \leq a, \quad (2.94)$$

где ΔW – ширина запрещенной зоны; λ – длина волны де Бройля; q – заряд электрона; a – параметр кристаллической решетки.

Вероятность туннельного перехода электрона $P_{\text{тун}}$ через запрещенную зону под воздействием внешнего электрического поля (*туннелирование Зинера*) определяется выражением [9]:

$$P_{\text{тун}} \approx \exp\left(-\frac{a \cdot m_e \cdot \Delta W}{h^2 \cdot q \cdot E}\right), \quad (2.95)$$

где m_e – эффективная масса электрона.

Характерной особенностью процесса генерации новых носителей заряда по туннельному механизму является независимость вероятности перехода электрона через запрещенную зону от температурного фактора. Этим электростатическая генерация носителей заряда отличается от традиционного термоэлектронного типа.

Согласно выражению (2.95), вероятность туннельного перехода электрона через запрещенную зону при значениях параметров $\Delta W = 1$ эВ и $a = 10^{-10}$ м (типичные величины для широкого круга конструктивных кристаллических материалов полупроводникового типа) становится пренебрежимо малой величиной, если напряженность электрического поля составляет значение менее $E = 10^7 - 10^8$ В/м (в части преемственности анализа уместно сравнить приведенное значение напряженности поля с аналогичными оценочными данными, представленными выше в табл. 2.15).

Интересно отметить, что, вопреки ожиданиям, при туннельном переходе не происходит скачкообразного повышения энергии перешедшего электрона, несмотря на то что электрон при этом переходит на значительно более высоко расположенный на зонной диаграмме энергетический уровень. Этот момент представлен на рис. 2.20, на котором отображены две соседние энергетические зоны, которые, в отличие от всех предыдущих отображений зонных энергетических диаграмм, имеют *наклонный* вид. Последнее обстоятельство объясняется проявлением влияния внешнего источника электрического потенциала (подключенного к конечным точкам проводящей среды) на энергетические состояния индивидуальных электронов.

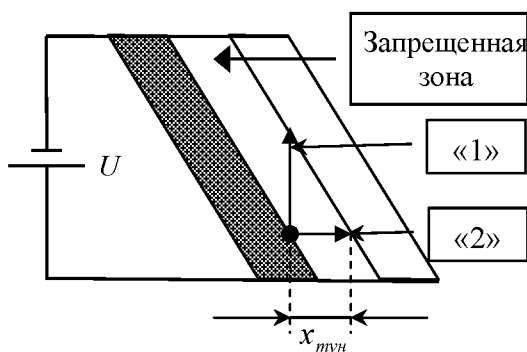


Рис.2.20. Схематическое отображение прямого и туннельного перехода электрона через запрещенную зону (U – внешний источник напряжения)

Согласно модельному представлению рис. 2.20, у электронов существует две альтернативные возможности преодоления запрещенной зоны:

- с помощью *прямого* перехода электрона (обозначен индексом «1» на рис. 2.20), реализация которого имеет ярко выраженный активационный характер и связана с необходимостью дополнительных энергетических затрат ΔW ;
- с помощью *туннельного* перехода электрона (обозначен индексом «2» на рис. 2.20), который имеет безактивационный характер и в этой связи не связан с необходимостью каких-либо энергетических затрат.

Из рисунка следует, что чем больше наклон энергетической диаграммы (т.е. величина приложенного внешнего электрического потенциала), тем меньше длина пути туннельного прохождения $x_{\text{тун}}$, а следовательно, и большей становится вероятность реализации этого механизма преодоления запрещенной зоны. Увеличение концентрации носителей заряда за счет туннельного механизма токопрохождения не носит на первых порах лавинообразного характера, так как этот процесс компенсируется процессом рекомбинации (обратными переходами) носителей заряда. Однако при очень высоких значениях напряженности электрического поля ($E = 10^8 - 10^9$ В/м) процесс генерации носителей заряда по туннельному механизму приобретает лавинообразный характер, что, как правило, приводит к электрическому пробоею проводника электричества.

Другим, интересным с практической точки зрения, фактором при туннельном механизме генерации (переходов) носителей заряда является возможность реализации *отрицательной дифференциальной проводимости* (эффект Гана) в полупроводниках. В обычных резистивных элементах (с положительной проводимостью) векторизованное движение носителя заряда в конечном итоге превращается в броуновское (хаотическое) движение, т.е. имеет место процесс «деградации» энергии от начальной организованной электрической формы в неорганизованную – тепловую. В резистивных элементах с отрицательной дифференциальной проводимостью происходит обратный процесс – повыше-

ние степени векторизации энергии в части превращения постоянного тока в переменный ток упорядоченного (в частности, синусоидального) типа.

В общем случае электрический ток в веществе I на основании выражения для статической проводимости (2.1) может быть выражен в виде (закон Ома):

$$I = q \cdot n \cdot \mu \cdot E = \sigma_0 \cdot E, \quad (2.96)$$

где σ_0 – статическая электропроводность вещества.

В соответствии с выражением (2.96) дифференциальная электропроводность вещества будет равна

$$\sigma_o = \frac{dI}{dE} = q \cdot n \cdot \mu + q \cdot n \cdot E \cdot \frac{d\mu}{dE} = \sigma_0 \left[1 + \frac{E}{\mu} \cdot \frac{d\mu}{dE} \right]. \quad (2.97)$$

В принципе, выражение, представленное в квадратной скобке, может иметь как положительный, так и *отрицательный* знак. Второй вариант имеет место в том случае, если с ростом напряженности электрического поля подвижность носителя заряда уменьшается настолько, что выполняется не только условие $\frac{E}{\mu} \cdot \frac{d\mu}{dE} < 0$, но и условие $\left| \frac{E}{\mu} \cdot \frac{d\mu}{dE} \right| > 1$.

С физической точки зрения выполнение этих условий свидетельствует о том, что внешнее электрическое поле «разогревает» электронный газ, вследствие чего его температура T_e может существенно превышать равновесную температуру T кристаллической структуры. Это обстоятельство приводит к деформации функции плотности распределения электронов по энергиям, т.е. сдвигу максимума функции распределения электронов в сторону более высоких значений энергии.

в) Влияние поверхности раздела физических сред на вид зонных энергетических диаграмм наноразмерных структур

В наноразмерных кристаллических структурах доля поверхностных атомов по отношению к количеству атомов, расположенных непосредственно в объеме структуры, может достигать довольно большого значения. В то же время наличие поверхностных атомов как таковых свидетельствует о факторах:

- *нарушения периодичности* потенциальной энергии кристаллической решетки в областях пространства, прилегающих к поверхностям раздела сред;

- появления на поверхностях раздела сред атомов, обладающих *ненасыщенными валентными связями*.

Оба эти фактора вносят свои коррективы на вид зонных энергетических диаграмм в наноразмерных кристаллических структурах. К общим моментам указанных корректив следует отнести факты появления на зонных энергетических диаграммах наноразмерных структур:

- большого количества энергетических уровней (*поверхностных состояний*), расположенных в запрещенной зоне энергетической диаграммы, в пространственном отношении – в непосредственной близости от поверхности раздела сред;

- изгиба (вверх или вниз) энергетических уровней (*изгиб зон*) на зонных диаграммах (обусловленный электростатическими силами), в областях пространства, непосредственно прилегающих к поверхностям раздела сред.

Поверхностные состояния Тамма

Этот тип поверхностных состояний характерен для *атомарно-чистой* поверхности кристаллической структуры, не содержащей дефекты атомного масштаба (вакансии, примесные атомы, выход линейных дислокаций и т.д.).

В разд. 1.2.2,б рассмотрен вопрос о характере движения электрона в периодическом потенциальном поле, образованном линейной цепочкой атомов *бесконечной* длины, т.е. при отсутствии границ раздела сред (модель Кронига – Пенни). Показано, что в этом случае энергетический спектр разрешенных состояний электрона имеет вид зонной диаграммы, представленной на рис. 1.11.

В рассматриваемом случае речь идет об идеализированной кристаллической структуре, например, линейной цепочке атомов, которая ограничена с одного конца и простирается по координатному направлению x в область значений $x > 0$, а в области значений $x \leq 0$ находится вакуумная среда с постоянным значением потенциала U_0 . Вакуумная среда необходима для обеспечения условия атомарной чистоты

поверхности. Графическое отображение потенциальной функции такой кристаллической структуры приведено на рис. 2.21.

Анализ спектра энергетических состояний представленной кристаллической структуры, с учетом наличия границы раздела физических сред, впервые проведен российским ученым И.Е. Таммом (1932 г.).

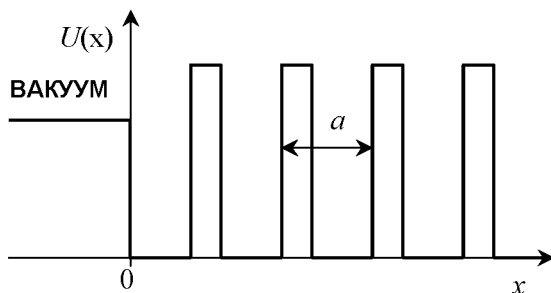


Рис. 2.21. Вид потенциальной функции для ограниченной с одной стороны кристаллической структуры

В частности, были выявлены следующие характерные моменты трансформации вида зонных диаграмм, обусловленные фактом обрыва (т.е. нарушением периодичности) кристаллической решетки:

- обрыв решетки приводит к тому, что в ограниченном кристалле появляются *новые разрешенные энергетические состояния* в тех областях энергий, которые запрещены в неограниченном кристалле;
- волновые функции электронов, расположенных на этих новых энергетических состояниях, *быстро затухают* при удалении от поверхности раздела сред как в область вакуума, так и вглубь кристалла.

Последнее обстоятельство свидетельствует о том, что вероятность обнаружения электрона, находящегося на этих новых энергетических уровнях, вдали от границы раздела сред ничтожно мала. Из чего следует, что эти электроны не могут проникать вглубь кристалла и *локализируются у поверхности раздела*. В этой связи эти новые энергетические уровни или состояния получили название *поверхностных уровней (состояний)*, или *уровней Тамма*.

Следует особо отметить, что поверхностные уровни *не являются дополнительными уровнями*, возникающими сверх того количества,

которое может содержать невозмущенная (т.е. без учета наличия границы раздела физических сред) разрешенная зона (напомним, что количество уровней в разрешенной энергетической зоне определяется количеством атомов, образующих кристаллическую структуру). Эти поверхностные уровни с малой пространственной протяженностью *отщепляются от соответствующих энергетических зон*, согласно модельному представлению, приведенному на рис. 2.22.

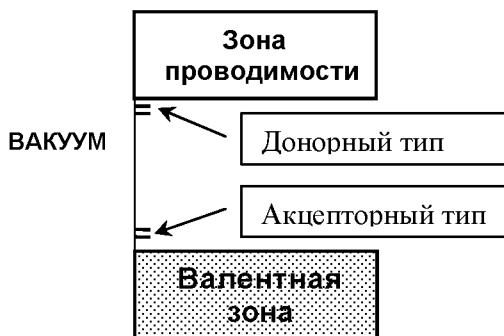


Рис. 2.22. Энергетический спектр поверхностных состояний Тамма

Поверхностные уровни Тамма, находящиеся в запрещенной зоне, могут быть как донорного типа (вблизи дна зоны проводимости), так и акцепторного типа (вблизи потолка валентной зоны).

В более общей постановке вопроса образование поверхностных уровней в запрещенной зоне наблюдается во всех случаях нарушения периодичности потенциала решетки, независимо от того обстоятельства, чем вызвано это нарушение: примесными или адсорбированными атомами, дислокациями линейного типа, границами кристаллических зерен, свободной поверхностью (уровни Тамма).

Поверхностные состояния от адсорбированных на поверхности атомов

Искусственно созданная атомарно-чистая поверхность является таковой недолго, порядка 10^{-9} с, при давлении окружающих газов 10^{-8} мм рт. ст. По истечении этого промежутка времени она покрывается частично или полностью слоем *адсорбированных* атомов или молекул различной химической природы.

Рассмотрим с позиции зонных диаграмм влияние процесса адсорбции атомов на энергетический спектр электронов. Пусть на поверхности кристаллической структуры произошел процесс хемосорбции атома типа A (рис. 2. 23). В результате химического взаимодействия атома A с поверхностными атомами кристаллической структуры их волновые функции перекрываются, что позволяет рассматривать атом A в качестве локальной примеси, нарушающей периодичность потенциального поля кристаллической решетки структуры. В свою очередь, нарушение периодичности потенциального поля приводит к появлению в запрещенной зоне дискретного (т.е. обладающего малой пространственной протяженностью) энергетического уровня B донорного или акцепторного типа.

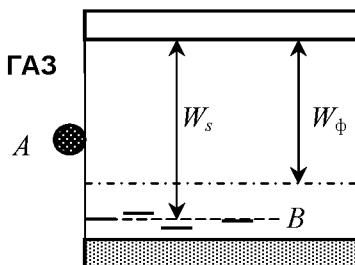


Рис. 2.23. Поверхностный уровень, инициированный адсорбированным атомом (W_Φ – энергия Ферми; W_s – энергия поверхностного уровня)

Для определенности положим, что появившийся поверхностный уровень B имеет акцепторный тип. Положение уровня B на энергетической диаграмме зависит от физической природы атомов структуры и сорбируемой частицы A и расстояния между частицей A и поверхностью структуры. С увеличением этого расстояния уровень B поднимается вверх, и при значительном удалении частицы A от поверхности он «втягивается» в зону проводимости. Волновая функция электрона, находящегося на уровне B , имеет *ярко выраженный максимум* вблизи положения атома A , причем величина максимума возрастает по мере перемещения уровня B вниз, к валентной зоне. Благодаря близости уровня B к валентной зоне вероятность ее заселения электроном доста-

точно велика, что приводит к появлению в приповерхностной области кристаллической структуры *избыточного отрицательного заряда*.

Аналогичное рассмотрение можно провести и для случая, когда в результате хемосорбции атомов на поверхности кристаллической структуры образуются поверхностные состояния донорного типа, с образованием в приповерхностной области *избыточного положительного заряда*.

В общем случае вероятность того, что поверхностный уровень B будет заполнен электроном, определяется функцией распределения частиц по энергиям Ферми-Дирака (см. разд. 2.1.3,в):

$$f^{(-)} = \left(1 + e^{-\frac{W_s - (-W_\Phi)}{k \cdot T}} \right)^{-1} = \left(1 + e^{\frac{W_\Phi - W_s}{k \cdot T}} \right)^{-1}. \quad (2.98)$$

Вероятность противоположного события (т.е. уровень B будет пустым, а адсорбированная частица A будет ионизированной), будет равна

$$f^{(+)} = 1 - f^{(-)} = \left(1 + e^{\frac{W_s - W_\Phi}{k \cdot T}} \right)^{-1}. \quad (2.99)$$

Умножая приведенные вероятности заполнения (2.98) и (2.99) на поверхностную концентрацию N_a поверхностных уровней B , получим поверхностную концентрацию заряженных частиц на поверхности раздела сред:

А) для адсорбированных частиц, создающих акцепторные уровни:

$$N^{(-)} = N_a \cdot \left(1 + e^{\frac{W_\Phi - W_s}{k \cdot T}} \right)^{-1}; \quad (2.100a)$$

Б) для адсорбированных частиц, создающих донорные уровни:

$$N^{(+)} = N_a \cdot \left(1 + e^{\frac{W_s - W_\Phi}{k \cdot T}} \right)^{-1}. \quad (2.100б)$$

Полагая, что каждая заряженная частица (на поверхностном уровне) имеет заряд, численно равный заряду электрона q , можно определить поверхностную плотность заряда σ на границе раздела сред в виде:

А) для адсорбированных частиц, создающих акцепторные уровни:

$$\sigma^{(-)} = -q \cdot N^{(-)} = -q \cdot N_a \cdot \left(1 + e^{\frac{W_\phi - W_s}{k \cdot T}} \right)^{-1}; \quad (2.101a)$$

Б) для адсорбированных частиц, создающих донорные уровни:

$$\sigma^{(+)} = +q \cdot N^{(+)} = +q \cdot N_a \cdot \left(1 + e^{\frac{W_s - W_\phi}{k \cdot T}} \right)^{-1}. \quad (2.101б)$$

Из выражений (2.101a), (2.101б) следует, что в результате хемосорбции атомов на поверхности кристаллической структуры эта поверхность теряет свою электрическую нейтральность и становится электрически заряженной.

Искривление энергетических зон у границы раздела кристаллической структуры

Рассмотренные выше поверхностные заряды $\sigma^{(-)}$ или $\sigma^{(+)}$ притягивают из толщи материала заряды противоположного знака, нейтрализующего поверхностный заряд, образованный в процессе хемосорбции атомов на поверхности. В результате этого приповерхностный слой материала *обедняется* носителями заряда, имеющими один знак со знаком поверхностного заряда, и обогащается носителями заряда противоположного знака. Толщина этого *обедненного носителями заряда* слоя материала находится в обратной зависимости от концентрации основных (для данного типа материала) носителей заряда. В частности, для металлов, у которых концентрация носителей заряда чрезвычайно высока (10^{22} см^{-3}), толщина обедненного слоя не превышает несколько параметров кристаллической решетки. Для материалов полупроводникового типа ($10^6 - 10^{18} \text{ см}^{-3}$) толщина обедненного слоя может быть весьма значительной.

В практическом плане принято считать толщину обедненного слоя материала равной *дебаевской длине экранирования* L_D , т.е. расстоянию, на котором электрическое поле в этом материале (в приближении свободных собственных носителей заряда) уменьшается в e раз. В общем

случае протяженность дебаевской длины экранирования определяется с помощью выражения [5]

$$L_D = \left(\frac{\varepsilon_0 \cdot \varepsilon \cdot k \cdot T}{e^2 \cdot n_i} \right)^{1/2}, \quad (2.102)$$

где n_i – концентрация собственных носителей заряда в веществе; ε – относительная диэлектрическая проницаемость материала кристаллической структуры.

Составить некоторое представление о порядке величин L_D в материалах металлического и полупроводникового типа можно на основании данных, приведенных в табл. 2.17.

Таблица 2.17. *Характерные значения дебаевской длины экранирования L_D для материалов металлического и полупроводникового типа (при $T = 300$ К; $\varepsilon = 5$)*

Параметр	Материал кристаллической структуры				
	Металл	Полупроводник			
$n_i, \text{см}^{-3}$	10^{22}	10^{12}	10^{14}	10^{16}	10^{18}
$L_D, \text{нм}$	≈ 0.1	10^4	10^3	10^2	10

Согласно данным табл. 2.17, в кристаллических структурах полупроводникового типа дебаевская длина экранирования может принимать значения, сопоставимые с пространственной протяженностью наноразмерных структур в направлениях с ограничениями размерного характера.

Имеющие противоположные знаки поверхностный заряд и обедненный слой образуют *двойной электрический слой*, т.е. своеобразный конденсатор, у которого одна из обкладок может иметь протяженный (объемный) характер. Тип носителей заряда (электроны, дырки), образующих обедненный слой в двойном электрическом слое, всецело определяется типом поверхностных состояний (донор, акцептор), образованных в результате хемосорбции атомов на поверхности кристаллической структуры.

На основании данных табл. 2.17 можно сформулировать ряд выводов (с позиции практических интересов наноразмерных структур) о том, что в конструкционных материалах:

а) металлического типа двойной электрический слой локализован непосредственно у границы раздела сред и поэтому его влияние на объемные параметры наноструктуры может проявиться весьма незначительным образом;

б) полупроводникового типа:

- протяженность двойного электрического слоя может существенно превосходить линейный размер наноструктуры в пространственном направлении, на которое распространяется размерное ограничение, поэтому негативное влияние двойного электрического слоя на электрические характеристики наноструктуры может иметь определяющее значение;
- для нейтрализации негативных эффектов, связанных с появлением двойного электрического слоя, необходимо использовать высоковакуумные технологии обработки материалов, исключающие возможность образования монокристаллических адсорбционных покрытий на основе атомов чужеродной природы.

Анализ вида функций напряженности электрического поля $E(x)$ и электрического потенциала $\varphi(x)$ в пределах толщины обедненного слоя можно провести на основании уравнения Пуассона [10]:

$$\frac{d^2 \varphi(x)}{dx^2} = - \frac{q}{\varepsilon_0 \cdot \varepsilon} \cdot \rho(x), \quad (2.103)$$

где $\rho(x)$ – функция объемной плотности распределения электрического заряда в пределах обедненного слоя ($0 < x < x_s$); x – координатное направление по нормали к плоскости раздела физических сред (рис. 2.24).

Не умаляя общности рассмотрения, а в целях упрощения анализа, примем допущения:

- функция плотности распределения электрического заряда $\rho(x)$ идентична функции плотности распределения поверхностных состояний в приповерхностной области пространства и удовлетворяет условию $\rho(x) = \rho_0 = \text{const}$;

- на всем протяжении обедненного слоя x_s , образованного фактом образования поверхностных состояний, указанные состояния «ионизованы», т.е. заняты основными носителями заряда кристаллической структуры (электронами или дырками, в зависимости от типа поверхностных состояний).

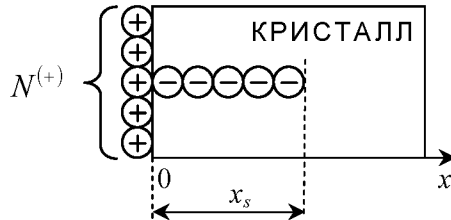


Рис. 2.24. Модельное представление двойного электрического слоя, образованного ионизированными адсорбированными атомами $N^{(+)}$

Для решения уравнения (2.103) воспользуемся граничными условиями:

$$E(0) = -\frac{d\varphi(0)}{dx} = 0; \quad \varphi(0) = 0. \quad (2.104)$$

Решениями уравнения (2.103) при граничных условиях (2.104) являются:

$$E(x) = -\frac{\rho_0}{\varepsilon_0 \cdot \varepsilon} \cdot x, \quad (2.105)$$

$$\varphi(x) = -\frac{\rho_0}{\varepsilon_0 \cdot \varepsilon} \cdot x^2. \quad (2.106)$$

Согласно выражению (2.106), диапазон изменения (по модулю) электрического потенциала $\Delta\varphi_s$ в пределах обедненного слоя $0 < x < x_s$ равен

$$\Delta\varphi_s = \varphi(x_s) = \left| -\frac{\rho_0}{\varepsilon_0 \cdot \varepsilon} \cdot x_s^2 \right|. \quad (2.107)$$

Для рассматриваемого случая объемная концентрация заряда ρ_0 является функцией поверхностной концентрации хемосорбированных

атомов $N^{(+)}$, приводящих к появлению поверхностных состояний, и толщины обедненного слоя x_s :

$$\rho_0 = \frac{q \cdot N^{(+)}}{x_s}. \quad (2.108)$$

Подставляя ρ_0 из выражения (2.108) в выражения (2.106) и (2.107), для потенциальной функции $\varphi(x)$ и поверхностного потенциала $\Delta\varphi_s$ соответственно имеем:

$$\varphi(x) = -\frac{q \cdot N^{(+)}}{\varepsilon_0 \cdot \varepsilon \cdot x_s} \cdot x^2; \quad (2.109)$$

$$\Delta\varphi_s = \left| \frac{q \cdot N^{(+)}}{\varepsilon_0 \cdot \varepsilon} \cdot x_s \right|. \quad (2.110)$$

В табл. 2. 18 приведены значения поверхностного потенциала $\Delta\varphi_s$, соответствующего различным значениям параметров $N^{(+)}$ и x_s .

Таблица 2.18. Зависимость $\Delta\varphi_s$ (В) от параметров $N^{(+)}$ и x_s (при $\varepsilon = 5$)

$N^{(+)}$, м ⁻²	x_s , нм					
	0.5	1.0	5.0	10	50	100
10^{12}	$1.8 \cdot 10^{-6}$	$3.6 \cdot 10^{-6}$	$1.8 \cdot 10^{-5}$	$3.6 \cdot 10^{-5}$	$1.8 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$
10^{13}	$1.8 \cdot 10^{-5}$	$3.6 \cdot 10^{-5}$	$1.8 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$	$1.8 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$
10^{14}	$1.8 \cdot 10^{-4}$	$3.6 \cdot 10^{-4}$	$1.8 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$	$1.8 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$
10^{15}	$1.8 \cdot 10^{-3}$	$3.6 \cdot 10^{-3}$	$1.8 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$	0.18	0.36
10^{16}	$1.8 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$	0.18	0.36	1.8	3.6
10^{17}	0.18	0.36	1.8	3.6	18	36

При больших значениях концентрации поверхностных состояний $N^{(+)}$ (более 10^{15} м⁻²) значения поверхностного потенциала $\Delta\varphi_s$ могут достигать довольно значительных величин (см. табл. 2.18), и в этом рассматриваемом случае объемная концентрация носителей заряда в слое полупроводника x_s может существенно отличаться от объемной концентрации невозмущенного материала. Потеря электрической нейтральности материала в слое x_s находит свое отражение и на виде зон-

ной энергетической диаграммы в диапазоне координат, находящемся в пределах этого слоя (рис. 2.25).

Появление электрически заряженной приповерхностной области x_s может привести к реализации ряда специфических эффектов относительно изменения количества и типа носителей заряда в этой пространственной области полупроводника:

- *обеднение* – возникает в том случае, когда на поверхности полупроводника появляется поверхностный заряд, совпадающий по знаку с основными носителями заряда, причем этот приповерхностный заряд не настолько велик, чтобы привести к пересечению кривой изгиба энергетических зон с уровнем Ферми;
- *инверсия* – случай, аналогичный приведенному выше, с тем лишь отличием, что при этом величина приповерхностного заряда становится настолько большой, что приводит к смене типа носителей заряда в этой области на противоположный тип ввиду пересечения кривой изгиба энергетических зон с уровнем Ферми материала;
- *обогащение* – возникает в том случае, когда на поверхности полупроводника появляется поверхностный заряд, противоположный по знаку с основными носителями заряда полупроводника, что приводит к эффекту обогащения приповерхностной области основными носителями заряда.

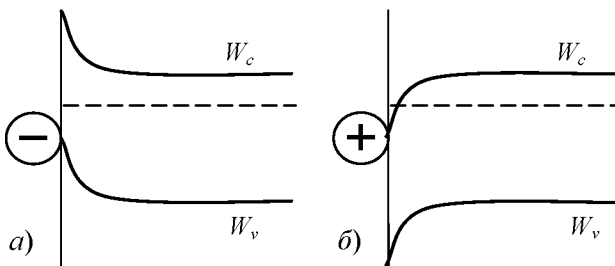


Рис. 2.25. Изгиб энергетических зон у поверхности полупроводника n -типа при зарядении поверхности отрицательными (а) и положительными (б) зарядами

Таким образом, появляющийся в области x_s пространственный заряд может существенно изменить параметры исходного полупроводника относительно его невозмущенного состояния. Это обстоятельство

может иметь весьма существенное значение применительно к кристаллическим структурам наноразмерного масштаба, поскольку в них в пространственных направлениях с размерным ограничением следует ожидать значительные отклонения электропроводности от объемных показателей полупроводника.

г) Специфика баллистического транспорта носителей заряда в наноразмерных структурах

В разд. 2.1.3 рассмотрен вопрос о подвижности носителей заряда для случая макроскопических структур. В частности было отмечено то обстоятельство, что параметр подвижности (пролетного типа) является усредненной характеристикой скорости движения индивидуального носителя заряда в процессе токопрохождения. При этом, прямо или косвенно, предполагалось, что векторизованное движение индивидуального носителя заряда может быть представлено в виде последовательности дискретных элементарных пространственных перемещений на расстояния микроскопического масштаба. Эта дискретность характера перемещения носителя обусловлена элементарными процессами рассеяния векторизованной формы движения на объектах, искажающих фактор периодичности потенциальной энергии кристаллической структуры, к числу которых следует отнести наличие в структуре:

- локальных точечных дефектов в виде инородных атомов примесного характера (атомы внедрения и замещения) и вакансий;
- пространственно протяженных дефектов линейного типа (краевые и винтовые дислокации);
- пространственно протяженных дефектов поверхностного типа (границы кристаллитов, макроскопические поверхности раздела сред);
- локальных тепловых колебаний (фононов) индивидуальных атомов, образующих кристаллическую структуру.

Для макроскопических монокристаллических структур с электронным типом проводимости пространственная протяженность элементарных перемещений (пролетов без рассеяния) характеризуется параметром *длины свободного пробега* электронов, которая, в силу высокого совершенства современных промышленных полупроводниковых монокристаллов, может достигать значений $\lambda \approx 50$ нм. Находясь в условиях

внешнего электрического поля с напряженностью E , электроны проводимости могут приобретать дополнительную энергию

$$\Delta W_E = q \cdot \lambda \cdot E, \quad (2.111)$$

а следовательно, и скорость векторизованного движения, равную в классическом приближении

$$v_E = \sqrt{\frac{2 \cdot q \cdot \lambda \cdot E}{m}}. \quad (2.112)$$

Для реализации этой возможности необходимо, чтобы на зонной энергетической диаграмме имелось достаточно большое количество *свободных* энергетических уровней ΔN (подробнее см. разд. 2.1.3,а; рис. 2.7), находящихся по соседству с энергетическим уровнем, на котором изначально находился ускоряемый электрическим полем электрон:

$$\Delta N = \frac{\Delta W_E}{\Delta W_1},$$

где $\Delta W_1 \approx 10^{-23}$ эВ – среднее расстояние между соседними энергетическими уровнями, расположенными в пределах разрешенной зоны.

В табл. 2.19 представлены значения количества свободных энергетических уровней ΔN , необходимых для *полной реализации* векторизованной формы движения электрона, в зависимости от значений напряженности внешнего ускоряющего электрического поля E .

Т а б л и ц а 2.19. *Зависимость параметра ΔN от E (при $\lambda \approx 10$ нм)*

E , В/м	0.1	1.0	10	100	1000
ΔN	10^{14}	10^{15}	10^{16}	10^{17}	10^{18}

Из данных табл. 2.19 следует, что для полной реализации векторизованной формы движения электрона необходимо значительное количество ($10^{14} - 10^{18}$) свободных энергетических уровней, расположенных выше начального, для обеспечения возможности ускорения электрона; или расположенных ниже – для замедления его движения.

Распространим приведенную выше методологию анализа на случай кристаллических структур, обладающих ограничениями размерного характера. Пусть в пространственном направлении с размерным ограничением атомная цепочка имеет длину $N_{\text{нано}}$. В первом приближении будем считать, что количество энергетических уровней в разрешенной зоне по данному пространственному направлению сопоставимо с $N_{\text{нано}}$, а расстояние между соседними энергетическими уровнями, как и прежде, составляет $\Delta W_1 \approx 10^{-23}$ эВ. В этом случае максимальная энергия векторизованной формы движения ансамбля электронов будет равна

$$\Delta W_E = \Delta W_1 \cdot N_{\text{нано}}, \quad (2.113)$$

а, соответственно, максимальная скорость векторизованного движения электрона равна

$$v_E = \sqrt{\frac{2 \cdot \Delta W_1 \cdot N_{\text{нано}}}{m}} = \sqrt{\frac{2 \cdot \Delta W_1 \cdot L_{\text{нано}}}{m \cdot a}}, \quad (2.114)$$

где a – параметр кристаллической решетки; $L_{\text{нано}} = a \cdot N_{\text{нано}}$ – протяженность кристаллической структуры в пространственном направлении с размерными ограничениями.

На практике может сложиться такая ситуация, что будет выполняться условие $L_{\text{нано}} < \lambda$ (характерного значения для данного материала в макроскопическом исполнении). В этом случае движение носителей заряда будет происходить без протекания процессов рассеяния их энергии на расстоянии $L_{\text{нано}}$. Такой процесс переноса носителей заряда принято называть *баллистическим транспортом* [11]. Следует отметить, что в качестве фактора, лимитирующего скорость векторизованного движения электрона по механизму баллистического транспорта, выступают:

- λ – для случая макроскопических кристаллических тел, не подверженных влиянию пространственных размерных ограничений;

- $L_{\text{нано}}$ (или $N_{\text{нано}}$) – для случая кристаллических структур, имеющих размерные ограничения хотя бы по одному из координатных направлений.

Проведение строгого анализа скорости движения электрона в более корректной постановке вопроса предполагает учет ряда дополнительных немаловажных обстоятельств, а именно:

- зависимости скорости движения электрона $v(k)$, находящегося в разрешенной энергетической зоне, от волнового вектора k с учетом специфики движения на границе зоны, т.е. при значениях $k = \pm(\pi/a)$, (см. разд. 2.2.2,б; рис. 2.19);

- наличия поверхностей раздела физических сред, приводящих к деформации периодического потенциала кристаллической структуры в областях пространства, примыкающих к границам раздела сред (см. рис. 1.7);

- значительных отклонений функций (статистики) распределения электронов по энергиям от равновесных состояний, характерных для случая отсутствия внешнего возмущающего электрического поля;

- наличия в наноразмерной структуре достаточно протяженной (с точки зрения размерных ограничений, свойственных наноструктурам кристаллического типа) приповерхностной области, содержащей большое количество поверхностных состояний, что приводит в конечном итоге к появлению обогащенного или обедненного основными носителями заряда электрически заряженного слоя и к искривлению энергетических уровней в разрешенных зонах (см. рис. 2.25).

Глава 3. ЭНТРОПИЙНЫЕ КРИТЕРИИ КАЧЕСТВА И МЕТОДЫ АНАЛИЗА НАНОТЕХНОЛОГИЙ И НАНОРАЗМЕРНЫХ СТРУКТУР

3.1. Замечания общего характера

Понятие «энтропия» широко используется в современной естественно-научной (физика, химия, молекулярная биология) и технической (вычислительные системы) практике. Используется настолько широко, что в ряде случаев ее ставят по значимости наравне с понятием «энергия». По своей смысловой сути это понятие характеризует меру беспорядка (хаоса), присутствующего в системе. В этой связи *уменьшение энтропии системы принято рассматривать как привнесение в нее определенного порядка в части упорядочения ее атомно-молекулярного состава и структуры.*

Поскольку любая инженерно-техническая деятельность непосредственно связана с процессами упорядочивания разнообразных материальных сред (с целью достижения определенного функционального качества изделия), то становится ясным, что понятие «энтропия» имеет самое непосредственное отношение к производственно-технической сфере. Это утверждение особенно актуально в тех случаях, когда речь идет об изделиях микро- и нанотехнологий, для которых характерна чрезвычайно высокая степень упорядоченности используемых материальных сред и структур.

Следует, однако, отметить, что несмотря на всю очевидность приведенных выше моментов, методы и подходы анализа технических объектов, основанные на энтропийных представлениях, не нашли широкого (адекватного их потенциальным возможностям) распространения в современной инженерной практике. Здесь в качестве едва ли не единственного примера успешного и плодотворного их использования в инженерной практике можно привести достижения при решении ряда прикладных задач в области теории связи (повышение помехоустойчивости сигналов). Причина подобного негативного положения в том, что для успешного использования энтропийных методов анализа в современной конструкторско-технологической практике необходимо ис-

пользовать комплексные подходы, вбирающие в себя основные положения таких дисциплин, как «теория информации», «теория измерительных процессов», «физика твердого тела», «неравновесная термодинамика». В то же время следует отметить, что указанные выше научные дисциплины представлены в современных учебных планах подготовки конструкторов и технологов различной отраслевой направленности в весьма скромных объемах или в специфическом виде, исключающем возможность комплексного подхода.

Настоящее пособие не ставит своей целью устранение пробела в преподавании указанных выше научных дисциплин, поскольку эту задачу невозможно решить в рамках одной учебной дисциплины. Тем не менее ряд основных положений этих дисциплин, необходимых для понимания общей методологии практического использования энтропийных подходов в конструкторско-технологической практике, нашли свое отражение в настоящем разделе в объемах, достаточных для формирования целостного представления о предмете рассмотрения. Последнее обстоятельство наложило свой отпечаток на общую структуру изложения материала раздела.

Междисциплинарный характер рассматриваемой проблематики создает известные трудности и в части выбора оптимальной схемы подачи материала, рубрикации разделов и структуры изложения материалов в подразделах. Ряд материалов имеет «пограничный» характер и их включение в тот или иной раздел обусловлено общностью либо методологического плана, либо – структурного.

В основной части первого подраздела рассматриваются (без приведения доказательств) общие положения теории информации, относящиеся к методам определения численных характеристик энтропийных показателей сложных структур, рассматриваемых как последовательности определенных случайных величин. Этим вопросам посвящено большое количество литературных источников (в рамках настоящей работы использована методология изложения материала, представленная в [1,2]). Вводятся понятия случайной и усредненной энтропии ансамбля случайных величин (объектов) и энтропии единичного размещения случайного события в последовательности.

Во втором подразделе приведены ключевые понятия теории информационной (энтропийной) емкости измерительных процессов [3]. Этот круг материалов необходим для последующей оценки информационной емкости и информационного дефицита технологических процессов.

Материалы третьего подраздела имеют вспомогательный характер и посвящены рассмотрению вопроса использования статистических методов анализа при исследовании кинетики индивидуальных физико-химических процессов в системах, в которых протекают процессы типа «рождение-гибель» [4]. Этот круг рассмотренных в разделе вопросов является конструктивной основой при определении энтропийных показателей твердотельных структур и технологий их создания.

В четвертом подразделе представлены вопросы, связанные с анализом термодинамической энтропии закрытых и открытых физико-химических систем (в рамках настоящей работы использована методология изложения материала, представленная в [5]). Показано, что термодинамическая энтропия физической системы является функцией состояния этой системы и тем самым однозначно характеризует ее текущее состояние. Представлена методология термодинамического анализа равновесных систем. В рамках линейной неравновесной термодинамики рассмотрены вопрос о скорости производства термодинамической энтропии в неравновесных структурах и зависимость ее интенсивности от градиентов интенсивных термодинамических параметров. Этот подход представляет определенный практический интерес при проведении анализа процессов деградации изделий (в ходе их эксплуатации) с позиций макрофизической кинетики [6].

Представленные в первых четырех подразделах материалы имеют вспомогательный и ознакомительный характер и поэтому изложены в реферативной форме и в объеме, необходимом для понимания общей методологии проведения энтропийного анализа в различных аспектах конструкторско-технологической деятельности.

С этой точки зрения, учитывая заявленные в названии раздела цели, пятый подраздел следует считать основным. Методологической основой этого подраздела является математический аппарат теории инфор-

мации [7]. В качестве первого шага в направлении использования энтропийных методов анализа применительно к инженерно-технической практике проведена формализация реальных физических структур (изделий), что позволило сблизить, в части достижения единообразия методов анализа, теоретические подходы теории информации и методы анализа, используемые конструкторами и технологами при разработке и производстве изделий [8]. Рассмотрен широкий комплекс энтропийных показателей качества (на микроскопическом и макроскопическом уровнях) технологических процессов создания изделий [9]. Представлена взаимосвязь энтропийных показателей качества технологических процессов с производственными показателями, что позволяет распространить результаты теоретических анализов в сферу практических интересов [10]. Определены значения энтропийных показателей качества технологических процессов в различных областях инженерной деятельности и проведено их сравнение с аналогичными показателями биологических объектов [11]. Показано, что для случая создания изделий наноскопического размера появляются дополнительные трудности, связанные с необходимостью реализации вырожденных элементов конструкций (количество реализаций близко к единице), что находится в противоречии с известными принципиальными физическими ограничениями флуктуационной природы. Дополнительно рассмотрены вопросы деградации изделий, с привлечением аппарата неравновесной термодинамики [12].

Материалы раздела адресованы широкому кругу инженерно-технических работников, связанных по роду своей деятельности с вопросами разработки и создания микро- и наноразмерных устройств различной отраслевой направленности. Именно в этой связи в настоящий раздел включен ряд дополнительных материалов (основы теории информации, информационная теория измерительных процессов, равновесная и неравновесная термодинамика, кинетика процессов типа «рождение-гибель»), способствующих восприятию рассматриваемых в учебном пособии энтропийных подходов, используемых в дальнейшем при проведении анализа нанотехнологий и наноразмерных систем.

3.2. Информационная энтропия

3.2.1. Элементы теории информации

а) Определение энтропии для случая равновероятных возможностей (дискретный случай)

Термин «информация», указанный в заглавии подраздела, понимается здесь не в том широком смысле, в каком его понимают работники средств массовой информации, а в узком научном смысле, какой ему придал основоположник математической теории информации К. Шеннон [7].

Главным свойством случайных событий, изучение которых составляет один из основных предметов математической теории информации, является отсутствие полной уверенности в их наступлении, создающее известную неопределенность при выполнении связанных с этими событиями опытов. Однако совершенно ясно, что степень этой неопределенности в различных случаях будет совершенно разной. *Для инженерной практики важно уметь численно оценивать степень неопределенности самых разнообразных опытов (ситуаций), чтобы иметь возможность сравнить их с этой стороны. Именно эта практическая задача составляет предмет теории информации [1].*

Рассмотрим основополагающие представления теории информации. Пусть имеется M равновероятных возможностей реализации (исходов) некоторого опыта (например, при бросании правильной игральной кости $M = 6$). Конечно, в практической инженерной деятельности, ввиду комплексного влияния множества возмущающих факторов различной физической природы, формализация условий проведения опыта достигается не так просто и однозначно, как в приведенном выше случае игральных карт [13]. Тем не менее предполагается, что оговоренная формализация успешно проведена и что действительно реализуется одна из M равновероятных возможностей. Предполагается, что различные реализации составляют полную группу событий, а следовательно, сумма вероятностей всех возможных реализаций удовлетворяет условию нормировки, т.е. равна единице. Представляется очевидным (даже на интуитивном уровне рассмотрения), что априорная неопределен-

ность результата опыта самым непосредственным образом связана со значением параметра M , и в частности, чем больше M , тем больше ожидаемая неопределенность результата опыта (при $M = 1$ процесс имеет детерминированный характер и априорная неопределенность исхода опыта заведомо равна нулю).

Таким образом, численная мера степени априорной неопределенности H опыта может быть представлена в виде

$$H = f(M), \quad (3.1)$$

где $f(M)$ – некоторая возрастающая от параметра M неотрицательная функция, определенная, по меньшей мере, для множества чисел натурального ряда.

Для более полного определения вида функции $f(M)$ надо предъявить к ней дополнительные требования. Рассмотрим два *независимых* опыта « α » и « β » (т.е. такие два опыта, что любые сведения об исходе первого из них никак не отражаются на вероятности появления исходов второго опыта). Пусть опыт « α » имеет k равновероятных исходов, а опыт « β » – l равновероятных исходов. Рассмотрим далее более сложный опыт « $\alpha\beta$ », суть которого состоит в одновременном выполнении опытов « α » и « β ». Представляется очевидным, что неопределенность исхода этого сложного опыта « $\alpha\beta$ » будет больше неопределенности опыта « α », так как к неопределенности опыта « α » добавляется еще неопределенность исхода опыта « β ». Вполне естественно принять, что степень неопределенности опыта « $\alpha\beta$ » равна сумме неопределенностей, характеризующих отдельные опыты « α » и « β ». Так как опыт « $\alpha\beta$ » имеет $(k \cdot l)$ равновероятных комбинаций исходов опыта (они получаются, если комбинировать каждый из k возможных исходов опыта « α » с каждым из l исходов опыта « β »), то дополнительное требование, которому должна удовлетворять функция $f(M)$, формулируется в виде условия, формируемого с учетом выражения (3.1):

$$f(kl) = f(k) + f(l). \quad (3.2)$$

Условие в виде выражения (3.2) наводит на мысль принять за меру неопределенности опыта, имеющего k равновероятных исходов, число $\log k$, поскольку в этом случае $\log(kl) = \log(k) + \log(l)$ (можно пока-

зять, что логарифмическая функция является единственной функцией, удовлетворяющей условию (3.2) [7]). Принятое определение меры неопределенности согласуется также с условиями, что при $k = 1$ она равна нулю и при возрастании k она также возрастает.

Численная величина $H = \log M$ получила название *энтропии события с равновероятными исходами* и характеризует *априорную* неопределенность опыта.

Количество пришедшей (в результате проведения опыта) информации I естественно измерять величиной исчезнувшей неопределенности:

$$I = H_{pr} - H_{ps}, \quad (3.3)$$

где индекс «*pr*» означает – «априори» (т.е. до проведения опыта), а индекс «*ps*» – «апостериори» (т.е. после проведения опыта).

Наличие фрагмента H_{ps} в выражении (3.3) отражает то обстоятельство, что в ряде случаев и после проведения опыта остается некоторая доля неопределенности (подробнее см. ниже) состояния рассматриваемой системы (объекта).

Заметим, что выбор основания логарифма при определении энтропии H несуществен, так как в силу известной формулы $\log_b k = \log_b a \cdot \log_a k$ переход от одной системы логарифмов к другой сводится к умножению функции $f(k) = \log(k)$ на постоянный множитель (модуль перехода). В настоящее время в практических приложениях используются три единицы измерения энтропии и информации:

- 1) «*бит*» – сокращение от «binary digit», основанием логарифма является число 2 (широко используется в вычислительной технике);
- 2) «*нат*» – сокращение от «natural digit», основанием логарифма является число e – натуральный логарифм (широко используется в физико-химических приложениях);
- 3) «*дит*», основанием логарифма является число 10 (используется в настоящей работе, исходя из удобства восприятия результатов количественных оценок).

Энтропию опыта при равновероятных исходах можно выразить не только путем использования количества *равновероятных исходов* k , но и значения вероятности индивидуального исхода $P = 1/M$. В этом слу-

чае можно выразить значение энтропии опыта через вероятность исхода опыта:

$$H = -\ln P. \quad (3.4)$$

В инженерно-физической практике при определении энтропии физических систем (процессов) часто используется «термодинамическая шкала» значений энтропии и вычисление энтропии проводится согласно выражению

$$H = k \cdot \ln M, \quad (3.5)$$

где $k = 1.38 \cdot 10^{-23}$ Дж/град – постоянная Больцмана.

Рассмотренное выше определение энтропии относится к самому простейшему случаю – опыту с равновероятными исходами. Ниже эти результаты будут распространены на более сложный и практически значимый случай неравновероятных исходов опыта.

*б) Энтропия для случая неравновероятных возможностей
(дискретный случай)*

Рассмотрим более сложный опыт, в котором вероятности различных реализаций (возможных исходов) неодинаковы. Если, как это было ранее, число различных исходов опыта равно M , то можно ввести в рассмотрение случайную величину ξ , принимающую одно из M значений. Приняв в качестве ξ номер исхода опыта, получим, что эти значения равны числовой последовательности $1, \dots, M$. По определению, вероятности индивидуальных исходов $P(\xi)$ являются положительными величинами и должны подчиняться условию нормировки вида

$$\sum_{\xi} P(\xi) = 1. \quad (3.6)$$

Если применить формальным образом выражение (3.4) к рассматриваемому случаю неравновероятных исходов, то получим, что каждому значению ξ (т.е. конкретному исходу) соответствует свое значение энтропии:

$$H(\xi) = -\ln P(\xi). \quad (3.7)$$

Таким образом, каждому индивидуальному исходу опыта ставится в соответствие свое значение энтропии. Поскольку ξ – случайная величина

на, то и эту совокупность индивидуальных энтропий $H(\xi)$ можно рассматривать как случайную величину, частота проявления которой определяется вероятностью $P(\xi)$. Из выражения (3.7) следует, что индивидуальная энтропия велика, когда априорная вероятность данной реализации мала, и наоборот. Этот факт вполне соответствует интуитивным представлениям, связанным с фактором неожиданности результата опыта. Если апостериорная энтропия, т.е. после проведения опыта, равна нулю, то, согласно выражению (3.3), информация, получаемая при выявлении реализации, численно равна первоначальной энтропии:

$$I(\xi) = H(\xi) = -\ln P(\xi). \quad (3.8)$$

Информация о конкретной реализации, как и рассмотренная выше энтропия, оказывается зависящей от типа реализации, т.е. тоже является случайной величиной. Энтропия в виде (3.7) получила название *случайной энтропии* и характеризует степень неопределенности появления конкретной реализации (исхода) опыта.

В практических приложениях нашла широкое распространение не случайная энтропия по выражению (3.7) (соответственно информация по (3.8)), а усредненная по всему ансамблю возможных реализаций опыта, определяемая выражением [1,7]

$$H_{\xi} = -\sum_{\xi} P(\xi) \cdot \ln P(\xi). \quad (3.9)$$

Случайная величина ξ , стоящая в индексе символа H_{ξ} (в отличие от $H(\xi)$), отражает тот факт, что в данном случае рассматривается *усредненная энтропия*, которая описывает весь комплекс возможных реализаций в целом. Такая система обозначений в расширенном виде будет использована и в дальнейшем.

Неопределенность типа $0 \cdot \ln 0$, появляющаяся в выражении (3.9), когда вероятности отдельных исходов равны нулю, всегда понимается в смысле $0 \cdot \ln 0 = 0$. Ввиду этого допущения множество из M возможностей всегда можно дополнить любым количеством возможностей с нулевой вероятностью, а также в индексы энтропии вписать детерминированные (неслучайные) величины.

Следует отметить, что различие между значениями усредненной энтропии и отдельных представителей случайной энтропии (из полного представительства значений последней) может достигать значительных величин, в чем можно убедиться непосредственным образом на основании данных табл. 3.1 (каждая строка характеризует один определенный опыт).

Т а б л и ц а 3.1. *Примеры значений случайной и усредненной энтропии для ряда опытов (значения энтропии приведены в битах)*

Вероятность индивидуального события / случайная энтропия события P_i / H_i					Усредненная энтропия H_{ξ}
0.001/10	0.999/0.001	-	-	2 исхода	0.011
0.01/6.6	0.99/0.01	-	-	2 исхода	0.081
0.1/3.2	0.9/0.15	-	-	2 исхода	0.469
0.5/1.0	0.5/1.0	-	-	2 исхода	1.0
0.25/2.0	0.25/2.0	0.25/2.0	0.25/2.0	4 исхода	2.0
0.2/2.3	0.2/2.3	0.2/2.3	0.2/2.3	0.2/2.3	2.322
0.01/6.6	0.01/6.4	0.01/6.4	0.01/6.4	0.96/0.06	0.322
1.0/0.0	-	-	-	1 исход	0.0

Данные таблицы свидетельствуют о следующем:

1) для случаев маловероятных исходов опыта случайная энтропия может иметь весьма большие значения, что вполне согласуется с интуитивными представлениями, связанными со степенью ожидания конкретного результата опыта;

2) диапазон возможных вариаций случайной энтропии (отношение H_{\max} / H_{\min}) при различных вариантах исходов опыта может достигать весьма больших значений (в частности 10^4);

3) усредненная энтропия может существенно различаться (как в большую, так и в меньшую сторону) от случайной энтропии отдельных реализаций опыта.

Для дальнейшего рассмотрения ниже представлен ряд теорем (без доказательств), определяющих свойства усредненной энтропии по (3.9), необходимые для проведения в дальнейшем анализа энтропии технических систем [1]:

Теорема №1. Как случайная (3.7), так и средняя (3.9) энтропия всегда неотрицательна.

Поскольку $P(\xi) \leq 1$ (по определению), то $\ln P(\xi) = H(\xi) \geq 0$. Последнее неравенство сохраняется и после усреднения.

Теорема №2. Усредненная энтропия имеет максимальное значение, равное $\ln M$, когда все возможные реализации опыта равновероятны, т. е. когда $P(\xi) = 1/M$.

Теорема №3. Если несколько случайных величин $\xi_1, \xi_2, \dots, \xi_n$ являются независимыми друг от друга, то полная совместная энтропия $H_{\xi_1, \xi_2, \dots, \xi_n}$ распадается на сумму энтропий индивидуальных случайных величин:

$$H_{\xi_1 \dots \xi_n} = \sum_{j=1}^n H_{\xi_j} . \quad (3.10)$$

Свойство, отмеченное в теореме №3, весьма примечательно в той части, что является проявлением *принципа аддитивности* при проведении независимых испытаний, который и ранее был использован и привел к логарифмической функции энтропии. Факт аддитивности энтропии при независимых испытаниях весьма широко используется на практике, поскольку это обстоятельство существенно упрощает анализ сложных процессов и структур. Например, если речь идет о технологическом процессе атомной сборки наноразмерных устройств (и при этом размещение индивидуальных атомов можно считать статистически независимыми событиями), то суммарная энтропия технологического процесса будет равна сумме энтропий элементарных технологических операций по размещению индивидуальных атомов в пределах наноразмерной структуры. Более того, как будет показано ниже, строгое требование о статистической независимости может быть ослаблено (и заведомо перекрыто) всего лишь требованием об *энтропийной устойчивости* (см. ниже) технологического процесса, что существенно расширяет возможный круг рассматриваемых процессов и структур.

*в) Энтропия для случая неравновероятных возможностей
(непрерывный случай)*

В рассмотренных выше разделах предполагалось, что случайная величина ξ с энтропийными параметрами $H(\xi)$ и H_ξ может принимать значения из некоторого ограниченного дискретного пространства, включающего в себя конечное или счетное число различных элементов. В то же время в технической практике нашли широкое распространение также непрерывные величины, т.е. скалярные или векторные величины, которые могут принимать значения из непрерывного пространства X , например, пространства действительных чисел. В этом случае в качестве основной характеристики случайной величины ξ выступает плотность распределения вероятностей $p(\xi)$, определяющая вероятность попадания ΔP случайной величины ξ в область ΔX указанного пространства X :

$$\Delta P = \int_{\xi \in \Delta X} p(\xi) \cdot d\xi = p(A) \cdot \Delta V, \quad (3.11)$$

где $d\xi = dV$ (дифференциал объема в пространстве состояний случайной величины); $A \in \Delta X$.

Формальным методом определения энтропии (для случая непрерывной случайной величины) является использование выражения

$$H_\xi = -M \cdot \ln P(\xi), \quad (3.12)$$

где M – оператор математического ожидания для случая дискретной случайной величины ξ по параметру $\ln P(\xi)$.

Суть формализма состоит в том, что вероятности $P(\xi)$ в выражении (3.9) формальным образом заменены на соответствующие плотности распределения вероятностей $p(\xi)$ – непрерывной функции в области допустимых значений X случайной величины:

$$H_\xi = -M \cdot \ln p(\xi) = - \int_X p(\xi) \cdot \ln p(\xi) \cdot d\xi. \quad (3.13)$$

Такой подход является общепринятым, но создает известные трудности в случае комплексного (одновременного) проявления непрерывных и дискретных факторов случайного процесса [1].

г) *Асимптотическая равновероятность и энтропийная устойчивость последовательности случайных событий*

Рассмотрим сложное событие, состоящее в последовательном проведении n независимых опытов [14]. В результате получается набор реализаций $\eta = (\xi_1, \dots, \xi_n)$ случайной величины ξ_j , принимающей одно из двух значений 1 или 0 с вероятностями соответственно $p < 1/2$ и $q = 1 - p$. Очевидно, число таких различных наборов (реализаций) равно 2^n . Обозначим посредством η_{n_1} реализацию, содержащую в своем составе n_1 единиц и $n_0 = n - n_1$ нулей. Вероятность этой реализации равна (согласно закону умножения вероятностей независимых событий):

$$P(\eta_{n_1}) = p^{n_1} \cdot q^{n-n_1}. \quad (3.14)$$

Эти вероятности $P(\eta_{n_1})$ при различных значениях n_1 будут различными. Отношение самой большой вероятности к самой малой, определяемое выражением

$$P(n_0) / P(n_n) = (q/p)^n, \quad (3.15)$$

велико и сильно возрастает с ростом n . В этой связи возникает вопрос, о какой равной вероятности (оговоренной в названии раздела) может идти речь? Суть дела в том, что по закону больших чисел количество единиц n_1 испытывает тенденцию принимать значения, близкие к своему среднему значению:

$$Mn_1 = \sum_{j=1}^n M\xi_j = n \cdot M\xi_j = np. \quad (3.16)$$

Определим дисперсию $Dn_1 = D\{\xi_1 + \dots + \xi_n\}$ числа единиц. В силу независимости слагаемых имеем

$$Dn_1 = n \cdot D\xi = n \cdot [M\xi^2 - (M\xi)^2] = npq. \quad (3.17)$$

Интересно отметить, что среднее отклонение, равное

$$\Delta n_1 = n_1 - np \approx \sqrt{npq}, \quad (3.18)$$

растет с ростом n , но медленнее, чем растет среднее значение произведения np и длина всего возможного диапазона $0 < n_1 < n$.

В пределах диапазона $|n_1 - np| \approx \sqrt{npq}$ неодинаковость вероятностей $P(\eta_{n_1})$ довольно велика

$$\frac{P(\eta_{n_1})}{P(\eta_{n_1 - \Delta n_1})} = \left(\frac{q}{p}\right)^{\Delta n_1} \approx \left(\frac{q}{p}\right)^{\sqrt{npq}} \quad (3.19)$$

и увеличивается с ростом n . Но это увеличение происходит относительно медленно, т.е. гораздо медленнее, чем увеличивается обратная величина самой вероятности. Соответствующее этому утверждению неравенство имеет вид

$$\ln \frac{P(\eta_{n_1})}{P(\eta_{n_1 - \Delta n_1})} \leq \ln \frac{1}{P(\eta_{n_1})}, \quad (3.20)$$

при значении $\Delta n_1 \approx \sqrt{npq}$ становится все более сильным с ростом n .

Для дальнейшего рассмотрения ниже представлен ряд теорем (без доказательств), позволяющих сформировать конструктивный подход при анализе реальных последовательностей случайных величин (нумерация продолжена, см. пункт (б)) [1]:

Теорема № 4. Все 2^n реализаций η можно разделить на два множества A_n и B_n так, что:

1) суммарная вероятность одного множества (A_n) исчезает:

$$P(A_n) \rightarrow 0 \text{ при } n \rightarrow \infty; \quad (3.21)$$

2) реализации второго множества B_n становятся относительно равновероятными в следующем смысле:

$$\left| \frac{\ln P(\eta) - \ln P(\eta')}{\ln P(\eta)} \right| \rightarrow 0, \quad \eta \in B_n, \eta' \in B_n. \quad (3.22)$$

Теорема №5. Пусть B_n – множество, описанное в теореме №4. Число M его элементов таково:

$$\frac{\ln M}{n} \rightarrow -(p \ln p + q \ln q) \equiv H_{\xi_1} \text{ при } n \rightarrow \infty. \quad (3.23)$$

Теорема №6. Если считать реализации множества A_n имеющими нулевую вероятность их появления, а реализации множества B_n равновероятными и вычислять энтропию H_η по простой формуле $H_\eta = \ln M$, то удельная энтропия H_η / n в пределе будет совпадать с энтропией, вычисленной по формуле (3.9), т.е.

$$\frac{\ln M}{n} \rightarrow -(p \ln p + q \ln q) \equiv H_{\xi_1} \text{ при } n \rightarrow \infty. \quad (3.24)$$

В случае, когда ξ_1 принимает одно из нескольких (m) значений, то имеется m^n различных реализаций процесса $\eta = (\xi_1, \dots, \xi_n)$ с независимыми значениями, из которых заслуживают внимания лишь $M = e^{nH_{\xi_1}}$ реализаций, которые можно считать равновероятными. Следовательно, подавляющее большинство реализаций при этом являются несущественными и ими можно пренебречь.

Представленные выше моменты, касающиеся *асимптотической эквивалентности* неравновероятных возможностей равновероятным, могут быть распространены на значительно более широкий класс случайных последовательностей и процессов, для которых:

- обязательно, чтобы случайные величины ξ_j , образующие отрезок последовательности $\eta = (\xi_1, \dots, \xi_n)$, принимали одно из двух значений;
- обязательно, чтобы они имели одинаковый статистический закон распределения исходов событий $P(\xi_j)$;
- обязательно, наконец, чтобы исходы событий были статистически независимыми;
- обязательно даже, чтобы η была последовательностью $\eta = (\xi_1, \dots, \xi_n)$.

Что же в таком случае является обязательным для описанной асимптотической эквивалентности?

Дать общую формулировку свойства асимптотической эквивалентности неравновероятных возможностей равновероятным возможностям

помогает понятие *энтропийной устойчивости* семейства случайных величин. Семейство случайных величин $\{\eta^n\}$ является энтропийно устойчивым, если отношение $H(\eta^n)/H_{\eta^n}$ при $n \rightarrow \infty$ сходится по вероятности к единице (где $H(\eta^n)$ – случайная энтропия единичной последовательности; H_{η^n} – средняя энтропия по ансамблю последовательностей). Это значит, что каковы бы ни были $\varepsilon > 0$, $\eta > 0$, найдется такое $N(\varepsilon, \eta)$, что будет выполнено неравенство

$$P\left\{\left|H(\eta^n)/H_{\eta^n} - 1\right| \geq \varepsilon\right\} \leq \eta \quad (3.25)$$

при любом $n \geq N(\varepsilon, \eta)$.

В приведенном выше определении подразумевается, что H_{η^n} не убывает с ростом n (обычно $H_{\eta^n} \rightarrow \infty$). Факт асимптотической равновероятности можно сформулировать при помощи понятия энтропийной устойчивости в виде следующей теоремы.

Теорема №7. Если семейство случайных величин $\{\eta^n\}$ является энтропийно устойчивым, то множество реализаций каждой случайной величины можно разбить на два подмножества A_n и B_n следующим образом:

1) суммарная вероятность реализации подмножества A_n становится исчезающей малой величиной:

$$P(A_n) \rightarrow 0 \quad \text{при } n \rightarrow \infty; \quad (3.26)$$

2) реализации второго подмножества B_n становятся равновероятными в смысле выполнения соотношения

$$\left| \frac{\ln P(\eta) - \ln P(\eta')}{\ln P(\eta)} \right| \rightarrow 0, \quad (3.27)$$

при $n \rightarrow \infty$, $\eta \in B_n$, $\eta' \in B_n$;

3) число реализаций M_n множества V_n связано с энтропией H_{η^n} соотношением

$$\ln \left(\frac{M_n}{H_{\eta^n}} \right) \rightarrow 1 \text{ при } n \rightarrow \infty. \quad (3.28)$$

Фактор энтропийной устойчивости семейства случайных последовательностей $\{\eta^n\}$ можно проверить, вычисляя дисперсию последовательности согласно выражению

$$DH(\eta^n) = MH^2(\eta^n) - H_{\eta^n}^2. \quad (3.29)$$

Если эта дисперсия достаточно медленно растет с ростом n , то случайная последовательность будет энтропийно устойчивой.

Рассмотренные выше моменты будут в дальнейшем использованы при определении энтропийных показателей технологических процессов, используемых при создании микро- и наноразмерных структур [9,10].

3.2.2. Основы информационной теории измерительных процессов

а) Применение основных положений теории информации для измерительных операций

Под измерением на практике понимается процесс получения человеком или машиной информации о количественном значении измеряемой величины. Однако дать точное словесное определение измерения оказывается далеко не просто. В настоящее время принято сущность процесса измерения определять следующим образом: «Измерением называется познавательный процесс, заключающийся в сравнении путем физического эксперимента данной величины с некоторым ее значением, принятым за единицу сравнения», т.е. эталоном для данного физико-технического параметра [3].

Для установления системы исходных положений теории измерений, собственно, как и любой другой научной теории, необходим аксиоматический выбор некоторых логических начал. Такие начала, как

известно, составляют систему исходных понятий, которые приходится принять без каких-либо доказательств, после чего последующие положения теории обосновываются уже логически. Истинность этих понятий не может быть логически обоснована сама по себе, а устанавливается лишь опытной проверкой всей системы результатов построенной на их основе теории. Для понимания сущности измерения исходными понятиями являются понятия *однородных величин, натурального ряда однородных величин, шкалы реперов, функциональной шкалы и измерительного преобразования.*

Познание, приобретенное человеком в результате повторяющихся практических наблюдений, убеждает в том, что различные предметы могут сравниваться между собой лишь по однородным свойствам. При этом *под однородными свойствами и характеризующими их величинами понимаются лишь такие, которые могут быть сопоставлены между собой по простому признаку: «больше — меньше».*

Сопоставляя информативные параметры (характерные для различных объектов) друг с другом, можно обнаружить некоторые признаки их отличия, например: данный предмет тяжелее или легче другого, тверже или мягче, теплее или холоднее, длиннее или короче и т. п. Разложим объекты в ряд так, чтобы каждый последующий был тверже предыдущего, или таким же образом, но по какому-либо другому признаку, относительно которого такое разложение взятых предметов представится возможным. *Ряды, составленные по однородным свойствам различных предметов, называются последовательными натуральными рядами.* Однородными называются сравнимые по признаку «больше — меньше» величины, которые могут быть расположены в последовательный ряд, составленный так, что каждая из входящих в него величин будет больше всех предыдущих и меньше всех последующих.

Если под измерением будем понимать лишь сравнение измеряемой величины с некоторым ее значением, принятым за единицу, то измерения по натуральным шкалам, в которых еще нет постоянной единицы измерения (например, определение чувствительности фотоматериалов), придется рассматривать как что-то «недостойное» называться измерениями. Однако в действительности эта оценка не так уж и груба.

(К примеру, если в результате опыта стало известно, что измеряемая температура была такова, что серебряный образец уже расплавился, а золотой еще оставался твердым, то, следовательно, температура находилась в интервале от $960,5$ до 1063°C . Этот результат можно записать как $t = 1010 \pm 50^\circ \text{C}$, а оставшуюся после измерения неопределенность характеризовать относительной погрешностью, равной $\pm 5\%$.)

Таким образом, информация, содержащаяся в результате измерения по натуральным шкалам, ничем не хуже той, которая может быть получена от других измерительных устройств с той же погрешностью измерения. Основным недостатком натуральных шкал является полное отсутствие уверенности в том, что интервалы между выбранными реперами хотя бы в какой-то мере являются равновеликими, т.е. отсутствие пропорционального деления (с понятием натуральных шкал еще не связано представление о пропорциональном делении величины). Именно таково положение с минералогической шкалой твердости.

Точность измерений обычно характеризуется числовым значением полученных при измерении или предполагаемых погрешностей. При этом используются понятия *абсолютной* и *относительной* приведенной погрешностей. Если измерительное устройство имеет диапазон измерения от X_1 до X_2 , т.е. может измерять величины, находящиеся в пределах от X_1 до X_2 , с абсолютной погрешностью $\pm \Delta$, не зависящей от текущего значения x измеряемой величины, то, получив результат измерения в виде показания $X_{\text{п}}$, записывают его как $X_{\text{п}} \pm \Delta$ и характеризуют относительной приведенной погрешностью:

$$\gamma = \pm \frac{\Delta}{X_2 - X_1}. \quad (3.30)$$

Рассмотрение этих же самых действий с позиций теории информации носит несколько иной характер, отличающийся тем, что всем перечисленным понятиям придается вероятностный, т.е. *статистический смысл*, а итог проведенного измерения истолковывается как сокращение области неопределенности измеряемой величины. В теории информации тот факт, что измерительный прибор имеет диапазон измерений от X_1 до X_2 , означает, что при использовании этого прибора могут быть получены показания $X_{\text{п}}$ только в пределах от X_1 до X_2 . Другими

словами, вероятность получения отсчетов, меньших X_1 и больших X_2 , равна нулю, а вероятность получения отсчета где-то в пределах от X_1 до X_2 равна единице.

Если предположить, что плотность вероятности распределения различных значений измеряемой величины вдоль всей шкалы прибора одинакова, то с точки зрения теории информации наше знание о значении измеряемой величины до измерения может быть представлено графиком распределения плотности вероятности $p(x)$ вдоль шкалы значений x , показанным на рис. 3.1.

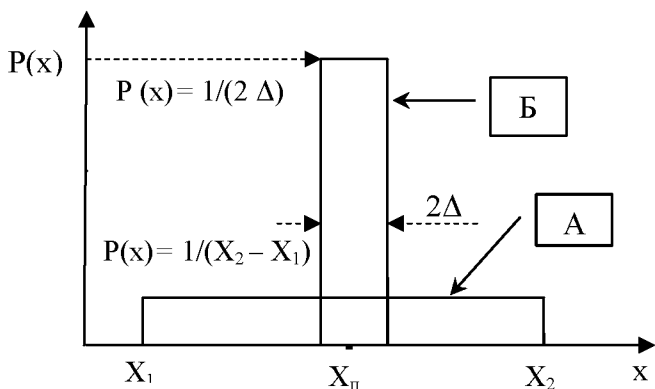


Рис. 3.1. Плотность распределения измеряемой величины до (А) и после (Б) проведения измерения

Так как полная вероятность получить отчет где-то в пределах от X_1 до X_2 равна единице, то под кривой $p(x)$ должна быть заключена площадь, равная единице (условие нормировки).

При равномерном распределении плотности вероятности измерительного процесса до проведения измерения вероятность определенного отсчета составляет

$$p(x) = \frac{1}{X_2 - X_1}. \quad (3.31)$$

После проведения измерения получаем показание прибора, равное $X_{п}$. Однако вследствие погрешности прибора, равной $\pm \Delta$, нельзя утверждать, что измеряемая величина точно равна значению $X_{п}$. Поэтому результат измерения представляется в виде $X_{п} \pm \Delta$. Это означает, что

действительное значение измеряемой величины X лежит где-то в пределах от $(X_n + \Delta)$ до $(X_n - \Delta)$, т. е. в пределах участка 2Δ , как показано на рис. 3.1.

С точки зрения теории информации результат измерения состоит лишь в том, что до измерения область неопределенности простиралась от X_1 до X_2 и характеризовалась малой плотностью вероятности $p_{\text{апр}}(x) = 1/(X_2 - X_1)$, а после измерения она сократилась до величины 2Δ и характеризуется намного большей плотностью вероятности $p_{\text{апост}}(x) = 1/2\Delta$.

Таким образом, получение какой-либо информации об интересующей нас величине заключается в уменьшении неопределенности ее значения. Теория информации облекает это логически понятное, но довольно расплывчатое рассуждение в строгую математическую форму. Формальный прием для математической записи этого логического заключения достаточно прост и состоит в определении количества информации I как фактора уменьшения энтропии от значения $H(X)$, которое характеризует неопределенность измеряемой величины до измерения, до значения $H(X/X_n)$, которое остается после получения показания прибора X_n :

$$I = H(X) - H(X / X_n). \quad (3.32)$$

Величина $H(X)$ представляет собой исходную энтропию, а значение $H(X/X_n)$ характеризует ту неопределенность, которая остается после получения показания прибора X_n и называется *условной энтропией* (т.е. при условии, что X_n известно).

В приведенном выше примере с равномерным законом распределения плотности вероятности как до, так и после измерения исходная, или безусловная, энтропия составляет

$$H(X) = \log(X_2 - X_1), \quad (3.33)$$

а оставшаяся, или *условная*, энтропия результата измерения после получения конкретного отсчета X_n равна

$$H(X/X_n) = \log(2 \cdot \Delta). \quad (3.34)$$

Отсюда следует, что полученное в результате проведения измерения количество информации, равное разности исходной и оставшейся энтропии, равно

$$I = -\log \frac{2 \cdot \Delta}{X_2 - X_1}. \quad (3.35)$$

В этой замене операции деления Δ на $(X_2 - X_1)$, используемой в традиционных подходах при определении *относительной погрешности измерения*, на операцию вычитания исходной и оставшейся неопределенностей, характеризующих соответствующими значениями энтропии, и заключается основной принцип теории информации применительно к измерительным процессам.

б) Процесс измерения как фактор сужения интервала неопределенности

Получение любой информации, в том числе и измерительной, теория информации трактует как устранение некоторой части неопределенности, а количество информации определяется как разность неопределенностей ситуаций до и после получения данного сообщения (измерения). В случае измерения по натуральным шкалам весь диапазон возможных значений измеряемой величины разбивается *реперными* (эталонными) точками на ряд интервалов. Неопределенность до измерения характеризуется тем, что изначально неизвестно, в каком из этих интервалов лежит значение измеряемой величины, т.е. область неопределенности до измерения простирается на все интервалы натуральной шкалы. Результатом измерения является указание на то, что измеряемая величина лежит в данном интервале, т.е. имеет место *сужение области неопределенности* от полной длины натуральной шкалы до длины указанного (после проведения процедуры измерения) интервала.

Таким образом, с точки зрения теории информации результат измерения заключается в выборе данного интервала из целого ряда возможных интервалов. Если при этом предположить, что вероятности попадания измеряемой величины в любой из интервалов равны между собой, то неопределенность исходной ситуации характеризуется безусловной энтропией, равной логарифму числа n интервалов (в качестве которых выступает *цена деления шкалы измерительного прибора*):

$$H(X) = \log n, \quad (3.36)$$

а получаемая в результате измерения информация, соответствующая устранению этой неопределенности, равна

$$I = \log n. \quad (3.37)$$

Подводя итог проведенному предварительному рассмотрению, можно заключить, что в самом общем случае измерение представляет собой сравнение измеряемой величины с тем или иным образом построенной шкалой возможных значений этой величины, а результат измерения состоит в выборе одного интервала из всего множества интервалов этой шкалы. При этом основная особенность результата измерения состоит в том, что он *никогда не может представлять собой точного значения измеряемой величины*, а является лишь указанием более или менее узкого интервала возможных значений.

Основной недостаток натуральной шкалы реперов состоит в отсутствии каких-либо данных о соотношении между собой отдельных интервалов, заключенных между реперными точками. Для интерполяции значений измеряемой величины внутри этих интервалов, а также для сравнительной оценки протяженности этих интервалов необходимо использовать какой-то определенный *принцип пропорционального деления*. В качестве такого принципа при теоретическом рассмотрении часто выдвигают принцип последовательного счета единичных значений измеряемой величины. При этом в качестве базового примера рассматривается процесс измерения длины в виде счета мерных отрезков, откладываемых по одной прямой линии. Поскольку счет конкретных вещей весьма нагляден, то и измерение, выполняемое счетом отдельных количественных содержаний величины, подкупает кажущейся ясностью. Но при внимательном ознакомлении обнаруживается, что такое представление не соответствует действительности и что в основе измерения лежит способ, принципиально отличный от последовательного счета.

В действительности основным методом интерполяции шкалы измеряемых величин между реперными точками является *метод измерительных преобразований*, представляющий собой определение измеряемой величины по значению другой величины, функционально с ней

связанной. Понятие об измерительном преобразовании является ключевым понятием современной теории измерительных устройств и с физической точки зрения означает, что измеряемая величина не может быть определена сама по себе, а может быть воспринята только вместе с тем физическим процессом, в котором она проявляется.

Согласно пункту (а) настоящего подраздела, вопрос о виде функции плотности распределения (случайной) измеряемой величины имеет первостепенное значение. В современных приложениях теории информации к задачам измерительной практики существует два представления о виде закона априорного распределения значений измеряемой величины. Рассмотрим вопрос, чему соответствуют эти два представления о законе распределения плотности исследуемых величин при информационном подходе к анализу метрологических возможностей измерительных устройств.

С точки зрения теории информации измерение представляет собой выбор узкого интервала, в котором находится действительное значение измеряемой величины, из большого числа интервалов, перекрывающих весь диапазон возможных ее значений. *При этом величина этих малых интервалов (независимо от желания наблюдателя) не может быть выбрана безотносительно к размеру самой измеряемой величины.* Для того чтобы размер измеряемой величины определить с какой-то заданной или выбранной точностью, протяженность намечаемых для последующего выбора малых интервалов должна быть в определенное число раз меньше текущего значения самой измеряемой величины. При нарушении этого соотношения корректная постановка процесса измерения становится невозможной.

Так, например, если при измерении длины за такой интервал принят 1 км, то измерение расстояния между двумя городами порядка 500 км будет сводиться к выбору одного значения из ряда возможных значений с шагом 1 км, т.е. из значений 498, 499, 500, 501, 502 км. Это обстоятельство объективно приводит к тому, что вследствие уменьшения протяженности различных градаций (малых интервалов) по мере уменьшения текущего значения измеряемой величины возрастает число этих интервалов на каждую единицу измеряемой величины, т.е. возрастает их плотность распределения.

Обратимся к другому примеру. Пусть при определении вольт-амперной характеристики какого-либо элемента были измерены токи 10, 20, ..., 90 и 100 мА, т. е. с интервалом 10 мА. Затем оказалось, что представляет интерес обследовать также и районы значений тока меньше 10 мА и больше 0.1 А, например до 1 А. При этом в районе от 0 до 10 мА сохранить прежний шаг, равный 10 мА, вообще невозможно, так как, по-видимому, должны быть обследованы точки 1, 2, ..., 8 и 9 мА. Исследование же области от 0.1 до 1 А с шагом 10 мА нецелесообразно, здесь более рационален шаг 0.1 А. Этот пример показывает, что при эксплуатации широкодиапазонных измерительных устройств действительно наблюдается изменение плотности распределения определяемых нами значений измеряемой величины в широких пределах. Так, в диапазоне от 0,1 до 1 А в нашем примере эта плотность составляла 10 (1/А), в диапазоне от 10 до 100 мА – 100 (1/А), а на участке от 1 до 10 мА достигала 1000 (1/А), т. е. изменялась в 100 раз обратно пропорционально текущему значению измеряемой величины.

Сам принцип выбора интервалов по геометрической прогрессии известен издревле. (Еще древнегреческий астроном Гиппарх (II век до н. э.) разделил все видимые на небосводе звезды на 6 степеней по яркости, используя для этого геометрическую прогрессию со знаменателем 2.5). В различных областях техники построение рядов предпочтительных чисел по геометрической прогрессии распространилось настолько широко, что Международная организация по стандартизации (ISO) ввела в действие стандарт на некоторые ряды предпочтительных чисел, учитывающие особенности десятичной системы счисления. Однако упомянутый стандарт – это лишь законодательное подтверждение интуитивно обнаруженного на основе длительного опыта объективно существующего в природе распределения величин в геометрической прогрессии. Более того, везде, где природа в течение длительного периода естественного отбора формировала наиболее совершенные информационные устройства, она также заложила принцип выбора градаций по геометрической прогрессии. Именно в силу этой закономерности естественно развившиеся информационные устройства имеют логарифмическую чувствительность. Логарифмический масштаб длины интервалов, из которых выбираются значения измеряемой величины,

объясняется тем, что для обнаружения существенны только относительные изменения измеряемой величины. Таким образом, общей закономерностью при измерении любой величины в широком диапазоне ее изменения является *необходимость изменения шага выбираемых малых интервалов пропорционально текущему значению самой измеряемой величины*. Коэффициент этой пропорциональности зависит от заданной относительной погрешности измерения, однако необходимость самого пропорционального изменения шага сохраняется при любом выбранном значении относительной погрешности измерения.

В свете вышеизложенного процедуру замены действительного закона распределения на равномерный закон распределения нельзя считать даже приближенно правомерной. Если бы этот закон приближался к равномерному распределению, то в логарифмических координатах он должен был бы изображаться горизонтальной прямой, в то время как в экспериментально полученных зависимостях закон распределения в широком диапазоне изменения измеряемой величины является гиперболическим. В результате приведенных общих соображений и рассмотренных данных статистики можно считать, что общей закономерностью изменения плотности вероятности различных значений одной и той же величины в широком диапазоне ее изменения является закон гиперболического изменения плотности вероятности, т.е. закон логарифмически-равномерного распределения. Интересно отметить, что гиперболический закон для плотности распределения вероятностей значений измеряемой величины в широком диапазоне ее изменения лежит в основе многих практически установившихся положений измерительной техники.

В тех частных случаях, когда действительное распределение плотности подлежащих определению значений измеряемой величины отклоняется от равномерного распределения, но остается достаточно симметричной, исходное предположение о равномерном распределении измеряемой величины можно рассматривать как более или менее правомерное приближение к действительному распределению. На практике это соответствует анализу приборов с очень малым рабочим диапазоном, например, приборов технологического контроля, постоянно установленных для измерения только одной величины, изменяю-

щейся вследствие свойств самого технологического процесса лишь в узких, ограниченных пределах. Действительное распределение значений в этих пределах не является, естественно, равномерным. Оно подчиняется, скорее всего, ограниченному нормальному закону или близкому к нему, несколько асимметричному распределению. Тем не менее в первом приближении замену действительного распределения равномерным при анализе подобных измерительных устройств с узким рабочим диапазоном можно считать правомерной.

Совсем другим является распределение плотности различных значений измеряемой величины при лабораторных измерениях или научных исследованиях, когда именно вследствие гиперболического распределения плотности экспериментатор вынужден использовать многопредельные универсальные измерительные устройства, а не однопредельные приборы. Этот характер закона распределения и следует учитывать при исследовании широкодиапазонных измерительных устройств.

Следует обратить внимание на то, что на краях диапазона измеряемых величин наблюдается снижение плотности $p(x)$. С математической точки зрения гиперболическое распределение величины x соответствует равномерному распределению величины $\lg(x)$, т.е. гиперболический закон распределения можно трактовать как логарифмически-равномерный закон распределения. Использование логарифмически-равномерного закона распределения означает признание факта, что плотность $\lg(x)$ сохраняется постоянной только в каком-то конкретном *ограниченном диапазоне* измеряемого параметра, а затем на границе этого диапазона скачком падает до нуля.

в) Практические методы определения энтропийного показателя погрешности измерения

Согласно рассмотренным выше моментам, энтропийный показатель погрешности измерительного процесса Δ_i может быть представлен в виде [3]

$$\Delta_i = \pm \frac{1}{2} \cdot \exp H(X / X_n). \quad (3.38)$$

Последнее выражение позволяет вычислить энтропийную погрешность процесса измерения по значению условной энтропии $H(X/X_n)$, т. е. энтропии $H(\Delta)$ закона распределения погрешности $p(\Delta)$. Следует, однако, отметить, что при оценке точности измерительных устройств или измерений наблюдатель располагает не самим законом распределения погрешности $p(\Delta)$, а лишь некоторым конечным числом n конкретных значений случайной величины Δ , подчиняющихся этому закону. На основании этого ограниченного числа результатов может быть построена лишь ступенчатая гистограмма (рис. 3.2), в той или иной степени приближающаяся к действительному закону распределения $p(\Delta)$.

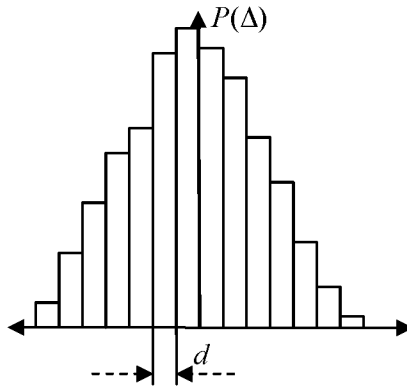


Рисунок 3.2. Гистограмма распределения погрешности измерительного процесса

На практике представляет интерес аналитическое выражение, необходимое для вычисления энтропийного параметра погрешности исходя из такой приближенной гистограммы. Если гистограмма состоит из m столбцов с границами в координатах $X_0, X_1, \dots, X_i, X_{i+1}, \dots, X_m$ и каждый столбец шириной $d_i = X_{i+1} - X_i$ включает в себя n_i результатов, то плотность вероятности на протяжении каждого из столбцов остается постоянной и равной величине:

$$p(\Delta) = \frac{n_i}{n \cdot d_i}. \quad (3.39)$$

Энтропия представленного ступенчатого распределения погрешностей измерения, согласно выражению (3.11), равна

$$H(\Delta) = - \int_{-\infty}^{+\infty} p(\Delta) \ln p(\Delta) dx = - \sum_{i=1}^m \int_{X_{i-1}}^{X_i} \frac{n_i}{nd_i} \ln \frac{n_i}{nd_i} dx = \sum_{i=1}^m \frac{n_i}{n} \ln \frac{nd_i}{n_i}. \quad (3.40)$$

Если ширина всех столбцов принята одинаковой, т.е. $d_i = d$, тогда

$$H(\Delta) = \sum_{i=1}^m \frac{n_i}{n} \ln \frac{n}{n_i} + \ln d. \quad (3.41)$$

Таким образом, энтропия гистограммы определяется как взятая с обратным знаком сумма произведений вероятности случайной величины на логарифм этой вероятности, а член $\ln d$ обеспечивает независимость значения величины энтропии от произвольного выбора величины d ступеньки гистограммы.

В ряде случаев при проведении практических расчетов выражение (3.41) представляют в виде

$$H(\Delta) = \sum_{i=1}^m \ln \left(\frac{n}{n_i} \right)^{\frac{n_i}{n}} + \ln d = \ln \left[d \prod_{i=1}^m \left(\frac{n}{n_i} \right)^{\frac{n_i}{n}} \right]. \quad (3.42)$$

С учетом выражения (3.42) энтропийный показатель погрешности измерения Δ_i приобретает вид

$$\Delta_i = \frac{1}{2} e^{H(\Delta)} = \frac{d}{2} \cdot \frac{n}{10^{\frac{1}{n} \sum_{i=1}^m n_i \lg n_i}}. \quad (3.43)$$

Из выражения (3.43) следует, что полученная оценка погрешности является смещенной, но фактор смещения убывает по мере увеличения n и m . При этом m и d могут выбираться произвольно, но с ограничением по параметру $n_i \geq 2$, так как наличие интервалов с параметрами $n_i = 0$ или $n_i = 1$ не изменяет энтропии, ибо $1^1 = 1$ и $0^0 = 1$.

Воспользоваться соотношением (3.43) можно практически только тогда, когда общее число наблюдений n настолько велико, что может быть разбито на m (по крайней мере 5–7) интервалов шириной d ,

в крайние из которых еще попадает по 1–2 наблюдения. Это приводит к тому, что вычислить энтропийный показатель погрешности исходя из основного определения можно лишь, располагая серией из 30–50 наблюдений. Это и понятно, так как для определения энтропийного показателя погрешности Δ , требуется такой объем испытаний n , при котором уверенно определяется вид закона распределения вероятностей погрешностей, а следовательно, и его энтропия. Поэтому *метод прямого вычисления энтропийного показателя погрешности может быть использован только при обстоятельном исследовании характеризующей погрешности, что связано с необходимостью использования достаточно большой выборки.*

г) Информационная емкость диагностических сигналов измерительного процесса

Одной из характерных составляющих погрешности измерительной аппаратуры, предназначенной для измерения как электрических, так и неэлектрических величин является погрешность от наводки на вход прибора помехообразующих сигналов (в частности синусоидального напряжения с частотой питающей сети). Эта помеха, являясь чисто аддитивной, складывается с полезным сигналом и в ряде случаев определяет порог чувствительности измерительных устройств.

Синусоидальное напряжение помехи от наводок питающего напряжения два раза за период частоты сети проходит через нуль, а между этими значениями достигает положительного и отрицательного максимумов. Однако при синусоидальном изменении напряжения во времени прохождение района нуля происходит относительно быстро, и поэтому вероятность отсчета именно в эти моменты времени оказывается малой. В районе же максимумов, наоборот, изменение напряжения помехи замедляется, напряжение долго сохраняет относительно большую величину, и поэтому вероятность проведения отсчетов именно в эти моменты относительно велика.

Таким образом, если напряжение помехи изменяется по закону $x = X_m \cdot \sin \omega t$ или при $X_m = 1$ по закону $x = \sin \omega t$, то, учитывая, что измерение с равной вероятностью может быть проведено в любой мо-

мент времени, *не синхронизированный по времени с кривой помехи*, вероятность появления значения x погрешности можно принять обратно пропорциональной скорости изменения x в период времени измерительной процедуры:

$$p(x) = \frac{A}{\cos \varphi} = \frac{A}{\cos(\arcsin x)} = \frac{A}{\sqrt{1-x^2}}. \quad (3.44)$$

Нормируя выражение (3.44) в пределах x от (+1) до (-1), для нормирующего множителя A получаем уравнение

$$\int_{-1}^{+1} p(x) dx = A \int_{-1}^{+1} \frac{dx}{\sqrt{1-x^2}} = A\pi = 1. \quad (3.45)$$

С учетом выражений (3.44) и (3.45) для закона распределения вероятностей показаний измерительного процесса имеем (для косиноидального сигнала)

$$p(x) = \frac{1}{\pi \sqrt{1-x^2}}. \quad (3.46)$$

График функции распределения вероятностей исхода измерительного процесса (согласно выражению (3.46)) представлен на рис. 3.3.

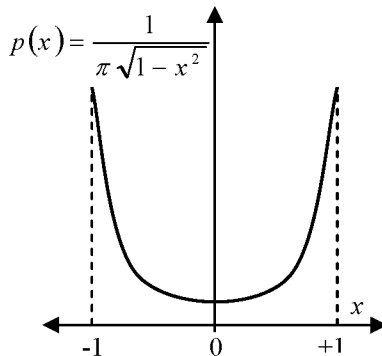


Рис. 3.3. Кривая распределения вероятности для косиноидального сигнала

Для вычисления неопределенности, вносимой в результат измерения погрешности величины x с косиноидальным законом распреде-

ления вероятностей исходов измерения, найдем условную энтропию измерительного процесса:

$$\begin{aligned}
 H(X / X_n) &= - \int_{-1}^{+1} p(x) \ln p(x) dx = \\
 &= \frac{\ln \pi}{\pi} \int_{-1}^{+1} \frac{dx}{\sqrt{1-x^2}} + \frac{1}{\pi} \int_{-1}^{+1} \frac{\ln \sqrt{1-x^2}}{\sqrt{1-x^2}} dx = \ln \pi - \ln 2 = \ln \frac{\pi}{2}.
 \end{aligned}
 \tag{3.47}$$

Используя выражение (3.38) для энтропийной погрешности, вносимой таким распределением, имеем

$$\Delta_i = \frac{\pi}{4} X_m = 0.785 X_m,
 \tag{3.48}$$

а энтропийный коэффициент K_c такого косинусоидального распределения равен:

$$K_c = \frac{\Delta}{\sigma} = \frac{\frac{\pi}{4}}{\frac{1}{2\sqrt{2}}} = \frac{\pi}{\sqrt{2}} = 1.11,
 \tag{3.49}$$

где σ – среднеквадратическая погрешность измерения (традиционно используемая в теории измерительных процессов).

Рассмотренные выше моменты могут быть распространены не только на помехообразующие факторы, но и на сами диагностические сигналы. В этом случае энтропия диагностического сигнала может выступать в качестве информативной меры параметра диагностируемой структуры (изделия).

3.3. Статистические методы анализа кинетики протекания элементарных физико-химических процессов

3.3.1. Физическая сущность статистического метода анализа макроскопических систем (технологический аспект)

Все конденсированные среды (твердые тела и жидкости) имеют концентрацию атомов (молекул) порядка $10^{22} - 10^{23}$ (атомов / см^3). Столь большое количество микрочастиц не позволяет проводить анализ

состояния макросистем с помощью механических подходов. Именно поэтому для этой цели используются статистические подходы и методы анализа. Учитывая значительную роль понятия энергии при рассмотрении динамики протекания практически всех физико-химических процессов, она (энергия) принята в качестве основного параметра статистического подхода [15].

Согласно выводам молекулярной физики *средняя кинетическая энергия теплового движения* микрочастиц равна

$$W_k = n \cdot \frac{k \cdot T}{2}, \quad (3.50)$$

где n – число степеней свободы у индивидуальной микрочастицы (поступательного, вращательного и колебательного типов); k – постоянная Больцмана ($1.38 \cdot 10^{-23}$ Дж/К); T – температура в градусах Кельвина.

Из выражения (3.50) следует, что каждая степень свободы микрочастицы обеспечивается энергией (в среднем по ансамблю частиц) равной $(kT / 2)$. Отметим, что индивидуальные микрочастицы в веществе (ввиду хаотического характера теплового движения) обладают различными энергиями, отличающимися от $(kT / 2)$ как в меньшую, так и в большую сторону. В этой связи более тонкий анализ состояния макросистемы предполагает определение количества частиц ΔN_i , энергия которых заключена в интервале значений $(W_i) \div (W_i + \Delta W)$. В физической и технической практике принято оперировать не параметром ΔN_i (который при устремлении ΔW к нулю теряет физический смысл), а параметром $\Delta N_i / \Delta W$, который получил название *полной функции распределения микрочастиц по энергиям* (подробнее см. в разд. №2). В классическом приближении статистики Максвелла-Больцмана с использованием предельного перехода ($\Delta \rightarrow d$) имеем [16]

$$\frac{dN}{dW} = \alpha \cdot W^{1/2} \cdot e^{-\frac{W}{kT}}, \quad (3.51)$$

где α – постоянная, выполняющая роль нормирующего множителя.

График зависимости функции $\varphi(W) = (dN/dW)$ при различных температурах ансамбля микрочастиц представлен на рис. 3.4.

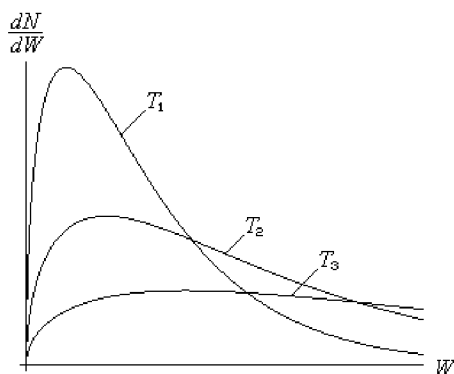


Рис. 3.4. Полная функция распределения частиц по энергиям при различных температурах ($T_1 < T_2 < T_3$)

Представленный на рис. 3.4 типичный вид функции распределения плотности микрочастиц по энергиям (с наличием максимума) обусловлен конкурентным влиянием двух частных функций:

- монотонно возрастающей функцией плотности квантовых энергетических состояний (фрагмент зависимости $W^{1/2}$);
- монотонно убывающей функцией вероятности заполнения квантовых энергетических состояний (фрагмент зависимости $-\exp(-W/kT)$).

Для последующего рассмотрения практических аспектов представляют интерес следующие моменты, нашедшие свое отражение на рис.3.4:

- площадь под графиком полной функции плотности распределения частиц по энергиям равна полному количеству частиц в макросистеме N_0 ;
- положение максимума на графике полной функции плотности распределения частиц по энергиям соответствует значению энергии $W = kT$;
- с увеличением температуры в макросистеме значение максимума плотности распределения частиц по энергиям снижается, а его положение смещается в сторону больших значений энергии.

Приведенная выше статистика Максвелла применима к атомам (молекулам), находящимся в пространстве, в котором отсутствуют внеш-

ние потенциальные поля. В противном случае (т.е. при наличии потенциальных полей) в числителе под экспонентой появляется сумма кинетической (тепловой) и потенциальной составляющих энергии частицы (этот случай не рассматривается в рамках настоящей работы).

3.3.2. Активационная модель протекания элементарных физико-химических процессов

Все элементарные физико-химические процессы, лежащие в основе технологии создания изделий (микро- и наноразмерных структур) и их последующей деградации в процессе эксплуатации, имеют активационный характер. Последнее означает, что для реализации этих процессов атомам (молекулам) изначально необходимо обладать энергией более некоторого определенного порогового значения (характерного для рассматриваемого процесса), называемого *энергией активации процесса* W_a . Именно этим обстоятельством обусловлены определенная устойчивость [17] технических объектов к внешним возмущающим факторам и технологические трудности создания изделий на этапах их производства.

Диапазон вариаций типичных значений энергий активации W_a для ряда (распространенных в технологической практике) элементарных физико-химических процессов приведен в табл. 3.2.

Т а б л и ц а 3.2. *Характерные значения энергии активации для ряда элементарных физико-химических процессов*

№ п/п	Наименование элементарного физико-химического процесса	Энергия активации процесса W_a (эВ)	
		min	max
1	Адсорбция (физическая)	0.005	0.01
2	Хемосорбция	0.05	3.0
3	Диффузия (поверхностная)	0.5	2.0
4	Диффузия (объемная)	1.0	3.0
5	Ионизация атомов	1.5	28
6	Химический процесс	1.0	5.0
7	Образование атомных вакансий в кристаллах	2.5	4.0
8	Перемещение атома в междоузлии кристалла	2.0	5.0

(Поскольку в ряде справочников значение энергии активации W_a приводится в единицах ккал / моль, в рамках выполнения контрольных заданий, приведенных в настоящем пособии, следует произвести перевод единиц измерения согласно равенству $1 \text{ эВ} = 23 \text{ ккал} / \text{моль}$).

Необходимо отметить, что приведенные в справочниках значения энергии активации W_a относятся к индивидуальным процессам как таковым, в то время как на практике зачастую имеет место параллельное (комплексное) протекание сразу нескольких индивидуальных процессов. В этом случае в качестве энергии активации принято рассматривать некоторое *эффективное* значение энергии активации процесса W_a , в котором интегрально нашли свое проявление все индивидуальные физико-химические процессы [18].

3.3.3. Методика определения концентрации энергетически активных микрочастиц

Согласно вышеизложенной активационной модели протекания индивидуальных физико-химических процессов следует, что в указанном процессе могут принимать участие только те микрочастицы, энергия которых превышает некоторое пороговое значение (энергия активации) W_a . В первом приближении можно принять, что скорость (интенсивность) протекания индивидуальных физико-химических процессов будет пропорциональна концентрации высокоэнергетических (активных) частиц N_a . Для определения концентрации энергетически активных микрочастиц N_a проинтегрируем выражение (3.51):

$$N_a = a \cdot \int_{W_a}^{\infty} W^{1/2} \cdot \exp\left(-\frac{W}{k \cdot T}\right) \cdot dW. \quad (3.52)$$

Для упрощения операции интегрирования учтем то обстоятельство, что в практическом плане интересен случай, когда выполняется условие $W_a \gg k \cdot T$, т.е. существенным является только фрагмент $\exp(W/k \cdot T)$:

$$N_a = \beta \cdot \int_{W_a}^{\infty} \exp\left(-\frac{W}{k \cdot T}\right) \cdot dW, \quad (3.53)$$

где β – некоторая нормировочная постоянная.

В результате интегрирования выражения (3.53) для концентрации энергетически активных микрочастиц имеем

$$N_a = N_0 \cdot \exp\left(-\frac{W_a}{k \cdot T}\right), \quad (3.54)$$

где N_0 – полная концентрация микрочастиц в рассматриваемой системе.

Графическая интерпретация концентрации активных частиц N_a представлена на рис. 3.5 в виде заштрихованного энергетического участка.

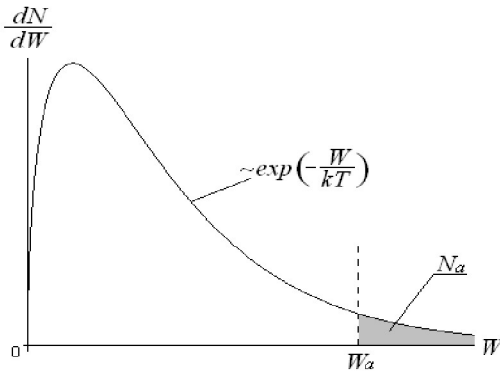


Рис. 3.5. Определение количества активных частиц с помощью статистики распределения частиц по энергии

Согласно рис. 3.5 следует, что концентрация активных частиц N_a (ей соответствует площадь заштрихованного участка на рис. 3.5):

- монотонно уменьшается с увеличением значения энергии активации W_a рассматриваемого физико-химического процесса;
- монотонно увеличивается с повышением температуры T ансамбля микрочастиц в макроскопической системе.

Из вышеизложенного следует тот негативный факт, что неблагоприятные физико-химические процессы (даже с большими значениями энергии активации W_a), приводящие к деградации элементов конструкций устройств, являются принципиально неустраняемыми в эксплуатационной практике.

3.3.4. Динамическое равновесие элементарных физико-химических процессов

Все физико-химические процессы, протекающие в макроскопических системах, являются обратимыми на атомном уровне рассмотрен-

ния. Последнее означает, что параллельно с каждым элементарным атомным процессом (назовем его условно *прямым*) протекает *обратный* ему процесс, противодействующий протеканию прямого процесса. Некоторые примеры обратимого характера протекания элементарных процессов приведены в табл. 3.3.

Т а б л и ц а 3.3. *Примеры обратимых физико-химических процессов*

№ п/п	Элементарный физико-химический процесс	
	<i>прямой процесс</i>	<i>обратный процесс</i>
1	Адсорбция атомов	Десорбция атомов
2	Ионизация атома	Рекомбинация иона
3	Поглощение фотона	Испускание фотона
4	Диссоциация	Ассоциация
5	Реакция разложения	Реакция соединения
6	Реакция окисления	Реакция восстановления

В результате протекания прямых и обратных процессов, в физико-химической системе устанавливается *динамическое равновесие*, при котором скорости протекания прямого и обратного процессов становятся *одинаковыми*. Анализ условия динамического равновесия между прямыми и обратными процессами, протекающими в атомно-молекулярных системах, можно проиллюстрировать на примере химического процесса, являющегося типичным представителем широкого круга элементарных физико-химических процессов, протекающих по механизму, получившему условное название «*рождение-гибель*» [19]:



где A – атомы реагента; B – атомы продукта; K_n , K_o – соответственно константы скорости протекания прямого и обратного химического процесса, определяемые в виде

$$K_n = K_{n0} \cdot \exp(-W_{an} / kT); \quad (3.56)$$

$$K_o = K_{o0} \cdot \exp(-W_{ao} / kT), \quad (3.57)$$

где K_{n0} , K_{o0} – постоянные для данного химического процесса величины; W_{an} , W_{ao} – соответственно энергии активации протекания прямого и обратного химического процесса.

Скорости прямого (ν_n) и обратного (ν_o) химических процессов можно представить в виде [19]

$$\nu_n = \frac{d[A]}{dt} = K_n [A]; \quad (3.58)$$

$$\nu_o = \frac{d[B]}{dt} = K_o [A], \quad (3.59)$$

где $[A]$, $[B]$ – текущие концентрации реагента и продукта в системе.

В условиях химического равновесия, т.е. при выполнении условия $\nu_n = \nu_o$, на основании выражений (3.58) и (3.59) имеем

$$K_p = \frac{K_n}{K_o} = \frac{[B_p]}{[A_p]}, \quad (3.60)$$

где $[A_p]$, $[B_p]$ – соответственно равновесные концентрации реагента и продукта; K_p – константа равновесия химического процесса при динамическом равновесии.

Таким образом, факт химического равновесия в системе свидетельствует о существовании определенного соотношения между равновесными концентрациями атомов (молекул) реагента и продукта (что свидетельствует о принципиальной невозможности достижения полной переработки реагента в продукт). Последнее обстоятельство представляется весьма важным в практическом плане, когда речь идет о технологии создания высокоупорядоченных структур, поскольку оно выдвигает принципиальные ограничения (по своей сути постулирует невозможность) на пути достижения 100% качества обработки материалов. Данный момент характерен не только для случая химических процессов, но и для всей совокупности физико-химических процессов, используемых в технологической практике создания изделий. В этой связи как конструктору, так и технологу изделия необходимо учитывать это обстоятельство на всех этапах конструкторско-технологической проработки изделия.

В рассмотренном случае статистические факторы нашли свое проявление в выражениях (3.56) и (3.57). Более наглядно этот момент представлен на рис. 3.6, на котором отображены полные функции энергетического распределения ансамблей атомов (молекул) реагента и продук-

та по энергии в различные моменты времени на пути установления динамического равновесия в системе.

Центральным моментом рис. 3.6 является то обстоятельство, что площади под кривыми распределения реагента уменьшаются (ввиду протекания прямого химического процесса) со временем, а площади продукта – увеличиваются. Этот процесс продолжается до тех пор, пока концентрации активных частиц для прямого и обратного процессов не сравняются.

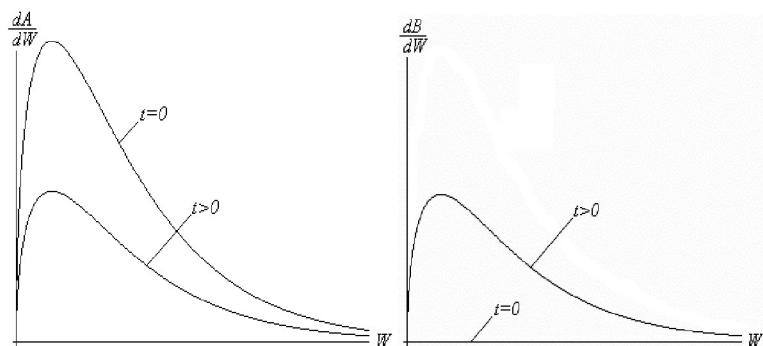


Рис. 3.6. Эволюция полных функций распределения молекул реагента и продукта по энергии со временем

В последнем случае говорят, что имеет место динамическое равновесие в химической системе. Таким образом, понятие о «динамическом равновесии» характеризует тот факт, что *равновесные концентрации реагента и продукта* удовлетворяют условию (3.60), но при этом в системе непрерывно происходят прямые и обратные процессы типа «рождение – гибель» на элементарном (атомном) уровне рассмотрения.

3.3.5. Оценка срока службы материалов и изделий на основе методов физической надежности

Изложенные выше методы физической статистики могут быть использованы для проведения оценки срока службы (долговечности) изделия при различных температурах эксплуатации. Предполагается, что процесс разрушения (деградации) изделия протекает по механизму термической активации, и следовательно, этому процессу деградации изделия можно поставить в соответствие некоторое интегральное зна-

чение энергии активации старения W_δ (на практике в качестве W_δ используется наименьшая энергия активация процесса, приводящего к деградации изделия по рассматриваемому функциональному показателю изделия).

Предположим, что на основании экспериментальных данных известен срок службы изделия τ_1 при температуре эксплуатации изделия T_1 . Требуется оценить срок службы изделия τ_2 при температуре эксплуатации изделия T_2 , в предположении, что температура T_2 не очень сильно отличается от T_1 в том смысле, что *физико-химический механизм деградации изделия остается неизменным*.

Примем гипотезу, что потеря функционального качества изделия наступает по мере накопления в изделии микроскопических (на атомном уровне) дефектов в количестве N_D . В этом случае в первом приближении можно принять, что значение N_D пропорционально сроку службы изделия τ и вероятности возникновения микродефекта, сопутствующего процессу нарушения функционального качества изделия:

$$N_D = \gamma \cdot \tau_1 \cdot \exp\left(-\frac{W_\delta}{k \cdot T_1}\right); \quad (3.61)$$

$$N_D = \gamma \cdot \tau_2 \cdot \exp\left(-\frac{W}{k \cdot T_2}\right), \quad (3.62)$$

где γ – постоянная величина.

Из выражений (3.61) и (3.62) после тривиальных алгебраических преобразований определяем срок службы изделия при температуре T_2 :

$$\tau_2 = \tau_1 \cdot \exp\left(\frac{W_c}{k} \cdot \frac{T_1 - T_2}{T_1 \cdot T_2}\right). \quad (3.63)$$

При использовании выражения (3.63) следует проявлять известную осторожность в выборе интервала ($T_1 - T_2$) с той точки зрения, чтобы при изменении температуры имело место сохранение физического механизма процесса деградации изделия.

Выражение (3.63) может быть использовано для определения режимов ускоренных испытаний изделий, с целью сокращения длительности и трудоемкости процесса испытаний, путем повышения температурных

режимов при определении срока службы изделия [20]. Здесь уместно отметить, что при $W_d > 1$ эВ (это значение характерно для современных устройств современной твердотельной электроники, изготавливаемых по планарной технологии) эффективный объем выборки, подтверждающий надежность и долговечность изделия, достигает значений $10^9 - 10^{10}$ (изделий·час) при проведении испытаний в условиях, соответствующих штатным режимам эксплуатации изделия.

3.4. Термодинамическая энтропия и упорядоченность структур

3.4.1. Элементы равновесной термодинамики

а) Основные понятия и определения

Методология термодинамики носит настолько всеобщий характер, что она с равным успехом может быть применена к объектам практически любой физической природы. Кроме того, в рамках этой теории можно составить представление о закономерностях протекания весьма широкого круга физико-химических процессов [5]. Вклад термодинамического подхода в представление о материи и физических явлениях может быть несколько условно разделен на три части:

- проведение количественных оценок макроскопических характеристик физико-технических систем и параметров протекающих в них физико-химических процессов;
- определение условий, при которых данные процессы становятся принципиально возможными, и установление направления развития равновесных и неравновесных макроскопических систем;
- выявление критериев устойчивости состояния изолированных и открытых физико-химических систем.

При изучении термодинамики (в особенности равновесной) обнаруживается, что в ее основе фигурирует не более четырех законов и значительное множество математических выражений частного характера. При ближайшем рассмотрении выясняется, что эти математические выражения и преобразования включают весьма ограниченный круг параметров, которые с первого взгляда кажутся сходными, но на самом деле сильно различаются. Эта «сухость» и феноменологический харак-

тер подхода в изложении основных принципов и методов термодинамики до некоторой степени является причиной ее малой популярности и абстрактности по сравнению с другими разделами физики.

Однако несомненным достоинством термодинамического подхода является тот факт, что он описывает и устанавливает соотношение непосредственно наблюдаемых параметров различных веществ, причем это можно сделать, *не имея глубоких знаний об их внутренней атомно-молекулярной структуре*. Опираясь всего несколькими законами и ограниченным кругом переменных, можно получить огромную информацию о сложных системах, содержащих множество индивидуальных молекул.

Другое преимущество термодинамики заключается в том, что ее выводы не зависят от физико-технического характера рассматриваемой системы. Специфическая природа веществ задается в теории только несколькими параметрами, например теплоемкостями, молярными объемами, электрической поляризацией и т.д.

Равновесная термодинамика имеет дело с описанием системы в состоянии термодинамического равновесия (или близком к нему) и совершенно не учитывает ее предысторию и специфику пути перехода от одного (начального) состояния к другому (конечному) состоянию.

Неравновесная термодинамика рассматривает динамический аспект эволюции физических систем и, следовательно, имеет дело с объектами, которые находятся в различной степени отклонения от состояния термодинамического равновесия. Следует отметить, что *все технические изделия являются по своей физической сути неравновесными системами (объектами)* и их рассмотрение требует привлечения аппарата неравновесной термодинамики.

Термодинамической системой принято называть конечную область пространства с находящимися в ней физическими объектами макроскопических размеров. Изменяемые свойства термодинамической системы (температура, давление, концентрация химических элементов и т. д.) называются термодинамическими параметрами состояния системы. Некоторые термодинамические параметры, такие как температура T и давление p , численно не зависят от массы и размера системы. Они яв-

ляются *интенсивными* свойствами системы. С другой стороны, существуют *экстенсивные* свойства, количественно пропорциональные массе.

Принято различать основные термодинамические параметры, которые являются постоянными на данный момент времени, и переменные термодинамические параметры, которые рассматриваются как функции основных параметров. Термодинамическая система может быть охарактеризована состоянием таких переменных параметров, связанных между собой функциональной зависимостью $f(V, T, p) = 0$, называемой *уравнением состояния*. Например, для сильно разреженных (идеальных) газов эти переменные связаны простым уравнением состояния:

$$p \cdot V = n \cdot R \cdot T, \quad (3.64)$$

где R – универсальная газовая постоянная; T – абсолютная, или термодинамическая, температура; n – количество молей вещества; V – объем системы; p – давление.

Уравнение (3.64) хорошо известно как уравнение состояния идеального газа. Справедливости ради следует отметить, что определение вида уравнения состояния для других физических систем (твердые тела или жидкости) является достаточно трудной задачей.

б) Внутренняя энергия термодинамической системы

Ключевым моментом равновесной термодинамики является представление о *внутренней энергии системы*. Рассмотрим это представление на примере простой системы «поршень – цилиндр».

Предположим, что поршень движется в цилиндре (рис. 3.7.) до тех пор, пока не займет положение S_1 , поддерживая при этом газ в сжатом состоянии при комнатной температуре. Если теперь отпустить поршень, он будет двигаться в обратном направлении, пока не достигнет уровня S_2 , при котором будет оставаться в положении покоя. В этом положении на поршень действуют три противоположно направленные силы: атмосферное давление и вес поршня толкают его внутрь цилиндра, в то время как газ внутри цилиндра действует в обратном направлении. Равенство этих сил определяет равновесное положение поршня S_2 . Именно в этом смысле следует понимать выражение – «состояние механического равновесия» термодинамической системы.

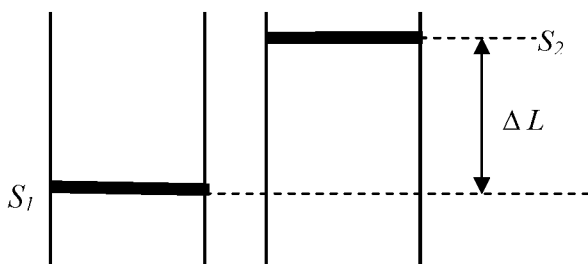


Рис. 3.7. Модель термодинамической системы поршневого типа

В общем случае, когда термодинамические процессы достигают окончательного и неизменного состояния покоя и когда *соответствующие потоки вещества и энергии прекращаются*, можно говорить о состоянии полного (в том числе химического) *термодинамического равновесия*. Концепция равновесия лежит в основе всех видов термодинамического описания равновесных систем.

Одним из основных параметров в термодинамике является температура. В выводах первого и второго законов термодинамики понятие температуры интуитивно связывалось с показаниями термометра и этим ограничивалось. Но поскольку термодинамика как наука развивалась все дальше, необходимо было установить новый закон, который независимо от любой конкретной термометрической шкалы вводил бы и определял понятие температуры как свойства, присущего любой термодинамической системе. Это выполняет нулевой закон термодинамики, который основан на следующих соображениях.

Из опыта известно, что множество объектов можно расположить в ряд в соответствии со степенью их нагрева. Водяной пар теплее жидкой воды, которая, в свою очередь, теплее льда. С другой стороны, при контакте этих тел их температура будет постепенно выравниваться до тех пор, пока в системе не будет теплового равновесия. Нулевой закон термодинамики основан на этом свойстве и может быть сформулирован следующим образом: *если два тела (условно A и B) независимо друг от друга находятся в тепловом равновесии с третьим телом C, то они также находятся в тепловом равновесии друг с другом*. Иными словами, тепловое равновесие характеризуется равенством температуры во всех точках системы.

Первый закон термодинамики вводит привычное и в то же время чрезвычайно абстрактное понятие энергии. Вообще говоря, этот термин описывает способность любого одушевленного или неодушевленного предмета производить работу. До разработки тепловых двигателей единственным источником полезной энергии была мускульная сила. Однако с появлением паровых двигателей была успешно продемонстрирована возможность превращения теплоты в механическую работу.

Очень трудно однозначно ответить на вопрос, что же такое энергия? На практике принято говорить об энергии только в том случае, когда имеет место ее взаимопревращение из одного типа в другой. Следовательно, можно говорить (и измерить) только разность энергий, и *только эта разность имеет физический смысл*. Количество энергии системы может измениться только в результате ее прибыли или убыли в результате ее взаимодействия с внешней средой (условно называемой «термостатом»). Этот факт имеет специфические формы выражения в различных областях физических и химических приложений, а в термодинамике он составляет сущность ее первого закона.

При макроскопическом подходе каждое тело ассоциируется с определенным количеством энергии. Эта энергия может быть подразделена на три вида. В результате сил гравитации (можно обобщить на потенциальные поля любого типа) возникает потенциальная энергия $\Delta E_{пот}$, величина которой зависит от положения тела в пространстве. Движущиеся объекты обладают кинетической энергией $\Delta E_{кин}$. И, наконец, каждое тело от природы наделено определенным количеством внутренней энергии $\Delta U_{внутр}$, которая обусловлена поступательным, вращательным и колебательным движением микрочастиц (атомов, электронов и ядер) и потенциальной энергией их взаимодействия. Следовательно, полная энергия системы равна

$$\Delta E = \Delta E_{кин} + \Delta E_{пот} + \Delta E_{внутр} . \quad (3.65)$$

Обычно термодинамическая система исследуется при таких условиях, когда $\Delta E_{пот}$ не вносит существенного вклада в полную энергию, а $\Delta E_{кин}$ приближается к нулю (не учитывается движение всей системы в целом). Это соответствует ситуации, которая наблюдается для поршня и цилиндра, твердо стоящих на столе. В таком случае $\Delta E = \Delta U_{внутр}$,

т. е., употребляя в термодинамике слово «энергия», имеем в виду именно разность внутренней энергии.

Прежде чем перейти к дальнейшему рассмотрению, необходимо дать математическое выражение обычной механической энергии, которая называется работой. Она определяется произведением силы F на бесконечно малое перемещение:

$$dW = F \cdot dx. \quad (3.66)$$

Универсальность термодинамического подхода обусловлена тем обстоятельством, что в качестве работы dW можно рассматривать всевозможные ее проявления, с учетом специфики рассматриваемых физико-химических процессов. В качестве примера в табл. 3.4 приведены формы представления dW для ряда процессов, обладающих различной физической сущностью и отражающих различные способы изменения внутренней энергии термодинамической системы при ее взаимодействии с внешней средой (термостатом). Целью табл. 3.4 является демонстрация широты охвата рассматриваемых физических явлений при использовании термодинамических методов анализа.

Согласно *первому закону термодинамики*, при изменении энергетического состояния термодинамической системы (например, в результате протекания в ней различного рода физико-химических процессов) сумма количества сообщенного системе тепла ΔQ и затраченной системой работы ΔW *зависит только от начального и конечного состояний системы*, но не зависит от пути перехода к конечному состоянию системы. Эта сумма определяет изменение внутренней энергии системы ΔU :

$$\Delta U = U_2 - U_1 = \Delta Q - \Delta W = \Delta Q - \sum_i X_i \cdot \Delta x_i, \quad (3.67)$$

где ΔQ и ΔW – соответственно количество тепла и работа; U_1 и U_2 – внутренняя энергия системы в начальном и конечном состояниях термодинамической системы. (Знак минус перед фрагментом ΔW обусловлен тем обстоятельством, что по историческим причинам, связанным с созданием паровых двигателей, в качестве положительной работы ΔW рассматривается работа, совершаемая системой над окружающей средой (термостатом)).

Т а б л и ц а 3.4. *Примеры представления элементарной работы dW для ряда физических процессов*

№ п/п	Форма проявления энергии dW элементарного процесса	Аналитическое представление dW
1	Процесс растяжения твердого тела (стержня)	$F \cdot dx$, F – механическая сила; x – длина стержня
2	Объемное расширение газа (поршень в цилиндре)	$p \cdot dV$, p – давление газа; V – объем газа
3	Процесс образования новой поверхности раздела сред	$\sigma \cdot dS$, σ – удельная поверхностная энергия границы раздела сред; S – площадь поверхности раздела
4	Процесс прохождения электрического тока	$q \cdot d\varphi$, q – электрический заряд; φ – электрический потенциал
5	Процесс намагничивания ферромагнетика	$H \cdot dB$, H – напряженность магнитного поля; B – магнитная индукция
6	Процесс поляризации диэлектрика	$E \cdot dD$, E – напряженность электрического поля; D – электрическая индукция
7	Процесс образования новой фазы вещества (кристаллизация)	$\eta \cdot dV$, η – удельная энергия образования новой фазы; V – объем новой фазы
8	Обобщенное представление параметра dW	$\sum_i X_i \cdot dx_i$, X_i – обобщенная сила ($F, p, \sigma, q, H, E, \eta$), x_i – обобщенная координата (x, V, S, φ, B, D)

Если выполняется условие $\Delta U = 0$, то система после совершения кругового цикла возвращается в свое исходное состояние, при этом $\Delta Q = \Delta W$, и первоначальные значения переменных величин восстанавливаются.

Уравнение (3.67) устанавливает общие количественные соотношения в энергетическом представлении физико-химических процессов. Однако следует отметить, что для проведения анализа процессов, про-

текающих в конкретной физической системе, необходимо предварительно установить функциональные соотношения (называемые *термодинамическими функциями*) между термодинамическими параметрами системы в виде строгих аналитических зависимостей.

Эквивалентность теплоты и работы, вытекающая из первого закона термодинамики, основана на том факте, что теплота и работа являются двумя формами передачи энергии. Однако эти формы передачи энергии имеют принципиальное различие, суть которого состоит в следующем:

- работа имеет ярко выраженный *векторизованный характер* в процессе передачи энергии от термостата к термодинамической системе;
- тепловая энергия, обусловленная *броуновским (хаотическим) движением* микрочастиц, является антиподом векторизованного движения микрочастиц.

В отличие от хаотической тепловой энергии векторизованная энергия может быть (в принципе, если не учитывать потери, обусловленные диссипацией энергии, в частности трение) возвращена системе путем реализации обратного процесса. Такой процесс называется обратимым и может быть определен следующим образом: *процесс называется обратимым, если в любой момент времени под воздействием бесконечно малого изменения условий окружающей среды он может менять свое направление на противоположное.*

Долгое время (до появления неравновесной термодинамики) обратимые процессы оставались главными объектами равновесной термодинамики. В рамках равновесной термодинамики оперируют с функциями состояния системы, не интересуясь предысторией состояния и путем перехода, который прошла система. Все положения, определяющие путь перехода от одного состояния системы к другому, лежат за пределами методов равновесной термодинамики и связаны с неравновесным состоянием системы и, соответственно, с протекающими в ней неравновесными процессами

в) Термодинамическая энтропия физико-химической системы

Первое начало термодинамики устанавливает то положение, что при протекании в системе любых физико-химических процессов вы-

полняется закон сохранения энергии и все виды энергии взаимно превращаются в *строго эквивалентных количествах*. Однако первое начало термодинамики не в состоянии ответить на вопросы:

- Почему тот или иной конкретный физико-химический процесс, связанный с трансформацией вида энергии, протекает именно в данном направлении?
- Насколько полно при протекании процесса реализуется фактор превращения энергии из одного вида в другой?
- Каковы реально допустимые пределы протекания индивидуальных физико-химических процессов?

В рамках равновесной термодинамики вводится в рассмотрение существование новой абстрактной функции состояния системы, называемой *термодинамической энтропией*. С помощью понятия энтропии можно дать современную формулировку второго закона термодинамики: *энтропия изолированной системы при приближении к состоянию термодинамического равновесия монотонно возрастает до тех пор, пока не достигнет своего максимального значения*.

Приведенное выше в *форме постулата* определение понятия энтропии *формально* вводит в рассмотрение новую функцию состояния термодинамической системы, но ни в коей мере не раскрывает физическое содержание введенного в рассмотрение понятия энтропии. Ответ на этот вопрос содержится в статистической физике (см. разд. 2.2.1): *параметр «энтропия» определяет меру неупорядоченности физико-химических систем (т.е. в контексте рассматриваемых далее задач она определяет меру фактора хаотичности расположения атомов в системе)* [21].

Энтропия является таким же экстенсивным параметром термодинамической системы, как например энергия, и ее значение растет с увеличением массы системы и равно сумме энтропий подсистем, образующих рассматриваемую систему.

Изменение энтропии dS термодинамической системы равно отношению количества теплоты δQ , сообщенной системе, к термодинамической температуре T , при которой происходит процесс передачи энергии (теплоты):

$$dS = \frac{\delta Q}{T}. \quad (3.68)$$

Сущность второго начала термодинамики для *изолированных систем* состоит в том, что *все необратимые процессы протекают в таком направлении, в котором энтропия системы увеличивается*. Последнее заключение означает, что далеко не вся внутренняя энергия термодинамической системы может быть переведена в работу (т.е. быть реализованной в форме векторизованной формы движения), а часть ее непременно будет рассеяна (деградирована) в виде теплоты (т.е. хаотической составляющей движения микрочастиц). В этом случае изменение энтропии подчиняется условию $dS > 0$. И только при *обратимых* процессах (что является идеализацией реальных процессов) имеет место сохранение энтропией своего первоначального значения, т.е. выполняется условие $dS = 0$.

Это свойство энтропии, характеризующее эволюцию состояния изолированной термодинамической системы, может быть обобщено и для случая необратимых переходов состояния системы:

$$S_B - S_A > \int_A^B \left(\frac{dQ}{T} \right), \quad (3.69)$$

где A и B – символы, характеризующие соответственно начальное и конечное состояния термодинамической системы в процессе ее эволюции.

Согласно выражению (3.69), термодинамическая система в ходе своей спонтанной эволюции приближается к такому состоянию, при котором ее энтропия имеет максимальное значение (*принцип максимальной энтропии*). Этот факт имеет большую практическую значимость, ибо из него следует, что:

- технологическим процессам создания технических изделий, направленных на уменьшение энтропии системы (т.е. повышению степени упорядоченности материальных сред с целью достижения функционального качества изделия), сопутствуют параллельно протекающие процессы спонтанной дезорганизации системы, преследующие цель повышения энтропии системы;

- срок службы технических изделий в процессе их эксплуатации всецело определяется интенсивностью протекания процессов дезорганизации материальных сред в изделии, т.е. скоростью возрастания энтропии;

- реализация 100 % функционального качества и бесконечный срок службы технических изделий являются принципиально недостижимыми событиями.

Распространим рассмотренные моменты на процесс эволюции энтропии *открытых систем*. *Термодинамическая система считается открытой, если она может обмениваться с окружающей средой (термостатом) энергией (в виде работы или теплоты) или веществом (атомы, молекулы), а в более широкой постановке и информацией* (когда речь идет о внешнем управляющем воздействии в технологической практике). В этом случае полное изменение энтропии dS термодинамической системы состоит из двух составляющих:

- продукции (*производства*) энтропии внутри самой термодинамической системы $d_i S$, обусловленной проявлением принципа максимальной энтропии;

- транспорта (*обмена*) энтропии термодинамической системы с окружающей средой (термостатом) $d_e S$.

С позиции этих представлений

$$dS = d_i S + d_e S. \quad (3.70)$$

С точки зрения практических интересов весьма важно отметить, что слагаемое $d_e S$ в выражении (3.70) может иметь как положительный, так и *отрицательный* знак. Возможность отрицательного знака у составляющей $d_e S$ свидетельствует о том, что в принципе существует возможность уменьшения энтропии термодинамической системы, а следовательно, повышения степени упорядоченности ее материальных сред. Эту функцию, как правило, в производственной деятельности выполняет совокупность технологий создания изделий (в частности, технологии специальной обработки материалов рабочих сред).

Факт производства энтропии внутри самой термодинамической системы ($d_i S$) играет большую роль при рассмотрении вопросов, связанных как с технологической практикой производства изделий, так и с процессами их деградации в ходе эксплуатации. В первом случае речь идет о дестабилизирующем воздействии фактора производства энтропии на качество технологических процессов в производственной практике. Во втором случае процессы нарушения атомно-молекулярного порядка в процессе срока службы изделия, обусловленные спонтанным ростом энтропии, приводят к расходованию начального «запаса качества» (ресурса) изделия. Учитывая практическую важность этих аспектов, анализ вклада фактора $d_i S$ в энтропию термодинамической системы будет продолжен ниже.

В заключение настоящего подраздела рассмотрим вопрос о влиянии фактора размерного «масштабирования» физической системы на расчетные показатели термодинамической энтропии. В основу термодинамической энтропии физических систем положена модель сплошной среды, чем и определяется степень ее идеализации при использовании макроскопических подходов. В этом случае теряется информация о движении индивидуальных микрочастиц в пределах локальных (но по своей сути еще макроскопических) «точек» сплошной среды, для которых можно ввести в рассмотрение совокупность термодинамических параметров, однозначно определяющих состояние материальной среды в пределах «точки». Оценим число используемых и теряемых степеней свободы системы при переходе от системы типа «ансамбль микрочастиц» к приближению типа «сплошная среда» [22]. Для системы из N атомов полное число степеней свободы равно $6N$. Здесь имеется в виду то обстоятельство, что состояние микрочастицы однозначно определяется совокупностью параметров $\{x, y, z, p_x, p_y, p_z\}$. Обозначим посредством N_c среднее количество частиц (атомы или молекулы) в физически бесконечно малом объеме – локальной «точке» сплошной среды. Относительное число используемых степеней свободы можно оценить как N/N_c . Тогда число «невостребованных» степеней свободы N_x определяется выражением [23]

$$N_x = 6 \cdot N \cdot \frac{N_c - 1}{N_c}, \quad (3.71)$$

и оно составляет львиную долю полного числа степеней свободы системы. Это обстоятельство следует иметь в виду при проведении анализа энтропии реальных физических систем и технических объектов.

Таким образом, можно заключить, что *«платой» за упрощение методов анализа термодинамических систем является существенная потеря степени детализации (на атомно-молекулярном уровне) процессов и объектов рассмотрения.*

3.4.2. Элементы неравновесной термодинамики

а) Энтропия и необратимость физико-химических процессов

Термодинамическая система, находящаяся в равновесном состоянии, является *«статичной»*, поскольку для нее не существует представления об ее эволюции во времени. Считается, что система когда-то перешла в это равновесное состояние из неравновесного состояния, при котором она была динамической системой и обладала *«силами»*, которые приводили к различным изменениям ее состояния в части материального состава и структуры и при этом постепенно подвели к ее закономерному финалу – термодинамическому равновесию. Эту тенденцию движения к состоянию равновесия можно избежать, если с помощью определенных *внешних сил* искусственно поддерживать систему в состоянии, далеком от термодинамического равновесия [24].

В неравновесной термодинамике понятия *«сила»* или *«обобщенная сила»* распространяется на все виды воздействия, включая и обычные механические силы. Эти силы, обусловленные наличием градиентов термодинамических параметров (концентрации, давления, химического или электрического потенциала и т.д.), могут поддерживать в системе постоянный приток или отток вещества и энергии [5]. Таким образом, эти силы образуют потоки энергии или вещества, которые в конце концов истощают сами силы, их породившие. Все градиенты со временем постепенно исчезают (если не поддерживаются искусственным образом) и система достигает окончательного состояния термодинамического равновесия. Например, источник движущей силы в виде темпера-

турного градиента между двумя точками в объеме вещества порождает перенос какого-то количества теплоты Q из горячей части системы в холодную. Этот поток теплоты увеличивает температуру холодного участка за счет охлаждения горячего участка, что постепенно приводит систему к состоянию теплового равновесия.

Наличие сил и потоков в неравновесной системе означает, что эта система неоднородна по своему составу и что в ней происходят различные физико-химические процессы. Поэтому в наиболее общем случае состав и все соответствующие переменные системы изменяются в пространстве и во времени. Возникает вопрос, каким же образом в неравновесных системах могут быть определены интенсивные параметры (подобные температуре или давлению), если они являются функцией координат и, вообще-то, введены в рассмотрение исходя из представлений о равновесном характере системы. Эти трудности могут быть преодолены, если воспользоваться представлением о *локальном равновесии* системы. В этом случае вся макроскопическая система условно расчленяется на совокупность большого количества *локальных подсистем* (остающихся по своей сути макроскопическими объектами) таким образом, что каждая отдельная подсистема находится в состоянии термодинамического равновесия (т.е. внутри ее термодинамические параметры сохраняют неизменное значение).

Примем допущение, что все выводы равновесной термодинамики могут быть применены к локально определяемым параметрам небольшого элементарного объема. Локальные переменные принято относить к единице массы вещества. Но удобнее перейти к переменным, соотнося их с объемом системы. Тогда, если ρ_i – парциальная масса на единицу объема (парциальная плотность) i -го компонента, то полная плотность ρ , удельная энтропия \hat{s} и удельная внутренняя энергия \hat{u} (отнесенные к единице объема среды) определяются выражениями [5]:

$$\rho = \sum_i \rho_i; \quad (3.72a)$$

$$\hat{s} = \sum_i (s_i \cdot \rho_i); \quad (3.72б)$$

$$\hat{u} = \sum_i (u_i \cdot \rho_i), \quad (3.72в)$$

где s_i и u_i – соответственно удельные энтропия и внутренняя энергия единицы объема компонента вещества i -типа.

Согласно принятой гипотезе относительно правомерности использования принципа локального равновесия применительно к термодинамической системе в целом, параметры локальных подсистем, определяемые с помощью выражений (3.72), по своей природе становятся аналогичными обычным параметрам, характеризующим равновесное состояние термодинамической системы. Таким образом, для каждого малого объема неравновесной термодинамической системы его локальная энтропия может быть связана с локальной энергией и парциальной массой соотношением, которое является частным случаем известного уравнения Гиббса [5]:

$$T \cdot d\hat{s} = d\hat{u} - \sum_j X_j \cdot dx_j - \sum_i \mu_i \cdot d\rho_i, \quad (3.73)$$

где μ_i – химический потенциал единицы объема компонента вещества i -типа.

Согласно выражению (3.72б) общая энтропия термодинамической системы, имеющей объем V , будет равна

$$S = \int_V \hat{s} \cdot dV. \quad (3.74)$$

Выше был продемонстрирован общий подход теоретического анализа неравновесной системы. Если термодинамическая система является пространственно-неоднородной по какому-то либо термодинамическому параметру, то необходимо предварительно определить функцию распределения для этого параметра, а затем проводить интегрирование выражения (3.74) по всему объему термодинамической системы.

б) Производство энтропии в неравновесной системе

Наиболее существенным признаком необратимого процесса является его динамика во времени. Свойства неравновесной системы постоянно изменяются во времени и пространстве, при этом эволюция термодинамической системы протекает в направлении непрерывного повышения энтропии системы (если на нее не распространены специальные меры внешнего управляющего воздействия).

Интересно рассмотреть, откуда вообще появляется фактор необратимости в поведении сложных атомно-молекулярных систем, если на уровне рассмотрения поведения единичных микрочастиц этот фактор отсутствует как таковой. Особенно наглядно этот момент проявляется при рассмотрении механических процессов, в которых движение микрочастиц подчиняется строго детерминированным законам Ньютона.

Следует отметить, что замена в механических закономерностях движения микрочастиц параметра t (время) на параметр $(-t)$ не приводит к парадоксальному результату, а, как это и можно ожидать, реализует обратный процесс вполне обратимым образом. Из этого следует, что на уровне сложности системы «одна микрочастица» фактор необратимости процессов эволюции систем не проявляется и нет места понятию неопределенности состояния системы (не принимая во внимание квантово-механические ограничения, обусловленные принципом неопределенностей Гейзенберга, рассмотренным в разд. 1).

Рассмотрим достаточно простую, но состоящую из множества микрочастиц термодинамическую систему типа «идеальный газ», уравнение состояния которой можно представить в виде [7]

$$P \cdot V = N \cdot k \cdot T, \quad (3.75)$$

где N – количество микрочастиц, содержащихся в объеме V идеального газа при температуре системы T .

Подвергнем газ *изотермическому* ($T = \text{const}$) расширению. В этом случае, согласно выражению (3.68), полное изменение энтропии газа составит

$$\Delta S = \int_{V_1}^{V_2} \frac{dQ}{T}. \quad (3.76)$$

Чтобы выполнить работу по расширению газа, необходимо придать системе (газу) дополнительную энергию в виде некоторого количества тепла, которое для малых изменений объема можно представить в виде

$$dQ = P \cdot dV. \quad (3.77)$$

На основании выражений (3.75) – (3.77) для полного изменения энтропии газа в результате его расширения получаем

$$\Delta S = \int_{V_1}^{V_2} P \cdot \frac{dV}{T} = N \cdot k \cdot \ln \frac{V_2}{V_1}. \quad (3.78)$$

Для определенности примем, что $V_2/V_1 = 2$, т.е. в результате расширения газа произошло удвоение его объема. В этом случае можно считать, что изменение энтропии, приходящееся на одну микрочастицу, составит величину $k \cdot \ln 2$ (число 2 появилось в результате того, что объем увеличился вдвое!). Весьма примечательно то обстоятельство, что в изменении энтропии никоим образом не нашли отражения индивидуальные свойства самой микрочастицы, а нашло проявление *свободного объема*, в пределах которого находится рассматриваемая микрочастица.

Отсюда следует, что энтропия системы увеличивается (при расширении газа) в условиях, когда температура и внутренняя энергия системы не меняются, а изменяются лишь условия пространственного размещения (изменение объема) ансамбля микрочастиц. Последнее становится более понятным, если учесть тот факт, что при увеличении объема термодинамической системы увеличивается также и *неопределенность наших знаний* относительно местоположения индивидуальной микрочастицы в пределах рассматриваемого объема. Иными словами, предельная локализация микрочастиц ($V_2 =$ объем ансамбля из N атомов) приводит к снижению неопределенности местоположения индивидуальных микрочастиц до возможного теоретического предела (аналог этой ситуации имеет место для случая совершенных кристаллических структур, широко используемых в современных электронных интегральных схемах).

в) Производство энтропии в линейных необратимых процессах

В равновесных термодинамических системах интенсивные параметры (такие как температура, давление, химический потенциал и т.д.), однозначно характеризующие состояние макросистемы, сохраняют (если абстрагироваться от процессов флуктуационной природы) свое постоянство во времени и в пространственных пределах макросистемы в целом. Последнее находит свое отражение в математической формулировке вида

$$\text{grad}U = \frac{dU}{dx} \cdot \bar{i} + \frac{dU}{dy} \cdot \bar{j} + \frac{dU}{dz} \cdot \bar{k} \equiv 0, \quad (3.79)$$

где U – интенсивный параметр макросистемы; $\bar{i}, \bar{j}, \bar{k}$ – соответственно единичные векторы по координатным осям $\{x, y, z\}$.

При нарушении условия (3.79), т.е. в случае $\text{grad}U \neq 0$, в термодинамической системе возникают необратимые процессы:

- *переноса*, т.е. возникновения потоков массы, электрического заряда, энергии (как в векторизованной форме в виде работ определенно-го типа, так и в не векторизованной форме в виде тепла);
- *химического* (электро- или фотохимического) *взаимодействия* различных компонентов системы, обусловленного наличием градиентов концентрации реагентов в системе (или внешних энергетических инициаторов процессов).

Из практического опыта известен широкий круг такого рода необратимых процессов, в которых интенсивность потоков (сюда же входит и химическое взаимодействие) *пропорциональна* градиентам соответствующих интенсивных переменных. Некоторые примеры таких потоков приведены в табл. 3.5.

Таблица 3.5. *Примеры необратимых процессов в линейном представлении*

Тип необратимого процесса	Поток	Движущая сила
Объемная диффузия	J_d	$\text{grad}(N)$, N – локальная концентрация частиц
Теплопроводность	J_q	$\text{grad}(T)$, T – локальная температура
Химический процесс	v_r	$\text{grad}(\mu)$. μ – химический потенциал (локальное значение)
Электрический ток	σ	$\text{grad}(\varphi)$, φ – электрический потенциал

В случае наличия в системе двух (или более) движущих сил появляются два необратимых процесса, которые налагаются друг на друга.

Например, от наложения теплопроводности и объемной диффузии поток массы возникает под воздействием и градиента концентрации, и градиента температуры:

$$J = -D \cdot \text{grad}N - \chi \cdot \text{grad}T. \quad (3.80)$$

Причины, вызывающие появление необратимых процессов, в неравновесной термодинамике принято называть *движущими силами процесса* и обозначать в виде символа X_i , где индекс i – условный номер, присваиваемый каждому типу градиента интенсивного термодинамического параметра. Количественные характеристики интенсивности протекания необратимых процессов, вызываемых силами X_i , называются *потоками*, которые обозначаются в виде символа J_i .

Термодинамика линейных необратимых процессов является обобщением равновесной термодинамики и линейных феноменологических закономерностей известных необратимых процессов (см. табл. 3.5 и выражение (3.80)). В линейном приближении предполагается, что наличие градиентов параметров термодинамической системы не приводит (ввиду их малости) к существенному изменению, а в более жесткой постановке – неправомерности использования как таковой равновесной функции распределения микрочастиц по энергиям. Последнее приближение оказывается допустимым в случае возможности использования принципа локального равновесия, частично рассмотренного выше.

Рассмотрим вопрос о границах применимости принципа локального равновесия. Представляется очевидным, что *время релаксации* термодинамической системы к равновесию возрастает с увеличением размеров системы. Поэтому отдельные, достаточно малые, но по своей сути – макроскопические, части системы приходят в равновесие гораздо раньше, чем устанавливается равновесие между самими этими частями. Благодаря этому можно говорить о локальном равновесии в частях системы и полагать, что между соседними частями имеют место *малые отклонения* от равновесного состояния. Точный смысл того, какие отклонения можно считать малыми, решаются с учетом специфики рассматриваемой задачи. В частности, в случае газовых систем в качестве малых можно считать такие отклонения, при которых изменение тер-

модинамического параметра ΔU на длине свободного пробега λ или за время свободного пробега τ гораздо меньше самого параметра:

$$\Delta U = \left| \frac{dU}{dx} \right| \cdot \lambda \ll U; \quad \Delta U = \left| \frac{dU}{dt} \right| \cdot \tau \ll U. \quad (3.81)$$

В случае выполнения принципа локального равновесия упомянутое выше обобщение равновесной термодинамики и линейных необратимых процессов сводится к ряду постулатов теории необратимых процессов [25].

Постулат №1. *Состояние отдельных частей системы, находящихся в локальном равновесии, в каждый момент времени определяется основным уравнением термодинамики для равновесных процессов (уравнение Гиббса):*

$$T \cdot d\bar{s} = d\bar{u} - \sum_j X_j \cdot dx_j - \sum_i \mu_i \cdot d\rho_i. \quad (3.82)$$

Согласно выражению (3.82) принимается, что энтропия неравновесной системы является той же функцией термодинамических параметров системы, как это уже было принято при рассмотрении случая полного термодинамического равновесия (закономерное следствие, вытекающее из принципа локального равновесия).

Постулат №2. *Поток J_i , вызванный действием сил X_i , пропорционален этим силам (линейный закон):*

$$J_i = \sum_{k=1}^n L_{ik} \cdot X_k. \quad (3.83)$$

Параметры L_{ik} называются *феноменологическими (кинетическими) коэффициентами* процесса. Коэффициенты типа L_{ii} (при выполнении условия $i = k$) характеризуют *прямые* процессы (тепло- или электропроводность, диффузия, химический процесс и т.д.). Коэффициенты вида L_{ik} (при выполнении условия $i \neq k$) характеризуют *налагающие* (вторичные) процессы (термодиффузия, термоэлектричество и т.д.).

Постулат №3. *При определенном выборе потоков J_i и сил X_i матрица кинетических коэффициентов становится симметричной (соотношение взаимности Онзагера):*

$$L_{ik} = L_{ki}, (i, k = 1, 2, \dots, n). \quad (3.84)$$

Выражение (3.84) свидетельствует о том, что имеется некоторая симметрия во взаимодействии налагающих процессов. Например, подобно тому как градиент температуры вызывает градиент концентрации, так и градиент концентрации порождает градиент температуры. Доказательство симметрии кинетических коэффициентов по индексам i и k основывается на принципе обратимости элементарных процессов на микрокопическом уровне рассмотрения.

Постулат №4. При выполнении уравнения Гиббса в виде (3.82) локальная скорость производства энтропии σ всегда имеет вид

$$\sigma = \sum_i J_i \cdot X_i. \quad (3.85)$$

Приведенные выше постулаты термодинамики линейных необратимых процессов образуют самосогласованную систему уравнений для определения как сопряженных потоков J_i и X_i сил неравновесных процессов, так и всего комплекса термодинамических функций, необходимых для практических целей.

В рамках равновесной термодинамики был сформулирован (в виде постулата) общий принцип эволюции энтропии для изолированных систем – принцип максимума энтропии в равновесном состоянии системы (см. разд. 3.4.1, в). В этой связи возникает вопрос о том, каким общим принципам подчиняется поведение энтропии при протекании неравновесных процессов? Для ответа на этот вопрос были введены (Онзагер) две дополнительные диссипативные функции (функции рассеяния) [25]:

$$\varphi(X, X) = \frac{1}{2} \cdot \sum_{i,k=1}^n L_{ik} \cdot X_i \cdot X_k \geq 0, \quad (3.86)$$

$$\Phi(J, J) = \frac{1}{2} \cdot \sum_{i,k=1}^n L_{ik}^{-1} \cdot J_i \cdot J_k \geq 0, \quad (3.87)$$

которые, как и параметр σ (см. выражение (3.85)), являются локальной мерой необратимости процессов.

Используя соотношения взаимности Онзагера (3.84), можно показать потенциальный характер функций (3.85) – (3.87). Для этого проведем дифференцирование функций (3.86) и (3.87):

$$\frac{d\varphi(X, X)}{dX_i} = J_i = \sum_{i,k=1}^n L_{ik} \cdot X_k, \quad (3.88)$$

$$\frac{dJ_i}{dX_k} = L_{ik} = L_{ki} = \frac{dJ_k}{dX_i}, \quad (3.89)$$

откуда следует, что функция $\varphi(X, X)$ является *неравновесной потенциальной функцией* состояния системы (в равной мере это относится к функциям $\Phi(J, J)$ и $\sigma(J, X)$). Обращает внимание тот факт, что производные (3.88) выражают линейный закон (3.83), а производные (3.89) – соотношение взаимности (3.84). Отсюда следует, что потенциалы $\varphi(X, X)$ и $\Phi(J, J)$ представляют собой функции, в которых находит свое отражение основные принципы термодинамики необратимых процессов.

з) Вариационный принцип наименьшего рассеяния энергии в стационарных процессах

Соотношение взаимности (3.84) для линейных процессов эквивалентно вариационному принципу *наименьшего рассеяния энергии*. Согласно этому принципу, для фактически протекающих (а не теоретически возможных) в системе процессов разностная форма [25]

$$\sigma(J, X) - \varphi(J, X) \quad (3.90)$$

имеет максимум по сравнению со всеми возможными виртуальными необратимыми процессами, протекающими с теми же самыми потоками, но различными сопряженными силами. Действительно, варьируя (3.90) по силам X_i при постоянных потоках J_i , получаем

$$\begin{aligned} \delta [\sigma(J, X) - \varphi(J, X)]_I &= \delta \left[\sum_{i=1}^n J_i \cdot X_i - \frac{1}{2} \cdot \sum_{i,k=1}^n L_{ik} \cdot X_i \cdot X_k \right]_I = \\ &= \sum_{i=1}^n J_i \cdot \delta X_i - \sum_{i,k=1}^n L_{ik} \cdot \delta X_i \cdot X_k = \sum_{i=1}^n \left(J_i - \sum_{k=1}^n L_{ik} \cdot X_k \right) \cdot \delta X_i. \end{aligned} \quad (3.91)$$

Последнее выражение с учетом выражения (3.83) приводит к условию экстремума разностной формы (3.12), определяемого в виде

$$\delta[\sigma(J, X) - \varphi(J, X)] \equiv 0. \quad (3.92)$$

Так как диссипативная функция $\varphi(J, X)$ является квадратичным и положительно определенным выражением относительно независимых сил X_i , то экстремум, определяемый выражением (3.92), может быть только *максимумом*.

Дифференциальный вариационный принцип в форме выражения (3.92) может быть представлен в интегральной форме:

$$\delta \int_V [\sigma(J, X) - \varphi(J, X)]_I \cdot dV \equiv 0, \quad (3.93)$$

или в виде, более востребованном на практике:

$$[P - \bar{\varphi}]_I = [\dot{S} + \dot{S}^* - \bar{\varphi}]_I = \max, \quad (3.94)$$

где

$$P = \int_V \sigma(J, X) \cdot dV = \int_V \rho \cdot \dot{S}_V \cdot dV + \int_V \text{div} J_s \cdot dV \equiv \dot{S} + \dot{S}^*; \quad (3.95a)$$

$$\bar{\varphi} = \int_V \varphi(J, X) \cdot dV. \quad (3.95b)$$

Из общей интегральной формы (3.94) можно сформулировать ряд выводов относительно ряда частных типов линейных необратимых процессов:

- В адиабатически изолированных неравновесных системах мощность потока термодинамической энтропии \dot{S}^* через внешнюю поверхность системы равна нулю (по определению), поэтому согласно выражению (3.94) для процессов этого типа выполняется условие максимума вида

$$[\dot{S} - \bar{\varphi}]_I = \max. \quad (3.96)$$

В изолированных системах протекание стационарных процессов невозможно, поскольку для поддержания этих процессов необходим постоянный приток энергии.

- В стационарных процессах (т.е. интенсивные параметры системы не являются функцией времени), протекающих в открытых систе-

мах, полная энтропия системы остается постоянной ($\dot{S} = 0$) и величина \dot{S}^* при постоянных потоках задана, поэтому для процессов этого типа выполняются условия:

$$\left[\dot{S}^* - \bar{\varphi} \right]_I = \max \quad (3.97a)$$

$$\int_V \varphi(J, X) \cdot dV = \min. \quad (3.97b)$$

- Из выражения (3.97b) следует, что при стационарных процессах в открытых системах, широко используемых в современной технологической практике создания микро- и наноструктур, диссипация (т.е. рассеяние векторизованной составляющей) энергии имеет минимальное значение.

По определению, система находится в стационарном состоянии, если характеризующие состояние системы параметры (например парциальные концентрации атомов определенного типа) не испытывают вариации во времени. В рамках использования принципа локального равновесия из этого условия вытекает, что локальные функции состояния системы типа плотности энтропии s или плотности производства энтропии σ также не зависят от времени. Для реализации устойчивого стационарного состояния характерно выполнение условия относительно транспорта (*обмена*) энтропии термодинамической системы с окружающей средой dS_e :

$$dS_e \leq 0. \quad (3.98)$$

Из выражения (3.98) следует, что для поддержания стационарного неравновесного состояния необходимо постоянно направлять в систему *отрицательный поток энтропии* от внешней среды (термостата), равный по величине внутреннему производству энтропии dS_i (см. также выражение (3.70)).

Как следует из экспериментальной практики, если термодинамическая система находится под воздействием не зависящих от времени возмущающих факторов, то по прошествии достаточно большого промежутка времени (по сравнению с продолжительностью элементарных взаимодействий на атомном уровне) она неизбежно приходит в стационарное состояние.

д) *Принцип минимума производства энтропии*

Вопрос о *специфической особенности* стационарных необратимых процессов сводится к установлению физического параметра, который при стационарном процессе имел бы *экстремальное значение*, подобно тому как равновесное состояние термодинамической системы может быть охарактеризовано максимальной энтропией. (Здесь и выше стационарные процессы рассматриваются подробно лишь по той причине, что для нестационарных процессов, ввиду недостаточного уровня разработки общей теории необратимых процессов, постановка подобного рода вопроса является преждевременной). Ответ на этот вопрос был дан в виде рассмотренного выше *принципа наименьшего рассеяния энергии (Онзагер)* и в виде *принципа минимума производства энтропии (Пригожсин)* [26].

Физическая суть принципа минимума производства энтропии состоит в следующем заключении: *стационарное состояние термодинамической системы, в которой протекают необратимые процессы, характеризуется тем фактором, что суммарная скорость производства энтропии в системе имеет минимальное значение при данной совокупности значений внешних граничных условий, препятствующих достижению системой равновесного термодинамического состояния.*

Рассмотрим этот принцип более детально, для чего определим производство энтропии при переносе энергии и вещества между двумя фазами, находящимися при разных температурах (принято, что других градиентов интенсивных параметров в системе нет!). Согласно выражениям (3.83) и (3.85) для упомянутой неравновесной системы для скорости производства энтропии и потоков в системе имеем:

$$\sigma = J_1 \cdot X_1 + J_2 \cdot X_2; \quad (3.99a)$$

$$J_1 = L_{11} \cdot X_1 + L_{12} \cdot X_2; \quad (3.99б)$$

$$J_2 = L_{21} \cdot X_1 + L_{22} \cdot X_2, \quad (3.99в)$$

где J_1 – поток теплоты; J_2 – поток вещества, обусловленный наличием градиента температуры; X_1 и X_2 – сопряженные этим потокам термодинамические силы.

При заданной разности температур между фазами ($X_1 = \text{const}$), если поток теплоты J_1 будет постоянным, а поток вещества J_2 равен нулю (термодиффузия компенсируется объемной диффузией):

$$J_2 = L_{21} \cdot X_1 + L_{22} \cdot X_2 = 0. \quad (3.100)$$

С учетом соотношения взаимности (3.84) производство энтропии в стационарном состоянии системы составит.

$$\sigma = L_{11} \cdot X_1^2 + 2 \cdot L_{12} \cdot X_1 \cdot X_2 + L_{22} \cdot X_2^2 > 0. \quad (3.101)$$

Дифференцирование функции (3.101) по параметру X_2 при постоянном параметре X_1 приводит к выражению:

$$\left(\frac{\partial \sigma}{\partial X_2} \right)_{X_1} = 2 \cdot (L_{21} \cdot X_1 + L_{22} \cdot X_2) = 2 \cdot J_2 = 0, \quad (3.102)$$

что свидетельствует о наличии экстремума у функции (3.101). Поскольку σ (в выражении (3.101)) является положительной квадратичной функцией, то полученный экстремум соответствует минимуму.

Полученный результат можно обобщить на случай n независимых сил (X_1, X_2, \dots, X_n) , из которых k типа силы (X_1, X_2, \dots, X_k) с помощью определенных внешних воздействий остаются постоянными, чему соответствует постоянство сопряженных с ними потоков (J_1, J_2, \dots, J_k) . В стационарном состоянии (при минимуме производства энтропии) все остальные потоки $(J_{k+1}, J_{k+2}, \dots, J_n)$ становятся равными нулю.

Если ни одна из n независимых сил (X_1, X_2, \dots, X_n) не фиксируется (т.е. отсутствуют внешние возмущающие факторы $k=0$), но считается, что выполняется условие минимума производства энтропии, то тогда все потоки и возникновение энтропии равны нулю и такая система является замкнутой (по определению) и равновесной. Таким образом, стационарное состояние нулевого порядка ($k=0$) соответствует равновесному состоянию изолированной термодинамической системы.

Следует отметить, что принцип минимального производства энтропии справедлив только в случае, когда кинетические коэффициенты L_{ik} постоянны во всем диапазоне сил X_i и удовлетворяют соотноше-

ниям Онзагера $L_{ik} = L_{ki}$ (см. вывод выражения (3.101)). Если эти соотношения не выполняются, то стационарное состояние термодинамической системы реализуется без удовлетворения условия минимального производства энтропии.

Из принципа минимального производства энтропии следует, что при установлении в системе стационарного состояния внутренние неравновесные процессы в системе действуют в направлении, противодействующем приросту энтропии. Из этого следует то обстоятельство, что система, находящаяся в стационарном состоянии, не может выйти из него путем самопроизвольного необратимого изменения своего состояния. Если в результате флуктуации система вышла из стационарного состояния, то в системе возникнут внутренние потоки энергии и вещества, которые стремятся вернуть ее в исходное состояние. Последнее обстоятельство свидетельствует о некоторой устойчивости стационарного состояния термодинамической системы (широко известный на практике принцип Ле Шателье [19]).

3.5. Энтропийные модели микро- и нанотехнологий и структур

3.5.1. Цели и задачи энтропийного подхода при анализе микро- и наноразмерных структур

Для современных интегральных электронных устройств характерна чрезвычайно высокая степень упорядоченности материальных сред, необходимая для обеспечения функционального качества устройства. В практическом плане это обстоятельство приводит к необходимости развития технологической базы, способной реализовать элементы конструкций субмикронных размеров. В этой связи все большую роль приобретают вопросы обеспечения топологического и континуального порядка физических сред, уровень проработки которых в настоящее время недостаточен для практических целей [27]. Достигнутые и перспективные уровни микроминиатюризации оперируют с физическими объемами менее 10^{-6} мкм³, что обуславливает проблематичность использования в конструкторско-технологической практике макроскопических подходов анализа, основанных на методах физики сплошных сред [28]. Состояние отмеченных выше моментов стимулирует разви-

тие нетрадиционных методов анализа технологических процессов, позволяющих объединить микро- и макроскопические подходы [9,10].

Целью настоящего раздела является рассмотрение энтропийных показателей качества микро- и нанотехнологий, основанных на базовых принципах теории информации [7]. Использование энтропийного подхода позволит:

- 1) ввести в конструкторско-технологическую практику универсальные критерии качества технологических процессов и сложности изделий, независимо от функционального назначения и отраслевой направленности последних;
- 2) установить уровень информационного обеспечения технологии создания изделия, необходимый для достижения приемлемых с практической точки зрения производственно-технических показателей по выходу годных изделий;
- 3) обеспечить эффективный диалог между конструктором и технологом изделий на основе единых с методологической точки зрения представлений о процессе создания изделий микро- и наноразмерного масштаба;
- 4) сблизить nano-, микро- и макроскопические подходы и методы анализа материальных сред и функциональных аспектов в конструкторско-технологической практике создания изделий новой техники;
- 5) определить необходимый уровень априорного запаса качества изделия, необходимого для обеспечения заданного срока службы изделия в регламентированных условиях эксплуатации;
- 6) оценить целесообразный (предельно достижимый) уровень обеспечения упорядоченности материальных (атомно-молекулярных) сред в изделии в зависимости от интенсивности воздействующих возмущающих факторов среды эксплуатации;
- 7) разработать в обозримом будущем методы и подходы, необходимые для анализа технологических процессов атомной сборки изделий, с учетом интенсивности проявления элементарных физико-химических процессов флуктуационной природы.

Уместно привести дополнительные пояснения относительно ряда вышеперечисленных пунктов в связи их непосредственного отношения

к конструкторско-технологическим и производственно-техническим аспектам, сопутствующим различным этапам создания изделий новой техники.

Замечания относительно пункта №1. В настоящее время, основываясь на известных аналитических подходах, очень трудно, а порой и просто невозможно, провести сравнительную оценку уровня технологических достижений для изделий не только различной отраслевой направленности, но и в пределах одной отрасли. Это положение обусловлено наличием весьма широкой гаммы технологических методов и аппаратных средств частного характера (как правило, основанных на использовании различных физических принципов), что исключает возможность создания единых с методологической точки зрения критериев оценки качества технологий. Отсутствие подобного рода сравнительных оценок затрудняет рациональное использование материальных ресурсов с целью выравнивания технологического фронта на внутри- и межотраслевом уровне рассмотрения.

Замечания относительно пунктов №2 и 3. Здесь следует отметить следующие два момента.

Во-первых, конструктор и технолог изделия разговаривают на «разных языках». Конструктор ставит во главу угла функциональные параметры изделия (энергетическая экономичность, оптическая эффективность, коэффициент усиления и т.д.) и, соответственно, руководствуется в своей деятельности специфическими научно-техническими дисциплинами типа: теория сигналов, теория механизмов и машин, теория двигателей и т.д. Технолог, в свою очередь, ставит во главу угла вопросы оптимизации процессов обработки материальных сред и, соответственно, руководствуется в своей деятельности дисциплинами типа: химическая термодинамика, физическая химия, физика процессов кристаллизации и т.д. При этом ни тот, ни другой не в состоянии в полной мере (и не ставят такую цель) навести мосты между функциональными параметрами изделия и технологическими режимами процесса их производства. В этой связи технологу в значительной мере приходится использовать эмпирические подходы и это, как правило, «метод проб и ошибок», что сказывается на экономической эффективности его деятельности.

Во-вторых, результатом работы конструктора является, фигурально выражаясь, «бумага», т.е. проект изделия, создание которого, как правило, не связано с большим объемом переработки материальных ресурсов. Ввиду этого инновационная деятельность конструктора новой техники может быть весьма высокой и сдерживается только динамикой развития перспективных научно-технических представлений. Иная ситуация складывается у технолога, инновационная деятельность которого *весьма инерционна*, поскольку требует большие объемы трудозатрат и материальных ресурсов, направленных на создание макетных и промышленных экземпляров аппаратных технологических средств. Поэтому технолог постоянно «отстает» в своих производственных возможностях от технологических требований (претензий) конструктора новой техники. Эта диспропорция решается путем принятия компромисса между технологом и конструктором, и для решения этого вопроса на конструктивной основе желательно знать текущий уровень «дефицита технологии».

Замечания относительно пункта №5. В процессе эксплуатации изделия в нем происходят разнообразные элементарные физико-химические (деградационные) процессы, нарушающие изначально созданный атомный порядок, ответственный за появление функционального назначения изделия. В этой связи необходимость обеспечения заданного срока службы изделия предполагает наличие в нем вполне определенного «запаса функционального качества». Установление количественной меры этого «запаса качества» позволяет прогнозировать ожидаемый срок службы изделия с учетом фактора интенсивности внешних возмущающих воздействий.

Рассматриваемая энтропийная модель и критерии оценки качества технологических процессов применимы к технологическим процессам и изделиям любого типа (вне зависимости от функционального назначения изделия), однако их практическая эффективность тем выше, чем выше сложность изделия в части уровня требований к атомно-молекулярной упорядоченности твердотельных структур. В этой связи приведенные ниже результаты количественных оценок относятся в основном к изделиям нано- и микроэлектроники, для которых характерна

весьма высокая степень упорядоченности материальных сред в элементном и структурном отношениях.

Замечания относительно пункта №7. В свете развития достижений нанотехнологий конструкторами в настоящее время просматривается возможность создания наноразмерных изделий, содержащих в своем составе весьма ограниченное количество атомов (вплоть до одного атома!). В этом случае, однако, следует иметь в виду то немаловажное обстоятельство, что существующие методы анализа материальных сред и физико-химических процессов, основанные на использовании физики сплошных сред (статистическая физика и термодинамика), становятся в принципе неприемлемыми для проведения анализа подобных малообъемных сред. Весьма ограниченное количество атомов, из которых образовано наноразмерное изделие, предполагает использование дискретных принципов анализа состояния этих систем и физико-химических процессов, протекающих в них. Более того, эти принципы должны быть основаны на вероятностных представлениях (учитывая возможность проявления факторов флуктуационной природы), при анализе элементарных физико-химических процессов. В особенности эти моменты относятся к случаю технологий атомной сборки изделий, в которых действительно (де-факто) имеет место заполнение физического пространства посредством размещения индивидуальных атомов. В этом случае на передний план выдвигаются методы *прямого компьютерного моделирования* технологических процессов и систем, в частности, основанные на широком использовании методологии вероятностных клеточных автоматов [10].

3.5.2. Формальное представление технических объектов и технологии их создания

Пусть имеется некоторый объем физического пространства V , в пределах которого необходимо реализовать техническое изделие с заданным функциональным назначением. Не умаляя общности рассмотрения, указанный объем может быть представлен в виде совокупности упорядоченных N ячеек размещения индивидуальных атомов определенного типа.

Имеется также рабочая среда, содержащая $(m + 1)$ типов атомов, как необходимых для заполнения конкретной индивидуальной ячейки размещения, так и нежелательного (примесного) характера (в количестве m типов). Процесс создания изделия состоит в размещении атомов определенного типа в индивидуальных ячейках объема V . *Результат размещения атомов считается удачным, если он обеспечивает появление в изделии необходимого функционального назначения с оговоренным уровнем качества, что возможно при обеспечении определенного порядка в размещении индивидуальных атомов.*

В качестве примера подобного упорядоченного размещения атомов можно рассмотреть диодную структуру, реализованную на основе полупроводникового плавного p - n перехода. На рис. 3.8 представлены (в одномерном приближении) зависимости концентраций донорной $C_d(x)$ и акцепторной $C_a(x)$ примесей от глубины залегания x слоя легирующей примеси в p - n переходе. Характерная особенность p - n перехода состоит в том, что он с термодинамической точки зрения является неравновесной структурой (как и любое другое изделие), которая во временной эволюции испытывает тенденцию к разрушению начальной упорядоченности, созданной с помощью использования технологии заполнения атомами элементарных ячеек объема.

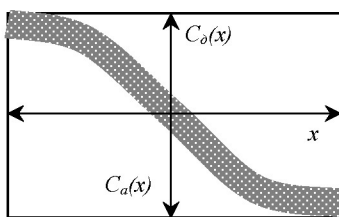


Рис. 3.8. Зависимости концентраций донорной $C_d(x)$ и акцепторной $C_a(x)$ примесей от глубины залегания x слоя легирующей примеси

Темной штриховкой на рисунке обозначена область допустимых (с точки зрения конструктора изделия) вариаций концентрационных зависимостей, обеспечивающих функциональное назначение (качество) диодной структуры. Следует отметить тот факт, что данное изделие является собой невырожденную конструкцию (ансамбль), поскольку в этом

случае необходимое функциональное назначение изделия может быть реализовано чрезвычайно большим количеством частных кривых распределения примесей, находящихся в пределах оговоренной штриховой области (негативным следствием вырожденности конструкции изделия является появление фактора начального разброса функциональных параметров изделия). Последнее обстоятельство облегчает (в ряде случаев существенно) процесс изготовления изделия, ибо допускает определенную нестабильность технологических процессов, а следовательно, способствует повышению выхода годных изделий. *Поэтому с точки зрения практических интересов технолога желательно, чтобы эта область была максимально широкой.*

Таким образом, видно, что интересы конструктора и технолога изделия диаметрально расходятся в части их отношения к допустимому уровню вырожденности изделия и их взаимоотношения должны строиться на компромиссных началах. Следует, однако, отметить, что в настоящее время конструктивная основа для принятия подобного рода компромиссов отсутствует, что переводит эту проблематику в плоскость волевых решений и, как правило, в пользу интересов конструктора изделия, поскольку именно он является инициатором изделий новой техники. Это приводит к тому, что технолог постоянно вынужден «догонять» в своей практической деятельности требования конструктора.

В подобной ситуации (с учетом тенденции непрерывного возрастания сложности изделий) отсутствие единой методологической базы на пути установления упомянутого компромисса выступает в качестве сдерживающего фактора, поскольку конструктор не может предъявить технологу объективную количественную меру необходимого уровня технологического обеспечения при создании изделий новой техники. Это вынуждает технолога «вслепую», методом проб и ошибок, разрабатывать необходимое технологическое обеспечение (как правило, на основании имеющейся у него технологической базы), не располагая возможностью точно оценить несоответствие информационных потенциалов (см. ниже) конструкции и технологии.

В рамках рассматриваемого энтропийного подхода под технологией создания изделия следует понимать *комплекс мер, направленных на по-*

вышение вероятности размещения атомов необходимого типа в определенных ячейках объема V (такой «нейтральный» подход не противоречит интересам и конструктора и технолога, и в то же время сближает их позиции в методологическом плане) [9].

Очевидно, что факту отсутствия технологии как таковой (назовем условно этот случай "случайной технологией") соответствует множество реализаций размещения атомов мощностью

$$N_c = (m + 1)^N, \quad (3.103)$$

где N – количество ячеек размещения в объеме V , необходимых для создания изделия.

Наличие в процессе создания изделия такого фактора как *технология* обеспечивает *избирательный характер процесса размещения атомов* в пределах элементарной ячейки объема V . Фактор избирательности технологии в производственной деятельности позволяет существенно уменьшить фактическое количество различных реализаций укладки атомов до уровня $N_t \ll N_c$. (Необходимо отметить, что понятие избирательности технологии включает в своем составе не только аспекты химического (элементного) состава, но и структурного плана. Последнее обстоятельство предполагает включение в указанную совокупность $(m + 1)$ типов атомов и структурных микрообъектов, например, дефектов кристаллической структуры типа «атомная вакансия»).

Как правило, не все реализации множества N_t обеспечивают необходимое функциональное качество (назначение) изделия. Обозначим посредством N_i совокупности реализаций, обеспечивающих функциональное качество изделия. Тогда, в свете принятого подхода, практическая эффективность технологии создания изделия будет всецело определяться степенью перекрытия множеств реализации N_t и N_i . В частности, при выполнении условия $N_t < N_i$ выход годных изделий составит 100% (если только множества N_t и N_i не смещены относительно друг друга).

На рис. 3.9 представлено условное отображение рассмотренных выше множеств атомных реализаций технологий N_c и N_t , и реализаций N_i , обеспечивающих функциональное назначение изделия.

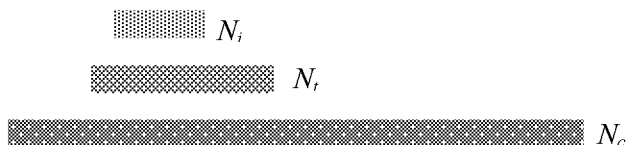


Рис. 3.9 Условное отображение мощности множеств атомных реализаций N_c , N_t и N_i

Согласно представлению рис. 3.9, исходя из целевой установки достижения необходимой вероятности выхода годных изделий можно сформулировать ряд рекомендаций общего характера:

- конструктору изделия необходимо закладывать в конструкцию изделия максимально возможное (с учетом ограничений экономического плана) значение реализаций N_i , обеспечивающих функциональное назначение изделия;
- технологу изделия необходимо закладывать в производственно-технологическую практику создания изделия значение избирательности технологии, обеспечивающее выполнение условия $N_t \leq N_i$ (знак «меньше» свидетельствует о наличии определенного запаса функционального качества изделия, который расходуется в процессе его эксплуатации, ввиду протекания неизбежных процессов деградации в процессе срока службы изделия).

Следует отметить, что в современной производственной практике создания изделий *высоких технологий* (к ним можно отнести изделия микроэлектроники, микрооптики, точной механики, приборостроения), в которых усложнение конструкции идет опережающими темпами относительно развития уровня технологии (в последнем случае имеются в виду инерционность и необходимость существенных затрат на развитие технологии), как правило, выполняется условие $N_t > N_i$, что и обуславливает вероятность выхода годных изделий существенно меньше единицы.

Допуски размерного характера (помимо факторов структурного и элементного плана) могут играть ключевую роль в вопросах выбора оптимальной с точки зрения *экономических* показателей технологии,

используемой при изготовлении изделия. К примеру, при изготовлении следующих изделий или элементов конструкций:

- гвоздей (при допуске на диаметр порядка ± 0.2 мм, что составляет $2 \cdot 10^6$ атомных монослоев материала), для производства которых можно, в принципе, использовать и токарную обработку материалов (точность обработки материалов: ± 10 мкм, т.е. 10^5 атомных монослоев), что обеспечит уверенное выполнение условия $N_i \ll N_i$ (выход годных 100%) за счет увеличения запаса качества до нецелесообразного уровня и в ущерб себестоимости производства изделия;

- «деталей» современных полупроводниковых интегральных схем (при требуемом допуске порядка ± 0.01 мкм, что составляет 10^2 атомных монослоев материала), для производства которых используются технологии, обеспечивающие реализацию допуска порядка ± 0.02 мкм, что приводит к выходу годных изделий существенно менее 100 %, а следовательно, и к потерям экономического характера, обусловленным «запланированным» технологией изготовления изделия появлением бракованной продукции.

Приведенные примеры свидетельствуют о том, что к вопросам взаимного соответствия уровня сложности изделия (закладывается конструктором на этапе проектирования изделия) и уровня технологии, достаточного для достижения приемлемых экономических показателей (обеспечивается технологом на этапе производства изделия), необходимо относиться с предельным вниманием.

3.5.3. Показатели качества технологий на микроскопическом уровне рассмотрения

Рассмотрим процесс размещения атома определенного типа в индивидуальной пространственной ячейке атомного масштаба объема как *случайное событие*. В этом случае, в соответствии с основополагающими положениями теории информации [7], для информационной энтропии процесса размещения единичного атома в соответствующей индивидуальной ячейке имеем (см. разд. 3.1):

$$h_t = - \sum_{i=1}^{(m+1)} p_i \cdot \lg p_i, \quad (3.104)$$

где h_t – энтропия размещения единичного атома при наличии технологии; p_i – вероятность размещения в индивидуальной ячейке атома i -го типа.

Выражение (3.2) является базовым для последующего рассмотрения энтропийного подхода в вопросах оценки качества микро- и нанотехнологий. Однако его практическое использование сопряжено с трудностями, обусловленными необходимостью определения значений параметров p_i для всей совокупности атомов примесного характера. Эту трудность в первом приближении можно обойти следующим образом. Введем в рассмотрение понятие *избирательности технологии α в виде отношения вероятности размещения необходимого атома к суммарной вероятности размещения остальных m типов атомов*, выступающих в качестве нежелательной примеси. Следует отметить, что для современного уровня развития технологии обработки материалов вероятность размещения необходимого атома $P \approx 1$. В этой связи количественную оценку параметра α можно в первом приближении провести согласно выражению

$$\alpha = \frac{n_0}{n}, \quad (3.105)$$

где n_0 – полная концентрация атомов в технологических рабочих средах, включая как необходимые атомы, так и атомы примесного характера; n – полная (по всей совокупности типов примесных атомов) концентрация атомов примесного характера.

С учетом допущения о равной вероятности размещения атомов примесного типа и выражения (3.2) для энтропии размещения единичного атома имеем [9]

$$h_t = \frac{\lg m + (\alpha + 1) \cdot \lg(\alpha + 1) \cdot \alpha \cdot \lg \alpha}{\alpha + 1}. \quad (3.106)$$

Примем в качестве показателя качества технологии на микроскопическом (атомном) уровне рассмотрения параметр

$$K = \frac{h_c - h_t}{h_c}, \quad (3.107)$$

где $h_c = \lg(m+1)$ – энтропия единичного размещения атома при случайной технологии.

В табл. 3.6 представлены значения рассмотренных показателей (K , α , h_b , h_c) при значении параметра $m = 100$, характерные для различных уровней технологии (определение параметра α проведено на основании предельно допустимой концентрации атомов примесного характера, согласно выражению (3.3)).

Первые две позиции табл. 3.6 характеризуют соответственно перспективный (нано-) и достигнутый (микро-) уровни технологий, используемых при создании современных интегральных схем.

Таблица 3.6. Зависимость избирательности технологии α , энтропии единичного размещения h_i и показатели качества технологии K от уровня технологии (при $m=100$)

№	Уровень технологии		α	h_i	K
1	Нанотехнология (нанoeлектроника)		$10^9 - 10^{10}$	1×10^{-9}	$0,9^{(9)}$
2	Микротехнология (микроелектроника)		$10^7 - 10^8$	1×10^{-7}	$0,9^{(7)}$
3	Макротехнология (машиностроение)		$10^4 - 10^6$	8×10^{-4}	0,9996
4	Случайный (отсутствие технологии)		10^{-2}	2 . 004	0
5	Биологический: (объект – ДНК)	а) уровень нуклеотида	$>10^{10}$	$<1 \times 10^{-9}$	$>0,9^{(9)}$
		б) атомный уровень	$>10^{12}$	$<1 \times 10^{-11}$	$>0,9^{(11)}$
6	Критический (обусловленный влиянием проявлений флуктуационного характера)				
	а) $W_a = 1$ эВ	$T = 300$ К	10^{21}	2×10^{-20}	$0,9^{(20)}$
		$T = 600$ К	10^9	1×10^{-8}	$0,9^{(8)}$
	б) $W_a = 0,5$ эВ	$T = 300$ К	10^9	1×10^{-8}	$0,9^{(8)}$
$T = 600$ К		10^5	1×10^{-5}	$0,9^{(4)}$	

Примечание: запись вида $0,9^{(3)}$ означает 0,999.

Данные позиции №3 позволяют оценить уровень технологий, используемых в микроелектронике, по отношению к возможностям макротехнологий, нашедших применение в других отраслях техногенной деятельности.

Данные позиции № 5 (исходные данные взяты из источника [29]) являются своеобразными ориентирами показателей качества антропогенных технологий. Сравнивая показатели избирательности α по пози-

циям №1 и №5 можно сделать вывод о том, что современный уровень нанотехнологий по своим показателям качества весьма близко отстоит от уровня природных «биологических технологий». Последнее обстоятельство становится понятным, если учесть, что одним из направлений развития современной нанотехнологии является *технология атомной сборки изделий*.

В позиции № 6 приведены значения предельно достижимых энтропийных показателей в зависимости от показателей температурных режимов проведения технологических процессов и энергии активации элементарных физико-химических процессов. При определении количественных показателей этой позиции использовано допущение о термоактивационном механизме протекания элементарных физико-химических процессов (см. разд. 3.3). Данные этой позиции свидетельствуют о недопустимости использования в нанотехнологиях высокотемпературных режимов обработки материалов. Эти же данные свидетельствуют о проблематичности достижения избирательности $\alpha > 10^8$ (характерной для производства устройств современной микроэлектроники) при температурах проведения технологических процессов более 600 К и значений энергий активации элементарных процессов (с учетом фактора одновременного протекания процессов деградации на этапе производства изделия) менее $0,5 \div 1,0$ эВ.

Данные позиции № 4 («случайная технология») приведены исключительно для сравнения и позволяют оценить абсолютный уровень развития различных типов (макро-, микро-, нано-) современных технологий.

3.5.4. Показатели качества технологий на макроскопическом уровне рассмотрения

Непосредственный анализ энтропийных показателей качества на макроскопическом уровне рассмотрения сопряжен со значительными трудностями, обусловленными необходимостью учета *условных энтропий* всевозможных реализаций атомных последовательностей длиной N . Однако учитывая тот факт, что в реальной технологической практике создания элементов конструкций микро- и наноразмерных устройств

значение параметра α удовлетворяет условию $\alpha > 10^7$ ($N > 10^6$, при минимальном объеме элементов устройства более 10^{-6} мкм³), можно считать, что *технологический процесс является энтропийно устойчивым*.

Принимая во внимание допущение о наличии фактора энтропийной устойчивости технологии, можно, согласно теореме об асимптотической равновероятности индивидуальных случайных реализаций (см. разд. 3.2.1, теорему №7: выражения (3.26), (3.27), (3.28)), полное множество возможных реализаций, включающих в себя цепочку из N атомов (элементов), представить в виде суммы двух частных подмножеств:

- одно из этих подмножеств, со значительной мощностью множества, имеет практически нулевую суммарную вероятность его проявления;
- другое подмножество, с существенно меньшей мощностью множества, состоит из достаточно ограниченной совокупности равновероятных реализаций.

Таким образом, для энтропийно устойчивого технологического процесса, используемого при создании атомной структуры, содержащей N элементов (атомов), имеем

$$H_t = N \cdot h_t = \lg N_t, \quad (3.108)$$

где N_t – мощность множества равновероятных реализаций с учетом фактора избирательности технологии.

В табл. 3.7 представлены значения энтропии H_t в зависимости от физического объема устройства и уровня технологии, используемого при его создании.

Т а б л и ц а 3.7. *Зависимость энтропии макроскопических объектов H_t от их объема и уровня технологии (при $m=100$)*

№	Уровень технологии	Объем устройства			
		1мм ³	1мкм ³	10 ⁻³ мкм ³	10 ⁻⁶ мкм ³
1	Перспективный (нанотехнология – $\alpha = 10^{10}$)	1·10 ¹⁵	1·10 ³	1·10 ⁰	1·10 ⁻³
2	Высокий (микротехнология – $\alpha = 10^8$)	1·10 ¹⁷	1·10 ⁵	1·10 ²	1·10 ⁻¹
3	Случайный ($\alpha = 10^{-2}$)	2·10 ²⁴	2·10 ¹²	2·10 ⁹	2·10 ⁶

Использование допущения в форме выражения (3.108) приводит к равенству показателей качества технологии на микро- и макроскопических уровнях рассмотрения. Представляется очевидным, что в реальной производственной практике показатель качества технологии на макроскопическом уровне рассмотрения должен быть непосредственно связан со сложностью изделия (см. ниже выражение (3.109)). С позиций энтропийного подхода в качестве объективной меры сложности устройства C следует принять показатель вида

$$C = \frac{H_c - H_i}{H_c}, \quad (3.109)$$

где $H_c = \lg N_c$ – энтропия в отсутствие технологии как таковой; $H_i = \lg N_i$ – энтропия реализаций, обеспечивающих функциональное качество устройств.

Определение значения параметра N_i относится к проблематике конструктора изделия. Следует, однако, отметить, что в настоящее время конструктор не готов к такой постановке вопроса и оперирует в своей деятельности макроскопическими параметрами (материальных сред и устройств) термодинамического характера. Подобный подход приемлем для устройств и технологий их создания при значениях избирательности $\alpha < 10^4$, для которых не играют существенной роли эффекты флуктуационной природы. В противном случае становится проблематичным использование в конструкторско-технологической практике представлений континуального характера. В этой связи возможность объединения микро- и макроскопических подходов в рамках единой методологии рассмотрения является позитивным моментом энтропийных моделей технологии создания изделий, в особенности для устройств микро- и нанотехнологии.

3.5.5. Показатели качества реальных микро- и нанотехнологий

Ввиду сложности определения параметра N_i прямыми теоретическими (конструктором изделия) и экспериментальными методами, проведем его оценку на основе использования традиционных производственных показателей. Из всей совокупности последних в качестве параметра, определяющего взаимосвязь между параметрами H_c и H_i ,

уместно рассматривать *вероятность выхода годных* изделий. Представляется очевидным, что в первом приближении вероятность выхода годных изделий P зависит от соотношения между избирательностью технологии α_t и избирательностью α_i , заложенной конструктором изделия для безусловного ($P = 1$) достижения функционального качества устройства (на практике, как правило, выполняется условие $\alpha_t < \alpha_i$):

$$P = \frac{\alpha_t}{\alpha_i}. \quad (3.110)$$

С учетом выражения (3.106) и (3.110) для необходимой энтропии единичного размещения h_i , при условии $\alpha_i \gg 1$ (что имеет место для случаев микро- и нанотехнологий) имеем

$$h_i = \frac{P}{\alpha_i} \cdot \lg \frac{m \cdot \alpha_t}{P}. \quad (3.111)$$

В табл. 3.8 представлены значения параметров α_i и h_i , в зависимости от уровня технологии и вероятности выхода годных устройств при значении параметра $m = 100$.

Т а б л и ц а 3.8. *Зависимость избирательности и энтропии единичного размещения h_i от уровня технологии и вероятности выхода годных устройств (при $m=100$)*

№	Уровень технологии	Показатель	Вероятность выхода годных устройств				
			10^{-3}	10^{-2}	10^{-1}	0,5	1,0
1	Перспективный (нано: $\alpha_t = 10^{10}$)	h_i	$1,5 \cdot 10^{-12}$	$1,4 \cdot 10^{11}$	$1,3 \cdot 10^{-10}$	$6,2 \cdot 10^{-10}$	$1,2 \cdot 10^{-9}$
		α_i	10^{13}	10^{12}	10^{11}	$2 \cdot 10^{10}$	10^{10}
2	Высокий (микро: $\alpha_t = 10^8$)	h_i	$1,3 \cdot 10^{-10}$	$1,2 \cdot 10^{-9}$	$1,1 \cdot 10^{-8}$	$5,2 \cdot 10^{-8}$	10^{-7}
		α_i	10^{11}	10^{10}	10^9	$2 \cdot 10^8$	10^8

Согласно данным табл. 3.8 следует, что достижение приемлемой в промышленной практике эффективности производства ($P > 0,1$) возможно для изделий с показателями $\alpha_i < 10^{11}$ и 10^9 соответственно для перспективного и высокого уровней технологии.

В табл. 3.9 представлены значения N_i в зависимости от объема и вероятности выхода годных устройств, полученные на основании данных табл. 3.8.

Т а б л и ц а 3.9. *Зависимость количества реализаций N_i от физического объема и вероятности выхода годных устройств (при $n=100$)*

№	Уровень технологии	Объем, мкм ³	Вероятность выхода годных устройств				
			10^{-3}	10^{-2}	10^{-1}	0,5	1,0
1	Перспективный (нанотехнологии: $\alpha_i = 10^{10}$)	1	31	(14)	(130)	(620)	(10^3)
		10^{-3}	1,00	1,03	1,34	4,23	15,8
		10^{-6}	1,00	1,00	1,00	1,00	1,00
2	Высокий (микротехнологии: $\alpha_i = 10^8$)	1	(130)	(10^3)	(10^4)	($5 \cdot 10^4$)	(10^5)
		10^{-3}	1,34	15,8	(11)	(52)	(10^2)
		10^{-6}	1,00	1,00	1,02	1,12	1,26

Примечание: запись вида (14) означает, что $\lg N_i = 14$.

Наличие в табл. 3.9 значений N_i близких единице свидетельствует о том, что для данных позиций конструктором закладывается, а технологю необходимо реализовать (при приемлемой вероятности выхода годных изделий) практически *вырожденные конструкции*, которым соответствуют единичные реализации размещения атомов. В то же время следует отметить, что для современных устройств ($V > 1$ мкм³) количество реализаций N_i при $P > 0,1$ достигает значений $N_i > 10^{130}$.

В свете вышеизложенного весьма важным моментом является установление взаимосвязи между сложностью изделий (закладываемой конструктором на этапе проектирования) и возможностью их практической реализации технологом. Указанная взаимосвязь может быть установлена на основании выражений (3.107), (3.109) и (3.110) в виде

$$P = \frac{1-C}{1-K}. \quad (3.112)$$

Последнее выражение устанавливает эффективность согласования конструкторских решений (параметр C) с технологическими возможностями изготовления (параметр K) устройств для типичных производственных ситуаций, удовлетворяющих условию $H_t > H_i$ (Это условие позволяет обеспечить нормировку параметра P в виде $0 < P < 1$. При выполнении условия $H_t \leq H_i$ имеет место тривиальный результат: $P = 1!$).

3.5.6. Информационный потенциал и информационный дефицит технологии

В повседневной практике часто встречаются ситуации, когда отсутствие информации затрудняет принятие оптимального решения. В этих случаях принято говорить о недостаточности информационного потенциала и, соответственно, о наличии информационного дефицита.

В соответствии с общими концепциями теории информации [1] в качестве *информационного потенциала технологий* уместно рассматривать параметр вида

$$I_t = H_c - H_t, \quad (3.113)$$

где, в отличие от вышеизложенного, целесообразно рассматривать информационные энтропии H_c и H_t в удельном представлении (т.е. для единицы объема преобразуемого пространства, а не фактического объема изделия). В этом случае параметр I_t выступает в качестве объективной меры качества процесса преобразования *единичного объема физического пространства*. Привязка к единице объема пространства позволяет объективно судить об абсолютном уровне качества различных типов технологий как таковых, а не во взаимосвязи со сложностью конкретных изделий, и, соответственно, проводить сравнительную оценку качества индивидуальных технологий (в том числе на межотраслевом уровне).

В такой постановке вопроса параметр I_t выступает в качестве объективной меры абсолютного уровня достижений технологии *в историческом разрезе*, т.е. с учетом всего практического опыта, накопленного человечеством до настоящего момента времени. Следует отметить, что этот абсолютный уровень испытывает значительные вариации в зависимости от отраслевой направленности технологий (см. данные табл. 3.6).

Однако в повседневной производственной практике технолога интересует не столько информационный потенциал I_t , сколько *информационный дефицит технологии* ΔI_t , ограничивающий выход годных для конкретного типа изделий (а не единичного объема преобразования

пространства). С учетом конкретики изделия информационный дефицит технологии ΔI_t может быть определен в виде

$$\Delta I_t = H_t - H_i, \quad (3.114)$$

где H_t и H_i – энтропийные характеристики, отнесенные к изделию (а не единицы объема).

Представляется очевидным, что если выполняется условие $H_t = H_i$, то имеется полное соответствие уровня технологии целям производства данного конкретного изделия и информационный дефицит технологии становится равным нулю.

С учетом выражений (3.107) – (3.109) и (3.112) информационный дефицит технологии можно связать с вероятностью выхода годных изделий с характеристическим параметром изделия N :

$$\Delta I_t = N \cdot h_t \cdot (1 - P). \quad (3.115)$$

Смысл параметра ΔI_t состоит в том, что он характеризует то количество информации, которое необходимо задействовать в технологии создания конкретного изделия, чтобы повысить вероятность выхода годных изделий от некоторого начального уровня $P < 1$ до предельно достижимого уровня $P = 1$.

Устранение информационного дефицита (с целью повышения вероятности выхода годных изделий) достигается посредством расширения *метрологического обеспечения технологии*. Под этим подразумевается не только расширение диагностической базы, непосредственно используемой при проведении технологических процессов, но и совокупность тех метрологических мероприятий, которые были задействованы при создании новых комплексов технологического оборудования.

Минимальное количество индивидуальных метрологических операций (всех типов), необходимых для ликвидации информационного дефицита текущей технологии (обладающей изначально энтропийным показателем качества h_t) при производстве изделия, можно оценить с помощью выражения

$$n = \frac{\Delta I_t}{I_1}, \quad (3.116)$$

где I_1 – среднее количество информации, получаемой с помощью единичной метрологической операции.

Устанавливая параметр I_1 в виде (подробнее см. в разд. 3.2.2 [3])

$$I_1 = \lg\left(\frac{L}{l}\right), \quad (3.117)$$

где L и l – соответственно диапазон измерения и погрешность измеряемого технологического параметра, и используя выражения (3.115) и (3.116), получаем количество метрологических операций n , которыми надо *дополнительно оснастить* технологию для безусловного достижения 100% выхода годных изделий:

$$n = \frac{N \cdot h_i \cdot (1 - P)}{\lg(L/l)}. \quad (3.118)$$

Для проведения количественных оценок можно использовать усредненное значение $\lg(L/l) = 6$, характерное для всех отраслей технической деятельности. Например, этому значению логарифма соответствует значение абсолютной погрешности 1 μA при измерении силы тока в интервале значений силы тока 0 – 1 А (подробнее этот вопрос рассмотрен в подразд. 3.2.2, выражение (3.35)). Параметр n определяет меру необходимых *дополнительных информационных потоков* в системе «рабочая среда – технология – изделие» с целью обеспечения *негоэнтропийного* характера технологии создания техногенных объектов.

В качестве иллюстрации применения выражения (3.118) для практических целей рассмотрим следующий пример. Пусть имеется изделие, реализуемое в объеме пространства $V = 1 \text{ мкм}^3$ с вероятностью выхода годных $P = 0,5$ (при начальном уровне технологии). Тогда, преследуя цель достижения $P = 1$, для устранения информационного дефицита технологии потребуется количество метрологических операций n порядка 10^2 и 10^4 соответственно для перспективного и высокого уровней технологии.

3.5.7. Информационный запас качества и его взаимосвязь со сроком службы изделия

Атомную структуру изделия можно рассматривать как своеобразную матрицу памяти, в которой «записано» функциональное назначение

ние изделия. Функциональное качество этой «записи» определяется начальной физической (конфигурационной) энтропией $S_0 = k N_i$, где k – постоянная Больцмана.

В процессе эксплуатации изделия в нем непрерывно протекают процессы деградации на атомном уровне, что приводит к возрастанию физической энтропии изделия и, следовательно, к «стиранию» начальной функциональной памяти (процесс деградации можно рассматривать как процесс нежелательной «технологии» с энтропийными показателями, представленными в позиции № 6 табл. 3.6). В результате протекания процессов деградации происходит снижение степени упорядоченности как специфики пространственного расположения атомов в изделии, обеспечивающей появление функционального назначения (качества) изделия.

Скорость процесса деградации изделия (по критерию его функционального качества) всецело зависит от интенсивности протекания нежелательных элементарных физико-химических процессов, приводящих к изменению элементного состава и атомно-молекулярной структуры материалов в элементах конструкций изделия. В свою очередь интенсивность упомянутых элементарных физико-химических процессов всецело зависит от энергии активации W_a и температурных режимов (если речь идет о термоактивационном механизме деградации) эксплуатации изделия. Подробнее вопросы кинетики протекания (по термоактивационному механизму) элементарных физико-химических процессов представлены в подразд. 3.3. Уместно отметить ключевую роль статистических факторов в вопросах кинетики протекания процессов деградации.

Для обеспечения длительного срока службы изделия необходимо, чтобы в изделии был изначально заложен определенный информационный запас качества, который будет расходоваться в процессе эксплуатации изделия.

Суммарное возрастание энтропии изделия ΔS , имеющее место в процессе эксплуатации изделия, может быть определено с помощью параметра «скорости производства энтропии» σ (подробнее см. в подразд. 3.4.2) [5]:

$$\Delta S(\tau) = \int_0^{\tau} \sigma dt, \quad (3.119)$$

где τ – длительность (срок службы изделия) процесса эксплуатации.

Введя в рассмотрение понятия о *критических уровнях* физической и информационных энтропий $S_k = k \cdot H_k$, при которых наступает полная потеря функционального качества изделия, и используя условие

$$\Delta S = S_k - S_0, \quad (3.120)$$

можно определить ожидаемый срок службы изделия в виде следующего уравнения интегрального типа:

$$k \cdot (H_k - H_t) = \int_0^{\tau} \sigma dt. \quad (3.121)$$

Таким образом, параметр $\Delta I_k = H_k - H_t$ выступает в качестве *информационного запаса качества изделия*, обеспечивающего его функциональное качество на протяжении срока службы изделия.

Следует отметить, что в настоящее время практическое использование выражения (3.121) может быть распространено только на ряд частных случаев, для которых известны (как правило, с помощью экспериментальных методов исследования) кинетические коэффициенты L_{ik} (см. подразд. 3.4.2) в рамках модели линейных неравновесных процессов. В этом случае, как было показано выше, скорость производства энтропии (в удельном представлении) определяется с помощью выражения

$$\sigma = \sum_i J_i \cdot X_i, \quad (3.122)$$

где

$$J_i = \sum_{k=1}^n L_{ik} \cdot X_k. \quad (3.123)$$

Проблематика использования выражений (3.122) и (3.123) усугубляется тем немаловажным обстоятельством, что в настоящее время, как правило, *конструктор изделия не располагает уверенными данными относительно пространственного распределения термодинамических сил X_i в пределах объема изделия.*

Рассмотренные выше представления относятся к случаю *«постепенного отказа»* изделия, характерной особенностью которого является непрерывное монотонное накопление микродефектов атомно-молекулярного масштаба. (Помимо этого типа отказов, на практике имеют место и *«внезапные отказы»*, появление которых обусловлено либо грубыми просчетами в конструкции или технологии производства изделия, либо существенным нарушением штатных режимов эксплуатации изделия. Характерной особенностью внезапного отказа изделия является его чрезвычайно высокая скоротечность ($\tau = 10^{-10} - 10^{-3}$ с), что обусловлено лавинообразным характером протекания процессов деградации изделия (наиболее ярким примером является электрический пробой конструкционных материалов в сильных электрических полях).

3.5.8. Процесс создания микроструктур в рамках представлений алгоритмической энтропии

Как упоминалось выше, упорядоченные атомные микроструктуры можно рассматривать как своеобразную матрицу памяти, в которой «записано» функциональное назначение изделия. С этой точки зрения можно считать, что в атомной структуре проведено кодирование определенной дискретной информации посредством упорядоченного (с точки зрения реализации необходимого функционального назначения) размещения атомов. Подобный подход позволяет при проведении анализа энтропийных параметров микроструктур использовать теорию кодирования дискретной информации. Возможность использования теории кодирования дискретной информации основана на том факте, что реальные микро- и наноструктуры могут быть представлены в виде совокупности отдельных атомно-молекулярных слоев (текстов), при этом позиция каждого атома в мономолекулярном слое может рассматриваться в качестве элемента сообщения (слова или буквы).

На рис. 3.10 представлено схематическое отображение возможного способа информационного «кодирования» твердотельной структуры на примере упорядоченного расположения индивидуальных атомов в мономолекулярном слое.

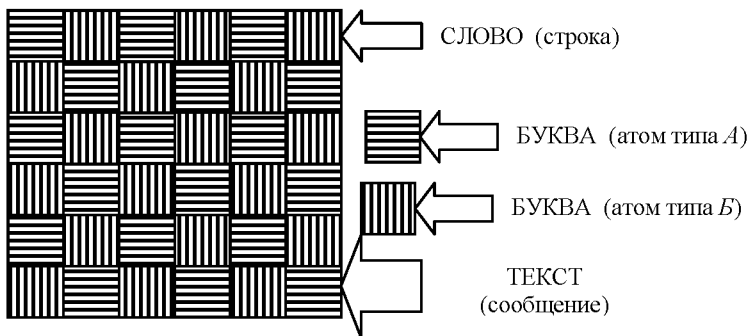


Рис. 3.10. Схематическое отображение фрагмента структуры изделия в рамках представлений теории кодирования сигналов

При выполнении кодирования предполагается заданной последовательность (как правило, большая) сообщений (объектов) $\xi_1, \xi_2, \dots, \xi_N$ вместе со своими вероятностями появления, т.е. как последовательность случайных величин. Следовательно, может быть вычислена соответствующая ей энтропия (и количество информации H , содержащееся в этой случайной последовательности, воспринимаемой как «сообщение»).

Эта информация может быть записана различными способами записи (см. рис. 3.10), в том числе, как это приведено для рассматриваемого случая – спецификой размещения индивидуальных атомов в пределах подложки (изделия).

Представленное выше определение количества информации целесообразно при рассмотрении процессов преобразования информации из одного вида в другой, т.е. при проведении процедуры кодирования информации (технологический процесс в этой расширенной трактовке также можно рассматривать в качестве своеобразного кодирования структуры изделия на атомном уровне физической реализации).

Существенно, что при таком преобразовании выполняется закон сохранения количества информации. Здесь можно провести аналогию с законом сохранения энергии. (Этот закон является ключевым основанием для введения понятия энергии). Однако закон сохранения информации сложнее закона сохранения энергии в двух отношениях. Закон сохранения энергии устанавливает точное равенство энергий при их

взаимопревращениях в процессе эволюции физической системы. При превращении же информации имеет место более сложное соотношение, а именно «не больше (\leq)», т.е. информация не может возрасти при проведении кодирования.

Знак равенства соответствует случаю оптимального кодирования. Поэтому при формулировке закона сохранения информации приходится говорить лишь о возможности существования такого кодирования. Таким образом, должно иметь место равенство количества информационного обеспечения технологии (применительно к рассматриваемому случаю речь может идти об адекватном соответствии информационного потенциала технологии в части совершенства оборудования и используемых материалов) и информационного потенциала изделия, изначально заложенного конструктором на этапе проектирования.

Второе обстоятельство состоит в том, что равенство является приближенным (а не точным), асимптотическим, т.е. справедливым для достаточно сложных сообщений (объектов). Точный знак равенства возможен в предельном случае. В последнем случае имеется полная аналогия с законами статистической термодинамики, которые справедливы для систем, состоящих из очень большого числа атомов. Информация может быть записана различными реализациями, в особенности если речь идет о невырожденной (см. выше) конструкции изделия. Если N – число таких реализаций, то указанный закон сохранения количества информации заключается в равенстве $H = \ln M$, осложняемом упомянутыми двумя обстоятельствами (в действительности $H < \ln M$).

Методология анализа качества технологических процессов, основанная на принципах кодирования случайных последовательностей, может быть распространена на все типы производственных процессов, но ее применение представляется более предпочтительным для случая технологий атомной сборки. Примем, что дискретный процесс создания изделия по технологии атомной сборки является стационарным, т.е. все законы распределения $P\{\xi_1, \dots, \xi_r\}$ (произвольной кратности r) не меняются при сдвиге:

$$P\{\xi_1, \dots, \xi_r\} = P\{\xi_{k_1+a}, \dots, \xi_{k_r+a}\}, \quad (3.124)$$

где a – любое натуральное число.

Величина сдвига a предполагается такой, что значения $\{k_1+a, \dots, k_r+a\}$ не выходят за пределы области K определения параметра k . Предполагается, что эта область определения параметра не ограничена в обе стороны.

В практическом плане понятие стационарности оговоренных выше, случайных последовательностей ассоциируется с представлениями об устойчивости (стабильности) реальных технологических процессов в производственной практике создания изделий и перекликается с введенными ранее представлениями об энтропийной устойчивости случайных процессов. Выполнение этого условия позволяет рассмотреть процесс создания изделия на основе методологии теории информации в части анализа процессов «передачи сообщений» при наличии помех и пропускной способности каналов связи. В последнем случае под пропускной способностью следует понимать взаимосвязь элементов из всей совокупности информационного обеспечения технологических процессов производства изделия.

3.5.9. Алгоритмическая энтропия и самоорганизация сложных атомных структур

В последнее время все настойчивее проявляется тенденция к использованию технологий атомной сборки применительно к процессам создания наноразмерных изделий (структур). При этом предлагается использовать механизмы самосборки изделий на атомно-молекулярном уровне обеспечения технологических процессов. Сочетание этих двух моментов сближает позиции в понимании понятия «технологический процесс» применительно к созданию как технических, так и биологических объектов (изделий). В этой связи уместно рассмотреть проблематику технологических процессов атомной самосборки наноразмерных изделий на примере формальных аспектов механизмов создания и эволюции элементарных биологических структур – «биополимеров». Рассмотрение нижеследующих материалов преследует цель ознакомления с общими методологическими принципами кинетического анализа, проводимого с привлечением фактора самоорганизации систем, нежели с достижениями в области теории биологической эволюции.

Рассмотрим [30] множество линейных объектов (в частности цепных макромолекул), построенных из элементарных «кирпичиков» А, В, С, D (в качестве последних могут выступать как индивидуальные атомы, так и отдельные молекулярные образования). При длине линейного объекта v (количество элементов цепи) возможны 4^v различных реализаций последовательностей. Если допустить, что длина цепи составит 10^6 «кирпичиков» (характерное значение для биополимеров), то число возможных последовательностей оказывается столь огромным и открывает такое число возможностей, которое за 10^{18} секунд существования Вселенной не могло быть реализовано и вполне может считаться «практически бесконечным».

Рассмотрим далее кинетическую модель эволюции системы последовательностей (например: {DABC}, {BDABD}, {DDABCD}), которая в результате взаимодействия с внешней средой могут самовоспроизводиться с некоторой определенной скоростью (в том числе с помощью копирования).

Дополнительно примем:

- при копировании в редких случаях могут спонтанно возникать случайные ошибки (мутации), из-за которых копия будет немного отличаться от оригинала;
- существует возможность присоединения к последовательности справа или слева одного из четырех мономеров (кирпичиков).

Далее на систему накладывается специфическое условие отбора (выступающего в качестве элемента самоорганизации):

- либо ограничивается общее число существующих в системе полимеров (проблема существования);
- либо ограничивается приток мономеров (кирпичиков) из окружающей среды (проблема сырья).

Множество всех последовательностей произвольной длины можно упорядочить, пронумеровав их натуральными числами. Из многих возможных способов выберем следующую процедуру условной нумерации i :

i	Тип последовательности
0	Ö – нулевая последовательность
1-4	A, B, C, D – единичная последовательность
5-20	AA, AB, AC, AD, BA, BB, BC, BD, CA, CB, CC, CD, DA, DB, DC, DD,
21-32	AAA, AAB, AAC, AAD, ABA, ABB, ABC, ABD, ACA, ACB, ACC, ACD и т. д.

Изучение этих последовательностей молекул тесно связано с задачами кодирования в теории информации. Молекулярные «кирпичики» A, B, C, D соответствуют символам четырехбуквенного алфавита. В теории информации всякую конечную последовательность букв принято называть словом, поэтому и в рассматриваемом случае последовательности молекул являются «молекулярными словами». Пользуясь такой формальной аналогией между цепными молекулами и «словами» из теории информации, можно позаимствовать многие понятия и теоремы теории кодирования.

Возьмем, например, понятие расстояния $d(i, j)$ между двумя последовательностями (реализациями): $i = \{z_1, z_2, \dots, z_k\}, j = \{y_1, y_2, \dots, y_L\}$, где $z_n, y_n \in \{A, B, C, D\}$, которое определяется как минимальное число несовпадений двух последовательностей при всевозможных наложениях их друг на друга. Для определения расстояния между последовательностями $i = \{DACB\}$ и $j = \{BDADB\}$ построим всевозможные наложения и проведем сравнение с целью выявления совпадающих фрагментов:

D	A	C	B	-	-	-	D	A	C	B	-	-	-	D	A	C	B
<u>B</u>	<u>D</u>	<u>A</u>	<u>D</u>	<u>B</u>	<u>B</u>	<u>B</u>	<u>D</u>	<u>A</u>	<u>D</u>	<u>B</u>	<u>B</u>	<u>B</u>	<u>D</u>	<u>A</u>	<u>D</u>	<u>B</u>	<u>B</u>

и, подсчитав минимальное число несовпадений, получим, что $d(i, j) = 3$.

Еще более важным представляется введенное Колмогоровым [31] понятие об алгоритмической энтропии $K(i)$ последовательности i , которую можно понимать как меру сложности последовательности. Понятие «сложность» системы (по Колмогорову) имеет очень большое значение для практических целей, поскольку оно позволяет *математически строго отличать упорядоченные и случайные последовательности*. Основная идея Колмогорова состоит в том, что *последовательность*

считается случайной, если ее нельзя построить с помощью программы, которая была бы короче самой последовательности. Таким образом, «сложность», образно говоря, есть мера длины кратчайшей программы, задающей последовательность. Другими словами, сложность последовательности i событий (например: размещения последовательности индивидуальных атомов) есть длина кратчайшего описания последовательности на языке L (имеется в виду конкретная версия используемого машинного языка).

К сожалению, общий алгоритм вычисления $K(i)$ для данной последовательности i неизвестен, однако к искомому минимуму можно приблизиться, сравнивая длину различных программ, представляющих последовательность. Понятие алгоритмической энтропии является существенным обобщением понятия информационной энтропии Шеннона, поскольку на уровне интуиции ясно, что параметр $K(i)$ можно рассматривать как меру содержащейся в последовательности информации. Параметр сложности системы $K(i)$ определяет меру упорядоченности множества последовательностей и показывает, увеличивается или уменьшается количество информации при переходе от последовательности $i^{\text{го}}$ типа к последовательности $j^{\text{го}}$ типа (например, в результате процессов мутации). Принято считать, что только сочетание представлений Эйгена [30] и Колмогорова [31] может помочь ответить на вопрос, каким образом в природе могли возникнуть молекулы нуклеиновых кислот, в которых закодирован большой объем информации (этот момент представляет большой практический интерес для случая самоорганизации техногенных наноразмерных систем).

Рассмотрим математическую модель кинетики процесса эволюции во времени системы таких последовательностей. Для этого обозначим посредством $N_i(t)$ число существующих в системе в момент времени t цепных молекул с последовательностью

$$i(z) = \{z_1 z_2, \dots, z_v\}; \quad z_k \in \{A, B, C, D\} \quad (3.125)$$

и их концентрацией:

$$x(t) = \frac{N_i(t)}{V}. \quad (3.126)$$

Полное число цепных молекул всех видов предполагается постоянным, так что в формальной математической записи условие отбора имеет вид

$$N = \sum_i N_i(t). \quad (3.127)$$

Условие постоянства общего числа цепных молекул приводит к конкуренции между различными их видами; исход этой борьбы определяется главным образом скоростью самовоспроизведения цепной последовательности атомов (молекул) E_i . В этой связи кинетика эволюции системы может быть представлена кинетическим уравнением вида

$$\Delta N_i(t) = (E_i - \langle E \rangle) \cdot N_i(t) \cdot \Delta t + q_i(t) \cdot \Delta t, \quad (3.128)$$

где $\Delta N_i(t)$ – изменение числа молекул вида i за интервал времени Δt , состоящее из вклада точного самовоспроизведения и вклада воспроизведения с ошибкой (мутацией) $q_i(t)\Delta t$.

Скорости воспроизведения E_i и мутационные функции q_i являются в рассматриваемой простой модели единственными величинами, которые предполагаются известными. Из этого вытекают все существенные для процесса эволюции свойства цепных молекул. Для известных цепных молекул истинный вид функций E_i и q_i , к сожалению, неизвестен. Однако с физической точки зрения представляются весьма правдоподобными два предположения:

- скорость воспроизведения E_i является возрастающей функцией сложности $K(i)$ цепной молекулы;
- вероятность a_{ij} перехода от последовательности j к последовательности i уменьшается с увеличением расстояния между последовательностями $d(ij)$.

Справедливость второго предположения очевидна, поскольку ошибочная копия должна всегда быть близкой к оригиналу: действительно, как правило, не на месте оказывается только один мономер, в худшем случае – несколько. Примером может служить мутация вида $\text{BDACAD} \rightarrow \text{BDABAD}$, которая имеет одно несовпадение с оригиналом и соответственно $d(ij) = 1$. Предположение об относительной малости изменений при прохождении процессов мутации обеспечивает известную

плавность процесса эволюции: вновь возникающие виды не слишком отличаются от прежних.

Значительно сложнее обосновать первое предположение, о связи между сложностью и скоростью воспроизведения. Следует исходить из предположения, что воспроизведение цепной молекулы из имеющихся в окружающей среде «кирпичиков» представляет сложный процесс, для осуществления которого необходима информация. Вообще говоря, для осуществления такого воспроизведения цепной молекуле необходимо, прежде всего, создать «воспроизводящий аппарат» (технологическое оборудование), принцип и способ действия которого должны быть заложены в последовательности. Подтверждением того, что такое предположение соответствует действительности, может служить анализ воспроизведения ДНК в клетке. В этом процессе в числе других участвует фермент, так называемая ДНК-полимераза. При синтезе белков содержащаяся в молекуле ДНК информация вначале «переписывается» на короткие, легко считываемые простые спирали молекулы РНК-носителя (аппарат транскрипции). За распознавание «кирпичиков» и установку их в правильном порядке целиком отвечает транспортная молекула РНК при участии активного фермента.

Таким образом, видно, что в самовоспроизведении ДНК принимает участие сложный аппарат, для создания которого молекула должна иметь его «чертежи» и «функциональные схемы», т.е. необходимое информационное обеспечение [32]. Отсюда можно сделать вывод, что скорость воспроизведения цепной молекулы должна увеличиваться с ростом ее сложности, т.е. с увеличением количества закодированной в ней информации. Такую же, или даже более важную роль, играет качество закодированной информации. Во всяком случае, некоторое количество закодированной информации является хотя и недостаточным, но необходимым условием для осуществления процесса самовоспроизведения молекулы.

На основе изложенного можно принять для достаточно простых математических моделей процесса эволюции (создания и совершенствования структур)

$$E_i = g [K(i)], \quad (3.129)$$

где $K(i)$ — сложность системы (по Колмогорову), а $g(K)$ — монотонно возрастающая функция действительного аргумента K .

Рассмотренные материалы не претендуют на исчерпывающее объяснение общих принципов самоорганизации, а уж тем более путей их использования в современной инженерной практике создания наноразмерных структур. Очевидно, что такая постановка вопроса является несколько преждевременной. Однако, если исходить из сложившихся реалий технологии атомной сборки изделий, то необходимо отметить следующее актуальное обстоятельство. Использование технологии атомной сборки изделий в том виде и по тем принципам, которые имеются в настоящее время, представляется крайне неэффективным с точки зрения ожидаемых технико-экономических показателей, когда речь идет о системах, содержащих в своем составе более $10^6 - 10^9$ атомов. Для таких систем «пошаговый» метод расстановки атомов приводит к неприемлемо низким показателям производственных процессов, связанных с выпуском больших партий изделий. Именно в этой связи использование принципов самоорганизации следует рассматривать в качестве возможности повышения экономической эффективности технологических процессов. Поэтому сближение точек зрения и методологических подходов, условно говоря, «биологической технологии» и «антропогенной технологии» может оказать плодотворное влияние на оба эти научно-технические направления.

ЗАКЛЮЧЕНИЕ

В настоящее время вполне очевидным фактом является тенденция разработчиков микросистем (и в первую очередь систем обработки информации) к освоению нанометровых размеров элементов конструкций устройств, что позволит существенно повысить тактико-технические данные современных вычислительных сред, а именно:

- степень интеграции элементов в единице объема физических сред, реализующих вычислительные процессы;
- быстрое действие элементарных вычислительных операций путем использования эффектов квантового параллелизма;
- расширение круга практических задач, решаемых методами вычислительного эксперимента.

Отмеченные выше моменты составляют предмет целенаправленного научно-технического поиска и, образно говоря, «лежат на поверхности» общей исторической тенденции развития вычислительной техники. С этой точки зрения процесс создания вычислительной техники можно рассматривать в качестве флагамена развития нанотехнологий и нанонауки как таковых. И тут уместно отметить важный привходящий фактор, имеющий большое значение в направлении развития ряда других научно-технических направлений, ранее не использовавших арсенал технологических возможностей микротехнологий.

Переход от микротехнологий к нанотехнологиям знаменует собой не просто банальное уменьшение размеров технических устройств. Этот переход создает известные предпосылки к разработке ряда специфических устройств, близких по своему функциональному назначению и конструктивному исполнению к природным (биологическим) объектам (в случае микротехнологий о такой близости не приходилось говорить ввиду существенного различия как технологических, так и функциональных принципов).

В этой связи можно отметить, что в настоящее время нанотехнологии одновременно становятся нужными весьма широкому кругу специалистов различной отраслевой направленности. С их помощью становится возможным «конструировать» на атомно-молекулярном уровне абсолютно новые (прорывные) изделия:

В медико-биологической отрасли:

- лекарственные средства, обладающие селективным механизмом взаимодействия с определенными органами человека и их транспортом, в соответствии со спецификой патологии;
- «биологические чипы», для целей мониторинга и диагностики состояния здоровья человека;
- «биологические нанороботы», для целей операбельного лечения болезней (выступающих в качестве альтернативы современной традиционной хирургии).

В химической и нефтехимической отрасли:

- устойчивые нанодисперсные среды (коллоиды, гели, золи) на органической и неорганической основе;
- молекулярные мембраны (атомно-молекулярное сито), для селективного разделения многокомпонентных сред;
- наноразмерные каталитические среды, для целей снижения энергоемкости и повышения эффективности современных химических производств и экологического назначения;
- искусственные атомно-молекулярные структуры, организованные по принципам супрамолекулярной химии.

В отрасли конструкционного материаловедения:

- нанокompозитные конструкционные материалы, позволяющие на порядок повысить статическую и динамическую прочность элементов конструкций;
- порошковая металлургия на основе нанодисперсных металлических порошков, позволяющая существенно снизить трудо- и энергоемкость технологических процессов;
- адаптивное управление (на основании управляемой анизотропии материалов) прочностными параметрами элементов конструкций.

В отрасли информационно-измерительной техники:

- квантовые вычислительные среды на базе «кубитов», основанные на принципах когерентной фазовой суперпозиции и перепутывания индивидуальных квантовых состояний;

- информационные сенсоры наноразмерного масштаба для весьма широкого круга технических параметров (производственно-технологической и эксплуатационной направленности);
- «технологические нанороботы», способные реализовать принципы технологии атомной сборки изделий (как альтернатива современным материало- и энергоемких технологических аппаратных средств).

Приведенный (далеко не полный) перечень направлений свидетельствует о чрезвычайно больших возможностях использования на практике нанотехнологий и наноструктурированных сред (изделий) даже на современном этапе развития наноинженерии. Более впечатляющих перспектив наноинженерии можно ожидать на этапе освоения технологии атомной сборки изделий.

Такая широта охвата объектов исследования предполагает использование адекватных по своей универсальности методов анализа нанотехнологий и наноразмерных структур. Подобная постановка вопроса тем более уместна, если учесть тот факт, что традиционные (макроскопические) методы анализа весьма ограничены в своих возможностях, когда речь идет о нанотехнологиях и наноразмерных структурах.

В этой связи рассмотренные выше подходы, основанные на энтропийных представлениях, могут оказаться достаточно эффективным инструментом при рассмотрении практических задач в области наноинженерии.

ЗАДАЧИ И УПРАЖНЕНИЯ

ЗАДАЧИ И УПРАЖНЕНИЯ К РАЗДЕЛУ №1

1.1 Вероятностные волны де Бройля

1) Вычислить длину волны де Бройля для газовых молекул в объеме, имеющем комнатную температуру.

№ варианта	1	2	3	4	5	6
Тип молекулы	H ₂	O ₂	N ₂	CO ₂	H ₂ O	OH

2) Определить энергию (в электронвольтах), которую необходимо дополнительно сообщить электрону, чтобы его длина волны де Бройля λ уменьшилась в n раз от начального значения λ_0 .

n	λ_0 , нм							
	1	10	10 ²	10 ³	10 ⁴	10 ⁵	10 ⁸	10 ¹⁰
10	№1	№2	№3	№4	№5	№6	№7	№8
100	№9	№10	№11	№12	№13	№14	№15	№16

Примечание: во внутренних позициях таблицы приведен порядковый № варианта.

3) Из катодной трубки на диафрагму с узкой прямоугольной щелью нормально к плоскости диафрагмы направлен поток моноэнергетических электронов. Определить анодное напряжение трубки, если известно, что на экране, отстоящем от щели на расстоянии L , ширина центрального дифракционного максимума $D = 100$ мкм. Ширину щели диафрагмы принять равной 0,01 мм.

№ варианта	1	2	3	4	5	6
L , м	0.5	0.1	0.2	0.5	1.0	2.0

4) Определить отношение длин волн де Бройля для электрона и протона, прошедших одинаковую ускоряющую разность потенциалов U .

№ варианта	1	2	3	4	5	6
U , кВ	10 ⁻²	10 ⁻¹	1.0	10	10 ²	10 ³

5) Кинетическая энергия микрочастицы равна удвоенному значению ее энергии покоя ($2m_0 \times c^2$). Вычислить длину волны де Бройля для случаев:

№ варианта	1	2	3	4	5	6
Тип микрочастицы	электрон	протон	H	O	OH	CO

1.2. Соотношение неопределенностей

6) Определить величину фазового объема, приходящегося на индивидуальную микрочастицу в импульсном подпространстве, при известных значениях неопределенностей координат расположения микрочастицы $\Delta x = \Delta y = \Delta z$.

№ варианта	1	2	3	4	5	6
ΔX , м	1.0	10^{-2}	10^{-4}	10^{-6}	10^{-8}	10^{-10}

7) Оценить минимальную кинетическую энергию электрона при его движении внутри одномерного потенциального ящика, имеющего протяженность L .

№ варианта	1	2	3	4	5	6
L , мкм	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1.0	10

8) Используя соотношение неопределенности, определить наименьшую погрешность в определении скорости электрона и протона, если координаты центра масс этих частиц могут быть установлены с неопределенностью ΔX .

№ варианта	1	2	3	4	5	6
ΔX , мкм	10^{-10}	10^{-6}	10^{-2}	1.0	10	100

9) Какова должна быть кинетическая энергия электрона в моноэнергетическом пучке, используемом для проведения микроскопических исследований структуры с линейными размерами L ?

№ варианта	1	2	3	4	5	6
L , мкм	10^{-11}	10^{-10}	10^{-9}	10^{-8}	10^{-4}	10^{-2}

10) Используя соотношение неопределенностей, оценить протяженность одномерного потенциального ящика, в котором задана минимальная энергия электрона E_{min} .

№ варианта	1	2	3	4	5	6
E_{min} , эВ	0.1	1.0	10	10^3	10^4	10^6

11) Среднее время жизни атома в возбужденном состоянии составляет $\Delta t \approx 10^{-8}$ с. При переходе атома в нормальное состояние испускается фотон, средняя длина волны λ которого равна n , нм. Оценить ширину $\Delta \lambda$ излучаемой веществом спектральной линии, если ее уширение не происходит за счет других процессов (воспользоваться дополнительным соотношением принципа неопределенности $\Delta \varepsilon \cdot \Delta t = \hbar$, где ε – энергия кванта света).

№ варианта	1	2	3	4	5	6
n	200	500	1000	2000	5000	10000

12) Для приближенной оценки минимальной энергии электрона в атоме водорода можно предположить, что неопределенность радиуса электронной орбиты и неопределенность импульса электрона на такой орбите соответственно связаны следующим образом: $\Delta r \approx r$ и $\Delta p \approx p$. Используя эти соотношения, найти значение радиуса электронной орбиты, соответствующее минимальной энергии электрона в атоме водорода (в приближении концепции свободных частиц).

13) Моноэнергетический пучок электронов в электронном литографе высвечивает на фотошаблоне пятно радиусом $r \approx 0.1$ мкм. Пользуясь соотношением неопределенностей, найти, во сколько раз неопределенность Δx координаты электрона в направлении, перпендикулярном направлению движения электронного луча, меньше размера r пятна. Длину электронного луча принять равной 0.6 м, а ускоряющее напряжение $U = n$ кВ.

№ варианта	1	2	3	4	5	6
n	0.1	1.0	10	20	50	100

1.3. Энергетическое состояние микрочастицы в потенциальном ящике

14) Частица находится в одномерном, прямоугольном потенциальном ящике с абсолютно непроницаемыми стенками. Найти величину разности ΔW между соседними энергетическими уровнями энергии $W(n_i)$ частицы для случаев: $n_1 = 1$; $n_2 = 2$; $n_3 = 3$; $n_5 = 5$; $n_\infty = \infty$.

15) Электрон находится в бесконечно глубоком, одномерном, прямоугольном потенциальном ящике шириной $L = n$, нм. Определить (в электронвольтах) наименьшую разность энергетических уровней электрона, находящегося в этом потенциальном ящике.

№ варианта	1	2	3	4	5	6
n	0.1	0.5	1.0	10	50	100

16) Частица находится в возбужденном состоянии ($n = 3$), в бесконечно глубоком, одномерном, прямоугольном потенциальном ящике шириной L . Определить, в каких точках интервала $0 < x < L$ плотность вероятности нахождения частицы принимает максимальное и минимальное значения.

17) В прямоугольной потенциальной яме шириной L с абсолютно непроницаемыми стенками находится частица в основном энергетическом состоянии. Найти вероятность местонахождения этой частицы в интервале длин $(L/4) < x < (3L/4)$.

18) Частица в бесконечно глубоком, одномерном, прямоугольном потенциальном ящике находится в основном состоянии. Какова вероятность обнаружения частицы в правой четверти ящика?

19) Амплитудная волновая функция, описывающая движение электрона в основном состоянии атома водорода, имеет вид $\phi(r) = A \cdot \exp(-r/a)$, где A — некоторая постоянная; a — первый боровский радиус орбиты. Найти для основного состояния атома водорода наиболее вероятное расстояние электрона от ядра атома.

20) Электрон находится в бесконечно глубоком, одномерном, прямоугольном потенциальном ящике шириной L . В каких точках в интервале значений $0 < x < L$ плотности вероятности нахождения электрона, расположенного на втором и третьем энергетических уровнях, будут одинаковы?

21) Волновая функция, описывающая движение электрона в основном состоянии атома водорода, имеет вид $\phi(r) = A \cdot \exp(-r/a)$, где A – некоторая постоянная; a – первый борковский радиус орбиты. Найти для основного состояния атома водорода среднее значение потенциальной энергии.

1.4. Зонные энергетические диаграммы твердого тела

22) Какова вероятность нахождения электрона на самом нижнем энергетическом уровне зоны проводимости полупроводника при температуре 300 К?

№ варианта	1	2	3	4
Тип полупроводника	Ge	Si	InSb	InAs
ΔW , ЭВ	0.66	1.12	0.18	0.35

23) Определить долю свободных электронов в металле (при температуре $T = 300$ К), энергия которых заключена в интервале значений $1/2 \cdot W_{\max} - W_{\max}$.

24) Германиевый кристалл, ширина запрещенной зоны в котором равна $\Delta W = 0,66$ эВ, нагревают от температуры $t_1 = 0^\circ\text{C}$ до температуры t_2 . Во сколько раз возрастет его удельная проводимость?

№ варианта	1	2	3	4	5	6
$t_2, ^\circ\text{C}$	10	20	50	100	150	300

25) При нагревании кремниевого кристалла от температуры $t_1 = 0^\circ$ до температуры $t_2 = 10^\circ\text{C}$ его удельная проводимость возрастает в 2,28 раза. По приведенным данным определить ширину ΔW запрещенной зоны в кремнии.

26) Найти минимальную энергию W_{\min} , необходимую для образования пары электрон–дырка в кристалле GaAs, если его удельная элек-

тропроводность σ изменяется в 10 раз при изменении температуры от 3 °С до 20°С.

27) Сопротивление кристалла PbS при температуре $T_1 = 20^\circ\text{C}$ равно $R_1 = 104 \text{ Ом}$. Определить его сопротивление R_2 при температуре $T_2 = 80^\circ\text{C}$.

28) Определить долю свободных электронов в металле при температуре $T = 10 \text{ К}$, энергия которых отличается от максимальной на $n \%$.

№ варианта	1	2	3	4	5	6
$n, \%$	1	5	10	20	30	50

29) Определить долю свободных электронов в металле при температуре $T = 100 \text{ К}$, значение энергии которых отличается от максимальной на $n \%$.

№ варианта	1	2	3	4	5	6
$n, \%$	1	5	10	20	30	50

30) Определить долю свободных электронов в металле при температуре $T = 10 \text{ К}$, энергия которых заключена в интервале значений от 0 до $1/n W_{\text{max}}$.

№ варианта	1	2	3	4	5	6
n	1	0.5	0.2	0.1	0.05	0.005

31) Определить длину волны излучаемого кванта света λ (в вакууме) для случая прохождения процессов рекомбинации в собственных полупроводниках.

№ варианта	1	2	3	4
Тип полупроводникового материала	Si	Ge	InSb	GaAs
Ширина запрещенной зоны (при $T = 300 \text{ К}$), эВ	1.21	0.72	0.17	1.42

ЗАДАЧИ И УПРАЖНЕНИЯ К РАЗДЕЛУ №2

2.1. Подвижность носителей заряда в кристаллических структурах

1) Подвижность носителя заряда при пролетном механизме его движения описывается выражением $\mu_{\text{прол}} = \frac{\lambda \cdot q}{2 \cdot \sqrt{3 \cdot k \cdot m \cdot T}}$. Определить значения подвижности $\mu_{\text{прол}}$ при различных длинах свободного пробега λ для случая ансамбля электронов, находящегося в кристаллической структуре при температуре 300 К.

№ варианта	1	2	3	4	5	6	7	8
λ , нм	1	2	5	10	20	50	100	200

2) Определить среднюю скорость движения носителя заряда (при механизме перемещения пролетного типа: $\lambda = 10$ нм; $T = 300$ К) при различных значениях напряженности электрического поля E .

№ варианта	1	2	3	4	5	6
E , В/м	1	10	10^2	10^4	10^6	10^8

3) Определить величину отношения векторизованной (при пролетном механизме движения) скорости v_E и броуновской v_T при различных значениях λ ($E = 10^3$ В/м; $T = 300$ К).

№ варианта	1	2	3	4	5	6	7	8
λ , нм	0.5	1	2	5	10	20	50	100

4) Подвижность носителя заряда в кристаллической структуре при прыжковом механизме его движения описывается выражением

$\mu_{\text{прыжок}} = \frac{v_0 \cdot q \cdot a^2}{k \cdot T} \cdot \exp\left(-\frac{W_a}{k \cdot T}\right)$. Определить значения подвижности $\mu_{\text{прыжок}}$ при различных значениях W_a для случая: $a = 10^{-10}$ м; $T = 300$ К, $v_0 = 10^{15}$ Гц.

№ варианта	1	2	3	4	5	6	7	8
W_a , эВ	0.5	1.0	1.25	1.5	2.0	2.5	3.0	5.0

5) Определить среднюю скорость движения носителя заряда (при механизме перемещения прыжкового типа: $a = 10^{-10}$ м; $T = 300$ К; $\nu_0 = 10^{15}$ Гц) при различных значениях напряженности электрического поля E .

№ варианта	1	2	3	4	5	6
E , В/м	1	10	10^2	10^4	10^6	10^8

6) Определить величину отношения векторизованной (при прыжковом механизме движения) скорости v_E и броуновской v_T при различных значениях W_a ($a = 10^{-10}$ м; $T = 300$ К; $\nu_0 = 10^{15}$ Гц).

№ варианта	1	2	3	4	5	6	7	8
W_a , эВ	0.1	0.2	0.5	1.0	1.5	2.0	2.5	3.0

7) Коэффициент прозрачности D для случая потенциального барьера конечной толщины определяется выражением $D = D_0 \cdot \exp\left(-\frac{4 \cdot \pi \cdot d}{h} \cdot \sqrt{2 \cdot m_e \cdot (U_0 - W)}\right)$, где D_0 – постоянный коэффициент, близкий по порядку величины единице. Построить зависимость $D = \varphi(d)$ при известных значениях параметра $(U_0 - W)$.

№ варианта	1	2	3	4	5	6	7	8
$(U_0 - W)$, эВ	0.1	0.2	0.3	0.4	0.5	1.0	2.0	3.0

8) Определить скорость движения носителя заряда (электронов) в различных полупроводниковых материалах при различных значениях напряженности электрического поля: E (В/м) = а) 10^2 ; б) 10^3 ; в) 10^4 ; г) 10^6 .

№ варианта	1	2	3	4	5
Тип полупроводника	Si	Ge	InSb	GaAs	InSb
Подвижность, $\text{м}^2/\text{В} \cdot \text{с}$	0.135	0.382	7.74	8.35	26.3

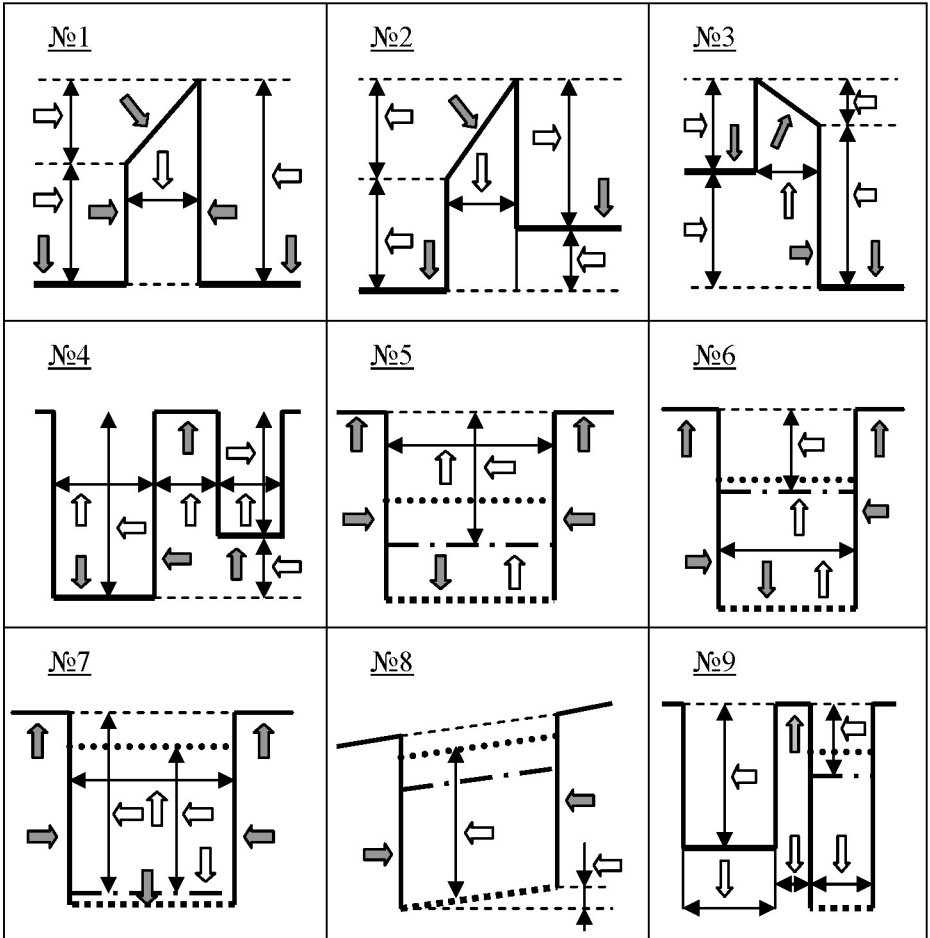
9) Определить время прохождения носителями заряда (электроны) слоя полупроводникового материала толщиной d при значениях напряженности электрического поля: E (В/м) = а) 10^2 ; б) 10^3 ; в) 10^4 ; г) 10^6 .

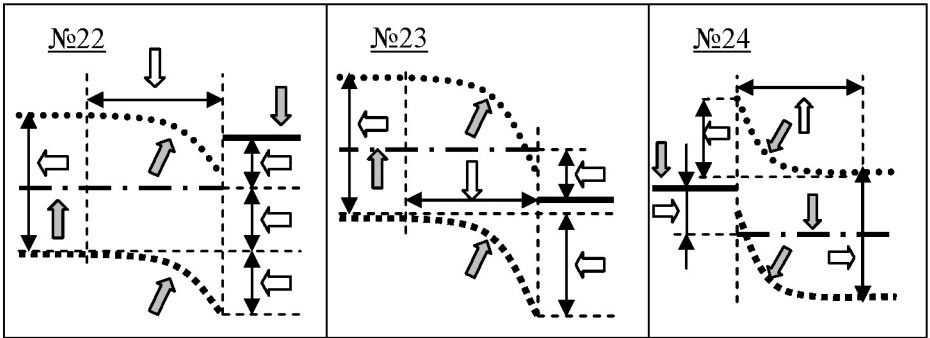
№ варианта	1	2	3	4	5	6	7	8	9	10
d , мкм	0.01	0.05	0.1	0.5	1.0	2.0	5.0	10.0	20.0	50.0

Примечание: при решении воспользоваться данными и результатами задачи № 8.

2.2. Контактные явления в кристаллических структурах

10) Провести анализ представленных ниже зонных энергетических диаграмм контактных твердотельных систем, выполненных на основе металлов и полупроводников (методология анализа представлена в приложении №3).





11) Построить зонные энергетические диаграммы для контактных твердотельных систем, выполненных на основе металлов и полупроводников (методология построения диаграмм представлена в приложении №3).

Вариант № 1 Система «M ₁ – M ₂ »		Вариант № 2 Система «M ₁ – M ₂ »		
<p>A)</p> $W_{\text{ВЫХ}1} = 4.5 \text{ эВ}$ $W_{\text{ВЫХ}2} = 1.0 \text{ эВ}$ $d = 10 \text{ А}$	<p>Б)</p> $W_{\text{ВЫХ}1} = 2.5 \text{ эВ}$ $W_{\text{ВЫХ}2} = 4.0 \text{ эВ}$ $d = 20 \text{ А}$	<p>A)</p> $W_{\text{ВЫХ}1} = 2.5 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3.0 \text{ эВ}$ $U_1 = 0 \text{ В}$ $U_2 = + 3 \text{ В}$	<p>Б)</p> $W_{\text{ВЫХ}1} = 2.5 \text{ эВ}$ $W_{\text{ВЫХ}2} = 4.0 \text{ эВ}$ $U_1 = 0 \text{ В}$ $U_2 = +2 \text{ В}$	
<p>В)</p> $W_{\text{ВЫХ}1} = 4.0 \text{ эВ}$ $W_{\text{ВЫХ}2} = 4.0 \text{ эВ}$ $d = 30 \text{ А}$	<p>Г)</p> $W_{\text{ВЫХ}1} = 0.5 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3.0 \text{ эВ}$ $d = 40 \text{ А}$	<p>В)</p> $W_{\text{ВЫХ}1} = 4.0 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3.0 \text{ эВ}$ $U_1 = 0 \text{ В}$ $U_2 = - 4 \text{ В}$	<p>Г)</p> $W_{\text{ВЫХ}1} = 0.5 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3.0 \text{ эВ}$ $U_1 = 0 \text{ В}$ $U_2 = - 5 \text{ В}$	
Вариант № 3 Система «M ₁ – M ₂ – M ₃ »				
<p>A)</p> $W_{\text{ВЫХ}1} = 4 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$	<p>Б)</p> $W_{\text{ВЫХ}1} = 1 \text{ эВ}$ $W_{\text{ВЫХ}2} = 4 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$	<p>В)</p> $W_{\text{ВЫХ}1} = 4 \text{ эВ}$ $W_{\text{ВЫХ}2} = 1 \text{ эВ}$ $W_{\text{ВЫХ}3} = 4 \text{ эВ}$	<p>Г)</p> $W_{\text{ВЫХ}1} = 4 \text{ эВ}$ $W_{\text{ВЫХ}2} = 4 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$	<p>Д)</p> $W_{\text{ВЫХ}1} = 6 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3 \text{ эВ}$ $W_{\text{ВЫХ}3} = 0 \text{ эВ}$
<p>Е)</p> $W_{\text{ВЫХ}1} = 4 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3 \text{ эВ}$ $W_{\text{ВЫХ}3} = 4 \text{ эВ}$	<p>Ж)</p> $W_{\text{ВЫХ}1} = 4 \text{ эВ}$ $W_{\text{ВЫХ}2} = 0 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$	<p>З)</p> $W_{\text{ВЫХ}1} = 0 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$	<p>И)</p> $W_{\text{ВЫХ}1} = 2 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$	<p>К)</p> $W_{\text{ВЫХ}1} = 1 \text{ эВ}$ $W_{\text{ВЫХ}2} = 3 \text{ эВ}$ $W_{\text{ВЫХ}3} = 1 \text{ эВ}$

Вариант № 4 Система «M ₁ – M ₂ – M ₃ »			
<p>А)</p> $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{ВЫХ } 2} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 3} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$		<p>Б)</p> $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{ВЫХ } 2} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 3} = 1 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$	
<p>В)</p> $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{ВЫХ } 2} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 3} = 3 \text{ эВ}$ $U_1 = 5 \text{ В}; U_2 = 4 \text{ В}; U_3 = -3 \text{ В}$		<p>Г)</p> $W_{\text{ВЫХ } 1} = 1 \text{ эВ}; W_{\text{ВЫХ } 2} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 3} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = 2 \text{ В}$	
<p>Д)</p> $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{ВЫХ } 2} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 3} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 6 \text{ В}; U_3 = -1 \text{ В}$		<p>Е)</p> $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{ВЫХ } 2} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 3} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 2 \text{ В}; U_3 = -1 \text{ В}$	
Вариант № 5 Система «M–П/П ₁ »			
<p>А)</p> $W_M = 4.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 3 \text{ эВ}$	<p>Б)</p> $W_M = 1.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 2 \text{ эВ}$	<p>В)</p> $W_M = 1.5 \text{ эВ}$ $W_{\text{П/П}} = 4.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 1 \text{ эВ}$	<p>Г)</p> $W_M = 0.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 0.5 \text{ эВ}$
<p>Д)</p> $W_M = 2.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 0 \text{ эВ}$	<p>Е)</p> $W_M = 1 \text{ эВ}$ $W_{\text{П/П}} = 2 \text{ эВ}$ $\Delta W_{\text{П/П}} = 1 \text{ эВ}$	<p>Ж)</p> $W_M = 2.5 \text{ эВ}$ $W_{\text{П/П}} = 3.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 2 \text{ эВ}$	<p>З)</p> $W_M = 2 \text{ эВ}$ $W_{\text{П/П}} = 2 \text{ эВ}$ $\Delta W_{\text{П/П}} = 3 \text{ эВ}$
Вариант № 6 Система «M–П/П _д »			
<p>А)</p> $W_M = 4.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 3 \text{ эВ}$	<p>Б)</p> $W_M = 1.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 2 \text{ эВ}$	<p>В)</p> $W_M = 1.5 \text{ эВ}$ $W_{\text{П/П}} = 4.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 1 \text{ эВ}$	<p>Г)</p> $W_M = 0.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 0.5 \text{ эВ}$
<p>Д)</p> $W_M = 2.5 \text{ эВ}$ $W_{\text{П/П}} = 1.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 0 \text{ эВ}$	<p>АЕ)</p> $W_M = 1 \text{ эВ}$ $W_{\text{П/П}} = 2 \text{ эВ}$ $\Delta W_{\text{П/П}} = 1 \text{ эВ}$	<p>Ж)</p> $W_M = 2.5 \text{ эВ}$ $W_{\text{П/П}} = 3.0 \text{ эВ}$ $\Delta W_{\text{П/П}} = 2 \text{ эВ}$	<p>З)</p> $W_M = 2 \text{ эВ}$ $W_{\text{П/П}} = 2 \text{ эВ}$ $\Delta W_{\text{П/П}} = 3 \text{ эВ}$

Вариант № 7 Система «М – П / П _а »				
А) $W_M = 4.5 \text{ эВ}$ $W_{\Pi/\Pi} = 1.0 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 3 \text{ эВ}$	Б) $W_M = 1.5 \text{ эВ}$ $W_{\Pi/\Pi} = 1.0 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 2 \text{ эВ}$	В) $W_M = 1.5 \text{ эВ}$ $W_{\Pi/\Pi} = 4.0 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 1 \text{ эВ}$	Г) $W_M = 0.5 \text{ эВ}$ $W_{\Pi/\Pi} = 1.0 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 0.5 \text{ эВ}$	
Д) $W_M = 2.5 \text{ эВ}$ $W_{\Pi/\Pi} = 1.0 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 0 \text{ эВ}$	АЕ) $W_M = 1 \text{ эВ}$ $W_{\Pi/\Pi} = 2 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 1 \text{ эВ}$	Ж) $W_M = 2.5 \text{ эВ}$ $W_{\Pi/\Pi} = 3.0 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 2 \text{ эВ}$	З) $W_M = 2 \text{ эВ}$ $W_{\Pi/\Pi} = 2 \text{ эВ}$ $\Delta W_{\Pi/\Pi} = 3 \text{ эВ}$	
Вариант № 8 Системы «М ₁ –П/П ₁ –М ₂ »				
А) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\Pi/\Pi} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	Б) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\Pi/\Pi} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	В) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 4 \text{ эВ}$	Г) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	Д) $W_{\text{ВЫХ } 1} = 6 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 0 \text{ эВ}$
Е) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 4 \text{ эВ}$	Ж) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 0 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$	З) $W_{\text{ВЫХ } 1} = 0 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$	И) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$	К) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$
Вариант № 9 Система «М ₁ –П/П _а –М ₂ »				
А) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}$ $W_{\Pi/\Pi} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	Б) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\Pi/\Pi} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$	В) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\Pi/\Pi} = 2 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 4 \text{ эВ}$	Г) $W_{\text{ВЫХ } 1} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 4 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	Д) $W_{\text{ВЫХ } 1} = 8 \text{ эВ}$ $W_{\Pi/\Pi} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 0 \text{ эВ}$
Е) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 4 \text{ эВ}$	Ж) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 0 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$	З) $W_{\text{ВЫХ } 1} = 0 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$	И) $W_{\text{ВЫХ } 1} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$	К) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 1 \text{ эВ}$
Вариант № 10 Система «М ₁ –П/П _д –М ₂ »				
А) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\Pi/\Pi} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	Б) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\Pi/\Pi} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 2 \text{ эВ}$	В) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}$ $W_{\Pi/\Pi} = 2 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 4 \text{ эВ}$	Г) $W_{\text{ВЫХ } 1} = 3 \text{ эВ}$ $W_{\Pi/\Pi} = 5 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$	Д) $W_{\text{ВЫХ } 1} = 8 \text{ эВ}$ $W_{\Pi/\Pi} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 0 \text{ эВ}$

Е) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\text{П/П}} = 4 \text{ эВ}$	Ж) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 0 \text{ эВ}$ $W_{\text{П/П}} = 1 \text{ эВ}$	З) $W_{\text{ВЫХ } 1} = 0 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\text{П/П}} = 1 \text{ эВ}$	И) $W_{\text{ВЫХ } 1} = 3 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 2 \text{ эВ}$ $W_{\text{П/П}} = 1 \text{ эВ}$	К) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $W_{\text{П/П}} = 1 \text{ эВ}$
---	---	---	---	---

Вариант № 11 Система «М₁-П/П₁-М₂»

А) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$	Б) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{П/П}} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$
В) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{П/П}} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 5 \text{ В}; U_2 = 4 \text{ В}; U_3 = -3 \text{ В}$	Г) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}; W_{\text{П/П}} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = 2 \text{ В}$
Д) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 6 \text{ В}; U_3 = -1 \text{ В}$	Е) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 2 \text{ В}; U_3 = -2 \text{ В}$

Вариант № 12 Система «М₁-П/П₁-М₂»

А) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{П/П}} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 2 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$	Б) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$ $U_1 = 3 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$
В) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 5 \text{ В}; U_2 = 3 \text{ В}; U_3 = -3 \text{ В}$	Г) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}; W_{\text{П/П}} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$
Д) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}; W_{\text{П/П}} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 6 \text{ В}; U_3 = -2 \text{ В}$	Е) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 2 \text{ В}; U_3 = -3 \text{ В}$

Вариант № 13 Система «М ₁ -П/П _а -М ₂ »	
А) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$	Б) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{П/П}} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 1 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = -2 \text{ В}$
В) $W_{\text{ВЫХ } 1} = 2 \text{ эВ}; W_{\text{П/П}} = 1 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 5 \text{ В}; U_2 = 4 \text{ В}; U_3 = -3 \text{ В}$	Г) $W_{\text{ВЫХ } 1} = 1 \text{ эВ}; W_{\text{П/П}} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 0 \text{ В}; U_3 = 2 \text{ В}$
Д) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 2 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 6 \text{ В}; U_3 = 1 \text{ В}$	Е) $W_{\text{ВЫХ } 1} = 4 \text{ эВ}; W_{\text{П/П}} = 3 \text{ эВ};$ $W_{\text{ВЫХ } 2} = 3 \text{ эВ}$ $U_1 = 1 \text{ В}; U_2 = 2 \text{ В}; U_3 = -1 \text{ В}$

2.3. Транспорт носителей заряда в наноразмерных структурах

12) Определить скорость векторизованного движения микрочастиц, находящихся во внешнем электрическом поле при значениях параметров: $q = 10^{-19}$ Кл; $m = 10^{-20}$ кг; $\lambda = 10^{-8}$ м (в рамках классического приближения воспользоваться выражением (2.112)).

№ варианта	1	2	3	4	5	6	7	8
$E, \text{ В/м}$	10	10^2	10^3	10^4	10^5	10^6	10^7	10^8

13) Определить значения максимальной скорости векторизованного движения носителя заряда в направлении пространственного ограничения наноразмерной структуры с протяженностью L при значениях параметров: $m = 10^{-20}$ кг; $a = 10^{-10}$ м; $\Delta W_1 = 10^{-23}$ эВ (воспользоваться выражением (2.114)).

№ варианта	1	2	3	4	5	6	7	8
$L, \text{ нм}$	0.5	1.0	5	10	20	50	100	500

14) Определить значения дебаевской длины экранирования L_D для кристаллических структур полупроводникового типа (воспользоваться выражением (2.102)).

№ варианта	1	2	3	4	5
ε	2	4	6	8	10

15) Оценить масштабы значений вероятностей реализации энергетических переходов в наноразмерных структурах (в рамках представлений об активационном характере процессов), стимулированных влиянием внешнего электрического поля (воспользоваться выражением (2.85)), при $q = 10^{-19}$ Кл; $T = 300$ К).

E , В/м	λ , нм							
	1	10	10^2	10^3	10^4	10^5	10^8	10^{10}
10^2	№1	№2	№3	№4	№5	№6	№7	№8
10^4	№9	№10	№11	№12	№13	№14	№15	№16
10^6	№17	№18	№19	№20	№21	№22	№23	№24
10^8	№25	№26	№27	№28	№29	№30	№31	№32
10^{10}	№33	№34	№35	№36	№37	№38	№39	№40

16) В направлении размерного ограничения наноструктура кристаллического типа представляет собой атомную цепочку в количестве N атомов. Оценить масштаб флуктуационных вариаций термодинамических параметров наноструктуры при различных значениях длины атомной цепочки наноструктуры N (воспользоваться выражением (2.79)).

№ варианта	1	2	3	4	5	6	7	8
N	5	10	20	50	100	200	500	1000

17) Определить значения частной статистической энтропии размещения атомов в пределах регулярной наноструктуры при различных значениях количества атомных ячеек N и размещаемых атомов n (воспользоваться выражением (2.80)).

№ варианта	1	2	3	4	5	6	7	8
N	5	10	15	20	25	30	35	40
n	3	7	10	12	6	9	14	8

18) Определить вероятность реализации макросостояния системы i^{20} типа (т.е. размещения n микрочастиц в объеме A и размещения $(N-n)$ микрочастиц – в объеме B) (воспользоваться выражением (2.81)).

№ варианта	1	2	3	4	5	6	7	8
N	5	10	15	20	25	30	35	40
n	3	7	10	12	6	9	14	8

19) Определить значения поверхностного потенциала $\Delta\varphi_s$ в зависимости от поверхностной концентрации адсорбированных атомов $N^{(+)}$ акцепторного типа и протяженности участка изгиба энергетической разрешенной зоны x_s у поверхности раздела сред при $\varepsilon = 10$; $q = 10^{-19}$ Кл (воспользоваться выражением (2.110)).

№ варианта	1	2	3	4	5	6	7	8
$N^{(+)}$, см ⁻²	10^4	10^5	10^6	10^7	10^8	10^9	10^{10}	10^{11}
x_s , нм	0.5	1.0	5.0	10	20	50	100	500

20) Определить значения критической напряженности электрического поля $E_{кр}$ для модели подвижности носителей заряда пролетного типа при условии справедливости допущения вида $v_T \gg v_E$ (воспользоваться выражением (2.11)).

№ варианта	1	2	3	4	5	6	7	8
T , °С	100	200	300	400	500	600	800	1000
λ , нм	5.0	10	50	100	5.0	10	50	100

ЗАДАЧИ И УПРАЖНЕНИЯ К РАЗДЕЛУ №3

1) Определить значения случайной и средней энтропии.

Варианты:

- а) для случая однократного выброса игральной кости;
- б) для случая однократного выбора определенной карты;
- в) для случая n -кратного выброса игральной кости;
- г) для случая n -кратного выбора определенной карты (с вложением предыдущей карты).

2) Вероятность события, состоящего в том, что студент сдаст экзамен, равна P . Определить значение случайной энтропии для случаев сдачи и несдачи экзамена.

№ варианта	$P, \%$	№ варианта	$P, \%$
1	0.999	5	0.500
2	0.990	6	0.100
3	0.900	7	0.050
4	0.800	8	0.010

3) Вероятность события, состоящего в том, что студент сдаст экзамен, равна P . Определить значение средней энтропии процедуры сдачи экзамена.

№ варианта	$P, \%$	№ варианта	$P, \%$
1	99.99	5	75.00
2	99.90	6	50.00
3	99.00	7	10.00
4	85.00	8	1.000

4) Суть технологии атомной сборки состоит в том, что на подложке избирательно размещаются атомы определенного типа. Вероятность размещения атома необходимого типа равна P . Определить вероятность и энтропию создания линейной атомной цепочки определенной конфигурации длиной N , содержащей n атомов необходимого типа.

№ варианта	N	1 – P
1	5	10^{-6}
2	10	10^{-7}
3	15	10^{-8}
4	20	10^{-9}
5	25	10^{-10}

5) Построить график зависимости относительной концентрации (N_a / N_0) активных частиц (атомов) от температуры при известном значении энергии активации элементарного процесса W_a .

№ варианта	W_a (эВ)	Температура вещества (К)			
1	0.25	100	150	200	300
2	0.50	100	150	200	300
3	0.75	200	300	400	500
4	1.00	200	300	400	500
5	1.20	200	300	400	500

6) Во сколько раз необходимо повысить температуру вещества (начальные температуры T_0), чтобы концентрация активных частиц возросла в десять раз? (При известном значении энергии активации элементарного процесса W_a .)

№ варианта	W_a (эВ)	Температура вещества T_0 (К)		
1	0.25	100	150	200
2	0.50	100	150	200
3	0.75	200	300	400
4	1.00	200	300	400
5	1.20	200	300	400

7) Определить концентрацию носителей зарядов в собственном полупроводнике при известном значении ширины запрещенной зоны W_n .

№ варианта	W_n (эВ)	Температура полупроводника (К)			
		100	150	200	250
1	1.0	100	150	200	250
2	1.1	100	150	200	250
3	1.2	100	150	200	250
4	1.3	150	200	250	300
5	1.4	150	200	250	300

8) Определить относительную концентрацию ионов в газовой фазе при известном значении энергии ионизации (W_u) и температуре газа.

№ варианта	Вещество	W_u (эВ)	T (К)
1	H	13.6	800
2	He	24.6	1500
3	N	14.5	900
4	O	13.6	700
5	H ₂	15.4	1000

9) Определить макроскопическую скорость движения носителей зарядов (ионов) v_E , принимая во внимание прыжковый (термоактивационный) механизм подвижности (μ) носителей зарядов в конденсированных средах.

(Примечание: $\mu = \mu_0 \cdot \exp(-W_n/k \cdot T)$, где $\mu_0 = 10^{-8}$ (м²/с/В).

№ варианта	W_n (эВ)	E (В/см)	T (К)
1	0.25	500	150
2	0.50	400	175
3	0.75	300	200
4	1.00	200	250
5	1.20	100	300

10) Отношение $N_{a1}/N_{a2} = 2$, где индексы 1 и 2 характеризуют различные температуры вещества T_1 и T_2 . Определить температуру T_2 , если известны T_1, W_{a1}, W_{a2} .

№ варианта	W_{a1} (эВ)	W_{a2} (эВ)	T_1 (К)
1	1.20	1.30	200
2	1.50	1.75	200

3	1.75	2.00	400
4	2.00	2.50	350
5	2.50	2.75	300

11) Разность температур $\Delta T = T_2 - T_1$ обеспечивает выполнение условия $N_{a1}/N_{a2} = 0.1$. Определить значение энергии активации процесса (W_a).

№ варианта	ΔT (К)	T_1 (К)
1	100	150
2	50	175
3	50	200
4	70	250
5	70	300

12) Определить энергию активации подвижности ионов в разбавленных растворах по известному значению подвижности ионов при температуре 18° С .

№ варианта	Тип иона		Подвижность * 10 ⁷ (м ² /с)/В
	Катион	Анион	
1	К		67
2	Na		44
3	Li		33
4	H		320
5	Ag		54
6	Zn		46
7		J	67
8		Br	68
9		Cl	66
10		OH	174
11		NO ₂	62
12		VnO ₄	53

13) Оценить равновесную концентрацию точечных дефектов (вакансий) в кристаллической решетке при условиях, что заданы: энергия активации образования точечного дефекта W_{ad} ; температура кристалла T ; $N_0 = n$ (моль/л) .

№ варианта	W_{ad} (эВ)	n	T (К)
1	0.50	11	900
2	0.75	10	800
3	0.75	9	700
4	0.75	8	600
5	0.80	7	500

14) Определить скорость процесса десорбции примесных атомов с поверхности твердого тела, если известны значения энергии активации процесса десорбции W_d и поверхностная концентрация примесных атомов N_0 .

№ варианта	T (К)	W_d (эВ)	N_0 (атом/см ²)
1	50	0.5 (1.0)	10^{12}
2	100		
3	200		
4	400		

15) Определить значения и построить график зависимости константы равновесия химического процесса от температуры.

№ варианта	W_{aO} (эВ)	$W_{aП}$ (эВ)	T (К)			
1	0.6	0.5	50	100	300	600
2	0.8					
3	1.0					
4	1.2					

16) Построить график зависимости скорости химического процесса от температуры в начальный момент времени (C_0 – начальная концентрация реагента).

№ варианта	1	2	3	4	5	6	7	8
W_a (эВ)	0.5	0.7	0.9	1.2	1.5	1.8	2.1	2.4
C_0 (моль/л)	0.01	0.1	1.0	1.5	2.0	3.0	4.0	5.0

17) Построить график зависимости скорости химического процесса от времени ($C_o = 1$ моль/л – начальная концентрация реагента).

№ варианта	1	2	3	4	5	6	7
W_a (эВ)	0.1	0.2	0.3	0.4	0.5	0.6	0.7
T (K)	50	100	150	200	250	300	350

18) Определить срок службы материалов τ при температуре эксплуатации T , если известен срок службы материала τ_0 при температуре эксплуатации T_0 .

№ варианта	τ_0 (с)	T_0 (K)	W_C (эВ)	T (K)
1	10^6	300	1.0	600
2			1.1	650
3			1.2	700
4			1.3	750

19) Во сколько раз увеличится срок службы изделия, если энергия активации процесса старения изделия W_C возрастет на 10 % (при температуре эксплуатации изделия T_0).

№ варианта	W_C (эВ)	τ_0 (с)	T_0 (K)
1	1.0	10^6	600
2	1.1		650
3	1.2		700
4	1.3		750

20) Определить температуру (T) режима ускоренных испытаний изделия, обеспечивающую сокращение продолжительности штатного режима испытаний (τ_0) в n раз, если известны: T_0 – температура штатного режима испытаний; W_C – энергия активации процесса старения изделия.

№ варианта	T_0 (K)	W_C (эВ)	τ_0 (с)	n
1	300	1.0	10^4	10
2	310			
3	320	1.2	$5 \cdot 10^4$	8
4	330			

21) Построить график зависимости количества возможных реализаций N_c от количества ячеек размещения N в наноструктуре при условии равной вероятности размещения каждого из m типов различных атомов, содержащихся в рабочей среде.

22) Определить количество реализаций $\lg N_c$ в зависимости от размеров рабочего поля наноструктуры (n_x, n_y) и количества типов размещаемых атомов m .

№ варианта	Количество ячеек рабочего поля		Количество типов размещаемых атомов m
	n_x	n_y	
1	5	5	10
2	10	10	10
3	20	20	10
4	50	50	10
5	5	5	50
6	10	10	50
7	20	20	50
8	50	50	50
9	5	5	100
10	10	10	100

23) Количество размещаемых типов атомов $m = 2$ (атомы типа А и В). Вероятности размещения атомов: $P_A = 0.9$; $P_B = 0.1$. Количество ячеек размещения атомов в рабочем поле наноструктуры (отображаем в виде линейной цепочки) при $N = 5$.

Определить вероятность получения наноструктуры типа:

№ варианта	а	б	в	г
1	BBBBB	AAAAA	ABBBB	BAAAA
2	ABBBA	BAAAB	AABAA	BBABBB
3	ABABA	BAVAB	AAABV	VBVAAA

24) Определить значения информационной энтропии размещения единичного атома при наличии технологии h_i при заданных значениях m и p_i .

№ варианта	Вероятность размещения в индивидуальной ячейке атома i -го типа p_i										
	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}
1	0.5	0.5									
2	0.2	0.2	0.2	0.2	0.2						
3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	
4	0.9	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
5	0.99	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
6	0.999	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}	10^{-4}
7	$0,9^{(4)}$	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-5}
8	$0,9^{(5)}$	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}
9	$0,9^{(6)}$	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}
10	0.9	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.091
11	0.999	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	10^{-6}	$9,91 \cdot 10^{-4}$
12	$0,9^{(4)}$	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	10^{-7}	$9,91 \cdot 10^{-5}$
13	$0,9^{(5)}$	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	$9,91 \cdot 10^{-6}$
14	$0,9^{(5)}$	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	$9,95 \cdot 10^{-6}$
15	$0,9^{(5)}$	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	10^{-8}	$4,97 \cdot 10^{-6}$
16								$4,97 \cdot 10^{-6}$	$4,97 \cdot 10^{-6}$		
	0.0909090 – для всех типов атомов («случайная» технология)										

Примечание: запись вида $0,9^{(4)}$ означает $0,9999$.

25) Определить значения энтропийных показателей (K , α , h_t) для случаев (вариантов), рассмотренных в задаче 24.

26) Определить значения информационной энтропии нано- и микроструктур H_t при различных значениях параметров N и h_t (в позиции « h_t » указан № варианта из задания № 2.1).

№ варианта	N	h_t	№ варианта	N	h_t
1	10	№6	11	100	№4
2	20	№6	12	100	№5
3	50	№6	13	100	№7
4	10^2	№6	14	100	№8
5	10^3	№6	15	100	№9
6	10^4	№6	16	10^8	№10
7	10^5	№6	17	10^9	№11
8	10^6	№6	18	10^{10}	№12
9	10^7	№6	19	10^{11}	№13
10	10^8	№6	20	10^{12}	№15

27) Определить меру сложности устройства C в зависимости от параметров N , m и величины отношения H_c/H_i .

№ варианта	N	m	H_c/H_i	№ варианта	N	m	H_c/H_i
1	10^{10}	100	10^7	11	5	10	1
2	10^8	100	10^7	12	10	10	5
3	10^6	100	10^7	13	20	10	10
4	10^4	100	10^7	14	50	10	10^2
5	10^3	100	10^7	15	10^2	10	10^3
6	10^2	100	10^7	16	$2 \cdot 10^2$	50	10
7	50	100	10^7	17	$5 \cdot 10^2$	50	10^3
8	10	100	10^7	18	10^3	50	10^6
9	5	100	10^7	19	10^4	50	10^9
10	1	100	10^7	20	10^6	50	10^{12}

28) Определить энтропию единичного размещения атомов h_t в нано- и микроструктурах, обеспечивающую необходимый процент выхода годных изделий P при известных значениях параметров α_i и m .

№ варианта	$P, \%$	m	α_i	№ варианта	$P, \%$	m	α_i
1	99	10	10^4	15	99	10	10^{10}
2	90	10	10^4	16	90	10	10^{10}
3	60	10	10^4	17	60	10	10^{10}
4	40	10	10^4	18	40	10	10^{10}
5	10	10	10^4	19	10	10	10^{10}
6	1.0	10	10^4	20	1.0	10	10^{10}
7	0.1	10	10^4	21	0.1	10	10^{10}
8	99	10	10^8	22	99	1	10^4
9	90	10	10^8	23	10	1	10^4
10	60	10	10^8	24	99	1	10^8
11	40	10	10^8	25	10	1	10^8
12	10	10	10^8	26	99	1	10^{10}
13	1.0	10	10^8	27	50	1	10^{10}
14	0.1	10	10^8	28	10	1	10^{10}

29) Определить процент выхода годных изделий P , если известно, что $h_c = 2.004$ и h_t, h_i (см. данные таблицы).

№ варианта	h_t	h_i	N	№ варианта	h_t	h_i	N
1	10^{-9}	10^{-9}	10^2	15	10^{-9}	10^{-9}	10^4
2	10^{-8}	10^{-9}	10^2	16	10^{-8}	10^{-9}	10^4
3	10^{-7}	10^{-9}	10^2	17	10^{-7}	10^{-9}	10^4
4	10^{-6}	10^{-9}	10^2	18	10^{-3}	10^{-9}	10^4
5	10^{-5}	10^{-9}	10^2	19	10	10^{-9}	10^4
6	10^{-4}	10^{-9}	10^2	20	10^{-9}	10^{-9}	10^8
7	10^{-3}	10^{-9}	10^2	21	10^{-8}	10^{-9}	10^8
8	10^{-2}	10^{-9}	10^2	22	10^{-7}	10^{-9}	10^8
9	2.004	10^{-9}	10^2	23	10^{-3}	10^{-9}	10^8

10	10^{-6}	10^{-7}	10^3	24	10	10^{-9}	10^8
11	10^{-4}	10^{-7}	10^3	25	10^{-9}	10^{-9}	10^{12}
12	10^{-3}	10^{-7}	10^3	26	10^{-8}	10^{-9}	10^{12}
13	10^{-2}	10^{-7}	10^3	27	10^{-7}	10^{-9}	10^{12}
14	10	10^{-7}	10^3	28	10^{-3}	10^{-9}	10^{12}

30) Определить информационный дефицит технологии создания микро- и наноструктур ΔI по значениям, заданным в упражнении 29.

31) Определить минимальное количество метрологических операций n , устраняющих информационный дефицит технологии, по значениям, заданным в упражнении 30.

32) Определить выход годных изделий P в зависимости от информационного дефицита ΔI , объема микро- и наноструктуры V и энтропии единичного размещения атомов в элементарной ячейке h_t .

№ варианта	h_t	ΔI	V , мкм ³	№ варианта	h_t	ΔI	V , мкм ³
1	10^{-9}	10^{-7}	10^{-9}	15	10^{-9}	0.4	10^{-3}
2	10^{-8}	$0.2 \cdot 10^5$	10^{-9}	16	10^{-8}	$0.3 \cdot 10$	10^{-3}
3	10^{-7}	$0.3 \cdot 10^4$	10^{-9}	17	10^{-7}	$0.2 \cdot 10^2$	10^{-3}
4	10^{-6}	$0.4 \cdot 10^3$	10^{-9}	18	10^{-3}	$0.3 \cdot 10^6$	10^{-3}
5	10^{-5}	$0.5 \cdot 10^2$	10^{-9}	19	10	$0.4 \cdot 10^{10}$	10^{-3}
6	10^{-4}	$0.6 \cdot 10^1$	10^{-9}	20	10^{-9}	$0.5 \cdot 10^3$	1.0
7	10^{-3}	$0.7 \cdot 1.0$	10^{-9}	21	10^{-8}	$0.6 \cdot 10^4$	1.0
8	10^{-2}	$0.8 \cdot 10$	10^{-9}	22	10^{-7}	$0.7 \cdot 10^5$	1.0
9	2.004	$2 \cdot 10^2$	10^{-9}	23	10^{-3}	$0.8 \cdot 10^9$	1.0
10	10^{-6}	0.6	10^{-6}	24	10	$0.9 \cdot 10^{13}$	1.0
11	10^{-4}	$0.9 \cdot 10^2$	10^{-6}	25	10^{-9}	$0.8 \cdot 10^6$	10^3
12	10^{-3}	$0.8 \cdot 10^3$	10^{-6}	26	10^{-8}	$0.7 \cdot 10^7$	10^3
13	10^{-2}	$0.7 \cdot 10^4$	10^{-6}	27	10^{-7}	$0.6 \cdot 10^8$	10^3
14	10	$0.6 \cdot 10^7$	10^{-6}	28	10^{-3}	$0.5 \cdot 10^{12}$	10^3

33) Определить суммарное возрастание энтропии изделия в процессе эксплуатации ΔS , если известны зависимости скорости производства

энтропии в процессе эксплуатации микро- и наноструктуры и срока службы изделия τ .

№ варианта	Зависимости скорости производства энтропии ν , Дж·(град·с) ⁻¹	τ , с
1	$\nu = \nu_0 = \text{const}$ ($\nu_0 = 10^{-9}$ Дж·(град·с) ⁻¹)	$2 \cdot 10^8$
2	$\nu = \nu_0 + \beta t$ ($\nu_0 = 10^{-10}$ Дж·(град·с) ⁻¹), $\beta = 10^{-15}$ Дж·град ⁻¹ ·с ⁻²	10^7
3	$\nu = \nu_0 + \gamma t^2$ ($\nu_0 = 10^{-9}$ Дж·(град·с) ⁻¹), $\gamma = 10^{-7}$ Дж·с ⁻³ ·град ⁻¹	$5 \cdot 10^7$
4	$\nu = \nu_0 + \beta t + \gamma t^2$ ($\nu_0 = 10^{-8}$ Дж(град·с) ⁻¹), $\beta = 10^{-8}$ Дж·град ⁻¹ ·с ⁻² $\gamma = 10^{-16}$ Дж·с ⁻³ ·град ⁻¹	10^8
5	$\nu = \nu_0 e^{\varepsilon t}$ ($\nu_0 = 10^{-6}$ Дж(град·с) ⁻¹), $\varepsilon = 10^{-8}$ с ⁻¹	10^9
6	$\nu = \nu_0 e^{-\varepsilon t}$ ($\nu_0 = 10^{-6}$ Дж(град·с) ⁻¹), $\varepsilon = 10^{-9}$ с ⁻¹	$5 \cdot 10^8$
7	$\nu = \nu_0 + \gamma t^2 - \sigma e^{\varepsilon t}$ ($\nu_0 = 10^{-6}$ Дж(град·с) ⁻¹), $\gamma = 10^{-16}$ Дж·с ⁻³ ·град ⁻¹ , $\varepsilon = 10^{-8}$ с ⁻¹	10^7

34) Определить срок службы изделия τ_k , если известны энтропия изделия с учетом технологического фактора H_t , критическая энтропия H_k и параметры скорости производства энтропии ν_0 , β , n .

№ варианта	H_t , дит	H_k/H_t	ν_0 , Дж(град·с) ⁻¹	β , Дж(град) ⁻¹	n	Примечание
Стационарный процесс деградации изделия: $\nu = \text{const}$						
1	10^5	1.1	10^{-12}			Температура эксплуатации $T=300$ К Энергия активации процесса деградации
2	10^4	1.1	10^{-12}			
3	10^3	1.1	10^{-12}			
4	10^2	1.1	10^{-12}			
5	10^1	1.1	10^{-12}			

						$W=2.0$ эВ
6	10^5	11	10^{-8}			Температура эксплуатации $T=600$ К Энергия активации процесса деградации $W=1.4$ эВ
7	10^4	11	10^{-8}			
8	10^3	11	10^{-8}			
9	10^2	11	10^{-8}			
10	10^1	11	10^{-8}			
Нестационарный линейный процесс деградации изделия: $\nu = \nu_0 + \beta t$						
11	10^5	1.1	10^{-12}	1		Температура эксплуатации $T=300$ К Энергия активации процесса деградации $W=1.5$ эВ
12	10^4	1.1	10^{-12}	5		
13	10^3	1.1	10^{-12}	10		
14	10^2	1.1	10^{-12}	20		
15	10^1	1.1	10^{-12}	50		
16	10^5	11	10^{-8}	1		Температура эксплуатации $T=600$ К Энергия активации процесса деградации $W=1.1$ эВ
17	10^4	11	10^{-8}	2		
18	10^3	11	10^{-8}	10		
19	10^2	11	10^{-8}	20		
20	10^1	11	10^{-8}	50		
Нестационарный нелинейный процесс деградации изделия: $\nu = \nu_0 + \beta t^n$						
21	10^5	1.1	10^{-12}	1	0.1	Температура эксплуатации $T=300$ К Энергия активации процесса деградации
22	10^4	1.1	10^{-12}	5	1	
23	10^3	1.1	10^{-12}	10	2	
24	10^2	1.1	10^{-12}	20	3	
25	10^1	1.1	10^{-12}	50	4	

						$W=1.3$ эВ
26	10^5	11	10^{-8}	1	0.1	Температура эксплуата- ции $T=600$ К Энергия активации процесса деградации $W=0.9$ эВ
27	10^4	11	10^{-8}	2	1	
28	10^3	11	10^{-8}	10	2	
29	10^2	11	10^{-8}	20	3	
30	10^1	11	10^{-8}	50	4	

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

Введение

1. *Taniguchi, N.* On the basic concept of nanotechnology [Text] / *N. Taniguchi.* // Proc. Int. Conf. Prog. Eng. Part 11. Tokyo: Jap. Soc. Pres. Eng., 1974.
2. *Фейнман, Р.* Внизу полным-полно места: приглашение в новый мир физики [Текст] / *Р. Фейнман* // Российский химический журн. – 2002. – Т.46. – №5. – С. 4-6.
3. *Гусев, А.И.* Наноматериалы, наноструктуры, нанотехнологии [Текст] / *А.И. Гусев.* – М.: Физматлит, 2005. – 416 с.
4. *Пул, Ч.* Нанотехнологии [Текст] / *Ч. Пул, Ф. Оунс*; пер. с англ.; под ред. И.П. Чернова. – М.: Техносфера, 2005. – 336 с.
5. Нанотехнологии в электронике [Текст] / под ред. *Ю.А. Чаплыгина.* – М.: Техносфера. 2005. – 446 с.
6. *Малинецкий, Г.Г.* Нанотехнологии. От алхимии к химии и дальше [Текст] / *Г.Г. Малинецкий* // Интеграл. – 2007. – № 5. – С. 4-5.
7. *Жоаким К.* Нанонауки. Невидимая революция [Текст] / *К. Жоаким, Л. Плевр*; пер. с англ.; под ред. В.А. Смирнова. – М.: КоЛибри, 2009. – 286 с.
8. *Борисенко, В.Е.* Нанoeлектроника [Текст] / *В.Е. Борисенко, А.И. Воробьева, Е.А. Уткина.* – М.: Бином. Лаборатория знаний, 2009. – 223 с.
9. *Лен, Ж.-М.* Супрамолекулярная химия. Концепции и перспективы [Текст] / *Ж.-М. Лен*; пер. с англ.; под ред. П.А. Ларионова. – Новосибирск: Наука, Сибир. отд-ние РАН, 1998. – 333 с.
10. *Хакен, Г.* Синергетика. Иерархии неустойчивостей в самоорганизующихся системах и устройствах [Текст] / *Г. Хакен*; пер. с англ.; под ред. Ю.Л. Климантовича. – М.: Мир, 1985. – 424 с.
11. *Гленсдорф, П.* Термодинамическая теория структуры, устойчивости и флуктуаций [Текст] / *П. Гленсдорф, И. Пригожин*; пер. с англ.; под ред. Ю.А. Чизмаджиева. – М.: Мир, 1973.
12. *Валиев, К.А.* Квантовые компьютеры: надежды и реальность [Текст] / *К.А. Валиев, А.А. Кокин.* – М., Ижевск: Регулярная и хаотическая динамика, 2004. – 320 с.
13. Физика квантовой информации [Текст] / под ред. *Д. Боумейстера, А. Экерта, А. Цайлингера*; пер. с англ. С.П. Кулика и Т.А. Шмаонова. – М.: Постмаркет, 2002. – 376 с.

По разделу №1

1. *Фейнман, Р.* Фейнмановские лекции по физике [Текст]: вып. 8,9: Квантовая механика / *Р. Фейнман, Р. Лейтон, М. Сэнде*; пер. с англ.; под ред. Я.А. Смородинского. – М.: Изд-во ЛКИ, 2008. – 528 с.
2. *Ландау, Л.* Квантовая механика (нерелятивистская теория) [Текст] / *Л. Ландау, Е. Лифшиц*. – М.: Физматгиз, 2004. – 800 с.
3. *Бейдер, Р.* Атомы в молекулах. Квантовая теория [Текст] / *Р. Бейдер*; пер. с англ.; под ред. А.В. Когана. – М.: Мир, 2001. – 532 с.
4. *Белоусов, Ю.М.* Курс квантовой механики. Нерелятивистская теория [Текст] / *Ю.М. Белоусов*. – М.: Физматгиз, 2006. – 547 с.
5. *Борисенко, В.Е.* Нанозлектроника [Текст] / *В.Е. Борисенко, А.И. Воробьева, Е.А. Уткина*. – М.: Бином. Лаборатория знаний, 2009. – 223 с.
6. *Пригожин, И.* Современная термодинамика. От тепловых двигателей до диссипативных структур [Текст] / *И. Пригожин, Д. Кондепуди*; пер. с англ.; под ред. Ю.А. Чизмаджиева. – М.: Мир, 2002. – 464 с.
7. *Епифанов, Г.И.* Физические основы микроэлектроники [Текст] / *Г.И. Епифанов*. – М.: Советское радио, 1971. – 376 с.
8. *Блейкмор, Д.* Физика твердого тела [Текст] / *Д. Блейкмор*; пер. с англ.; под ред. Д.Г. Андриянова и В.И. Фистуля. – М.: Мир, 1988. – 608 с.
9. *Физика квантовой информации* [Текст] / под ред. *Д. Боумейстера, А. Экерта, А. Цайлингера*; пер. с англ. С.П. Кулика и Т.А. Шмаонова. – М.: Постмаркет, 2002. – 376 с.
10. *Степанов, Н.Ф.* Квантовая механика и квантовая химия [Текст] / *Н.Ф. Степанов*. – М.: Мир, 2001. – 519 с.
11. *Phys. Rev. Lett* [Test] / *M.N. Baibich, J.M. Broto, A. Fert [и др.]* – 1988.61(21). – P2472-2475.
12. *Наноматериалы. Нанотехнологии. Наносистемная техника* [Текст] / под ред. П.П. Мальцева. – М.: Техносфера, 2006. – 367 с.
13. *Kejn, B. Nature* [Test] / *B. Kejn*. – 1998. – Vol. 393. – P. 133-137.
14. *Ganithcev, S. Phys. Rev. Lett.* [Test] / *S. Ganithcev*. – 2001. – Vol. 86, P. 4358-4361.
15. *Perpendicular hot electron spin-valve effect in a new magnetic field sensor-valve transistor* [Test] / *D.J. Monsma, J.C. Lodder, T.J.A. Popma [и др.]* // *Phys. Rev. Lett.* – 1995. – Vol. 74. – P. 5260–5263.
16. *Electrical spin injection ferromagnetic semiconductor heterostructure* [Test] / *Y. Ohno, D.K. Young, B. Beschoten [и др.]* // *Nature (London)*. – 1999. – Vol. 402. – P. 790–792.

17. *Johnson, M.* Bipolar spin switch [Текст] / *M. Johnson // Science.* – 1993. – Vol. 260. – P. 320–323.

По разделу №2

1. *Фейнман, Р.* Фейнмановские лекции по физике [Текст]: Вып. 4: Кинетика. Теплота. Звук / *Р. Фейнман, Р. Лейтон, М. Сэнде*; пер. с англ.; под ред. Я.А. Смородинского. – М.: Изд-во ЛКИ, 2008. – 386 с.

2. *Френкель, Я.И.* Кинетическая теория жидкостей [Текст] / *Я.И. Френкель.* – М.-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2004. – 424 с.

3. *Фейнман, Р.* Фейнмановские лекции по физике [Текст]: Вып. 7: Физика сплошных сред / *Р. Фейнман, Р. Лейтон, М. Сэнде*; пер. с англ.; под ред. Я.А. Смородинского. – М.: Изд-во ЛКИ, 2008. – 386 с.

4. *Елифанов, Г.И.* Физические основы микроэлектроники [Текст] / *Г.И. Елифанов.* – М.: Советское радио, 1971. – 376 с.

5. *Ансельм, А.И.* Введение в теорию полупроводников [Текст] / *А.И. Ансельм.* – СПб.: Лань, 2008. – 624 с.

6. *Левич, В.Г.* Курс теоретической физики [Текст] / *В.Г. Левич.* – М.: Физматгиз, 1999. – 695 с.

7. *Новиков, Ф.А.* Дискретная математика [Текст] / *Ф.А. Новиков.* – СПб.: Питер, 2001. – 304 с.

8. *Китель, Ч.* Введение в физику твердого тела [Текст] / *Ч. Китель*; пер. с англ.; под ред. А.А. Гусева. – М.: Главная ред. физ.-мат. литературы изд-ва «Наука», 1999. – 792 с.

9. *Блейкмор, Д.* Физика твердого тела [Текст] / *Д. Блейкмор*; пер. с англ.; под ред. Д.Г. Андриянова и В.И. Фистуля. – М.: Мир, 1988. – 608 с.

10. *Тамм, И.Е.* Основы теории электричества [Текст] / *И.Е. Тамм.* – М.: Физматгиз, 2002. – 563 с.

11. *Борисенко, В.Е.* Нанoeлектроника [Текст] / *В.Е. Борисенко, А.И. Воробьева, Е.А. Уткина.* – М.: Бином. Лаборатория знаний, 2009. – 223 с.

12. *Жабреев, В.А.* Введение в нанотехнологию (общие сведения, понятия и определения) [Текст]: учеб. пособие / *В.А. Жабреев, В.И. Марголин, В.С. Павельев.* – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2007. – 172 с.

По разделу №3

1. *Стратанович, Р.Л.* Теория информации [Текст] / *Р.Л. Стратанович.* – М.: Сов. радио, 1975. – 424 с.

2. *Бриллюен, Л.* Наука и теория информации [Текст] / *Л. Бриллюен*; пер. с англ. – М.: Физматгиз, 1960. – 392 с.
3. *Новицкий, П.В.* Основы информационной теории измерительных устройств [Текст] / *П.В. Новицкий*. – Л.: Энергия, 1968. – 248 с.
4. *Эммануэль, Н.М.* Курс химической кинетики [Текст] / *Н.М. Эммануэль, Д.Г. Кнорре*. – М.: Высшая школа, 1998. – 372 с.
5. *Пригожин, И.* Современная термодинамика. От тепловых двигателей до диссипативных структур [Текст] / *И. Пригожин, Д. Кондепуди*; пер. с англ. – М.: Мир, 2002. – 461 с.
6. *Эйринг, Г.* Основы химической кинетики [Текст] / *Г. Эйринг, С.Г. Лин, С.М. Лин*. – М.: Мир, 1999. – 528 с.
7. *Шеннон, К.* Работы по теории информации и кибернетики [Текст] / *К. Шеннон*; пер.с англ. – М.: ИЛ, 1963. – 547 с.
8. *Кадомцев, Б.Б.* Динамика и информация [Текст] / *Б.Б. Кадомцев* // Избранные труды. – 2003. – Т. 2. – С. 508-515.
9. *Саноян, А.Г.* Энтропийные модели и критерии оценки качества обработки материалов в микро- и нанотехнологиях [Текст] / *А.Г. Саноян* // ФХОМ – 2006. – № 6. – С. 69 – 74.
10. *Саноян, А.Г.* Энтропийные модели конструкторско-технологической деятельности [Текст] / *А.Г. Саноян* // Мировое сообщество: проблемы и пути решения. – Уфа. – 1999. – Вып. № 4-5. – С.107–116.
11. *Саноян, А.Г.* Энтропийные модели биологических структур [Текст]: сб. науч. тр. / *А.Г. Саноян*. – Самара: Самар. мед. ин-т «Реавиз», 2004. – С. 168 – 176.
12. *Воробьев, В.Л.* Термодинамические основы диагностики и надежности микроэлектронных устройств [Текст] / *В.Л. Воробьев*. – М.: Наука, 1989. – 160 с.
13. *Гнеденко, Б.В.* Курс теории вероятностей [Текст] / *Б.В. Гнеденко*. – М.: Наука, 1999. – 408 с.
14. *Стратанович, Р.Л.* Количество информации и энтропия отрезков стационарных гауссовых процессов [Текст] / *Р.Л. Стратанович* // Проблемы передачи информации. – 1967. – Т.3. – Вып. 2. – С. 92–104.
15. *Седов, Л.И.* Механика сплошной среды [Текст] / *Л.И. Седов*. – М.: Наука, 1994. – Т.1. – 528 с.
16. *Ландау, Л.Д.* Статистическая физика [Текст] / *Л.Д. Ландау, Е.М. Лифшиц*. – М.: Наука, 1976. – Ч. 1. – 584 с.
17. *Глансдорф, П.* Термодинамическая теория структуры, устойчивости и флуктуаций [Текст] / *П. Глансдорф, И. Пригожин*. – М.: Мир, 2001. – 398 с.

18. *Барыбин, А.А. Физико-технологические основы электроники [Текст] / А.А. Барыбин, В.Г. Сидоров. – СПб.: Изд-во «Лань», 2001. – 272 с.*
19. *Эммануэль, Н.М. Курс химической кинетики [Текст] / Н.М. Эммануэль, Д.Г. Кнорре. – М.: Высшая школа, 1998. – 372 с.*
20. *Федюкин, В.Н. Методы термодинамики в задачах теории надежности [Текст] / В.Н. Федюкин // Основные вопросы теории и практики надежности. – М.: Сов.радио, 1985. – С. 228-246.*
21. *Волькенштейн, М.В. Энтропия и информация [Текст] / М.В. Волькенштейн. – М.: Наука, 1986. – 579 с.*
22. *Фейнман, Р. Фейнмановские лекции по физике [Текст]: Вып. 4: Кинетика / Р. Фейнман, Р. Лейтон, М. Сэнде; пер. с англ.; под ред. Я.А. Смородинского. – М.: Книжный дом «ЛИБРОКОМ», 2009. – 264 с.*
23. *Климонтович, Ю.Л. Энтропия и информация открытых систем [Текст] / Ю.Л. Климонтович // УФН. – 1999 – Т. 169. – № 4. – С. 443 – 452.*
24. *Стратанович, Р.Л. Нелинейная неравновесная термодинамика [Текст] / Р.Л. Стратанович. – М.: Наука, 1994. – 473 с.*
25. *Базаров, И.П. Термодинамика [Текст]: учебник.- 3-е изд., перераб. и доп. / И.П. Базаров. – М.: Высшая школа, 2003. – 352 с.*
26. *Николис, Г. Самоорганизация в неравновесных системах [Текст] / Г. Николис, И. Пригожин; пер. с англ. – М.: Мир, 1979. – 512 с.*
27. *Валиев, К.А. Физика субмикронной литографии [Текст] / К.А. Валиев. – М.: Наука, 1990. – 541 с.*
28. *Займан, Дж. Модели беспорядка [Текст] / Дж. Займан; пер.с англ. – М.: Мир, 1982. – 592 с.*
29. *Волькенштейн, М.В. Молекулярная биофизика [Текст] / М.В. Волькенштейн. – М.: Наука, 1975. – 616 с.*
30. *Эйген, М. Самоорганизация материи и эволюция биологических макромолекул [Текст] / М. Эйген. – М.: Мир, 2003. – 284 с.*
31. *Колмогоров, А.Н. Три подхода к определению понятия «Количество информации». Новое в жизни, науке, технике [Текст] / А.Н. Колмогоров. – Сер. «Математика, кибернетика». – 1991. – № 1. – С.24 – 29.*
32. *Эбелинг, В. Образование структур при необратимых процессах: Введение в теорию диссипативных структур [Текст] / В. Эбелинг. – М.-Ижевск: Институт компьютер. исследов., Мир, 2004. – 256 с.*

ПРИЛОЖЕНИЕ № 1. Справочные данные общего характера

А) Некоторые основные физические постоянные

Наименование	Обозначение	Значение
Скорость света в вакууме	c	$2.9979 \cdot 10^8$ м/с
Постоянная Больцмана	k	$1.3806 \cdot 10^{-23}$ Дж/К
Постоянная Планка	h	$1.3806 \cdot 10^{-34}$ Дж·с
Электрическая постоянная	ε_0	$8.8541 \cdot 10^{-12}$ Ф/м
Магнитная постоянная	μ_0	$1.2566 \cdot 10^{-6}$ Гн/м
Газовая постоянная	R	8.3144 Дж/(К·моль)
Постоянная Авогадро		$6.0220 \cdot 10^{23}$ моль ⁻¹
Элементарный электрический заряд	e	$1.6021 \cdot 10^{-19}$ Кл
Масса покоя электрона	m_e	$9.1095 \cdot 10^{-31}$ кг
Масса покоя протона	m_p	$1.6726 \cdot 10^{-27}$ кг
Атомная единица массы	а.е.м.	$1.6605 \cdot 10^{-27}$ кг

Б) Единицы энергии и работы

Дж	ккал	эВ
1	$2.39 \cdot 10^{-4}$	$6.24 \cdot 10^{18}$
$4.19 \cdot 10^3$	1	$2.61 \cdot 10^{22}$
$1.6 \cdot 10^{-19}$	$3.83 \cdot 10^{-23}$	1

В) Приставки к единицам системы СИ

Приставка	Обозначение	Значение	Приставка	Обозначение	Значение
дека	да	10^1	деци	д	10^{-1}
гекто	г	10^2	санتي	с	10^{-2}
кило	к	10^3	милли	м	10^{-3}
мега	М	10^6	микро	мк	10^{-6}
гига	Г	10^9	нано	н	10^{-9}
тера	Т	10^{12}	пико	п	10^{-12}
пета	П	10^{15}	фемто	ф	10^{-15}
экса	Э	10^{18}	атто	а	10^{-18}

Г) Буквы греческого алфавита

α – альфа	ι – йота	ρ – ро
β – бета	χ – каппа	σ – сигма
γ – гамма	λ – ламбда	τ – тау
δ – дельта	μ – мю	υ – ипсилон
ε – эpsilon	ν – ню	φ, ϕ – фи
ζ – дзета	ξ – кси	χ – хи
η – эта	\omicron – омикрон	ψ – пси
θ – тета	π – пи	ω – омега

ПРИЛОЖЕНИЕ № 2. Справочные данные о параметрах некоторых материалов

(для случая макроскопических объемов вещества)

А) Ширина запрещенной зоны некоторых полупроводников и диэлектриков (материал) / (электронвольт)

InSb 0.18	InAs 0.35	Ge 0.66	GaSb 0.68	Si 1.12	InP 1.34
GaAs 1.42	CdTe 1.44	AlSb 1.60	CdSe 1.7	InN 1.9	AlAs 2.15
ZnTe 2.26	GaP 2.27	CdS 2.42	AlP 2.45	ZnSe 2.67	GaN 3.42
ZnS 3.58	Si ₃ N ₄ ≈ 5	AlN 6.2	Al ₂ O ₃ ≈ 8	SiO ₂ 8.8	CaF ₂ 12.2

Б) Подвижность носителей зарядов в некоторых собственных полупроводниках (при температуре 300 К)

Тип носителя заряда	Подвижность носителя заряда, м ² / В·с			
	Si	Ge	InSb	Cu (для справки)
Электроны	0.135	0.38	7.70	0.005
Дырки	0.040	0.18	0.13	-

В) Энергия Ферми некоторых металлов

Металл	Na	Li	Ag	Cu
W_F , ЭВ	3.1	4.7	5.5	7.1

Г) Электропроводность некоторых материалов (при температуре 300 К)

Проводник	σ (Ом·м)	Проводник	σ (Ом·м)
Серебро	$1.6 \cdot 10^{-8}$	Железо	$1.0 \cdot 10^{-7}$
Медь	$1.7 \cdot 10^{-8}$	Платина	$1.1 \cdot 10^{-7}$
Золото	$2.2 \cdot 10^{-8}$	Сталь (литье)	$1.3 \cdot 10^{-7}$
Алюминий	$2.7 \cdot 10^{-8}$	Свинец	$2.1 \cdot 10^{-7}$
Магний	$4.4 \cdot 10^{-8}$	Нейзильбер (сплав)	$3.3 \cdot 10^{-7}$

Иридий	$4.7 \cdot 10^{-8}$	Манганин (сплав)	$4.3 \cdot 10^{-7}$
Молибден	$5.4 \cdot 10^{-8}$	Константан (сплав)	$5.0 \cdot 10^{-7}$
Вольфрам	$5.5 \cdot 10^{-8}$	Ртуть	$9.6 \cdot 10^{-7}$
Цинк	$5.9 \cdot 10^{-8}$	Чугун (серый)	$1.0 \cdot 10^{-6}$
Никель	$8.7 \cdot 10^{-8}$	Графит	$8.0 \cdot 10^{-6}$

Диэлектрик	σ (Ом·м)	Диэлектрик	σ (Ом·м)
Вода (дистиллированная)	$\approx 10^4$	Плексиглаз	10^{13}
Вода (морская)	0.3	Полистирол	10^{16}
Стекло	$10^{10} - 10^{12}$	Полихлорвинил	10^{12}
Стекло (кварцевое)	10^{16}	Полиэтилен	$10^{10} - 10^{13}$
Слюда	10^{14}	Парафин	$10^{14} - 10^{16}$
Керамика	$10^{12} - 10^{16}$	Масло (трансформатор.)	$10^{10} - 10^{13}$
Янтарь	10^{18}	Масло (силиконовое)	10^{14}

ПРИЛОЖЕНИЕ № 3. Общие рекомендации по решению задач и упражнений по теме «контактные явления в твердых структурах»

В настоящем приложении рассматриваются вопросы построения энергетических диаграмм в контактных системах с участием металлов и полупроводников. Материалы представлены в виде задач и упражнений (преимущественно в графическом виде).

В ряде графических примеров специально допущены ошибки в построении энергетических диаграмм. В этих «неправильных» случаях необходимо обосновать физическую суть допущенных в приведенных рисунках противоречий.

Контактные явления в системах «Металл №1 – Металл №2»

Построение энергетических диаграмм производится на основании зонных энергетических диаграмм металлов, принимающих участие в создании контактной системы. При построении энергетических диаграмм в контактных системах с участием только металлов необходимо учитывать следующие моменты:

- В начальный момент времени ($t = 0$) оба вещества имеют одинаковое расположение "0" энергетического уровня.
- После установления равновесия в системе оба металла имеют одинаковое расположение уровня Ферми.
- Вещество с меньшей работой выхода выступает в качестве эмиттера электронов, а вещество с большей работой выхода выступает в качестве акцептора электронов.
- После установления равновесия в контактной системе приповерхностная область (толщиной порядка 1 ангстрем) материала с меньшей работой выхода приобретает положительный заряд, а материал с большей работой выхода – отрицательный заряд.
- Величина поверхностного заряда, появляющаяся на поверхностях раздела (благодаря перешедшим электронам из вещества с меньшей работой выхода в вещество с большей работой выхода), в результате установления равновесия составляет по порядку величины 10^{12} электрон/см² (для справки: количество электронов, находящихся

на поверхности металла, составляет по порядку величины 10^{15} электрон/см², а в объеме – 10^{23} электрон/см³).

- После установления равновесия в контактной системе в вакуумном зазоре устанавливается контактная разность потенциалов (КРП), обусловленная разностью работ выхода веществ, принимающих участие в создании контактной системы.

- Отрицательный потенциал внешнего источника электропитания (батареи) смещает уровень Ферми вещества вверх на величину, равную qU , где q – заряд электрона; U – потенциал внешнего источника.

Контактные явления в системах «Металл – Полупроводник»

Построение энергетических диаграмм производится на основании зонных энергетических диаграмм металлов и полупроводников, принимающих участие в создании контактной системы. При построении энергетических диаграмм в контактных системах с участием полупроводников (хотя бы в качестве одного из элементов системы) необходимо учитывать следующие моменты:

- Общая методология построения энергетических диаграмм контактных систем с участием полупроводников аналогична методике построения диаграмм для систем «Металл №1 – Металл №2».

- Основное отличие энергетических диаграмм контактных систем с участием полупроводников (в отличие от диаграмм контактных систем с участием только металлов) состоит в том, что объемная концентрация носителей зарядов в собственных ($10^{10} - 10^{14}$ электрон/см³) и примесных ($10^{15} - 10^{19}$ электрон/см³) полупроводниках гораздо ниже объемной концентрации носителей зарядов в металлах (10^{23} электрон/см³).

- В энергетических диаграммах контактных систем с участием полупроводников не отображается участок, соответствующий области вакуумного зазора (последнее обусловлено тем обстоятельством, что протяженность области вакуумного зазора d (порядка 10 ангстрем) существенно меньше протяженности приповерхностной области объемного заряда в полупроводниках d_n ($10^3 - 10^5$ ангстрем)).

- В свете вышеизложенного значение контактной разности потенциалов, приходящейся на вакуумный зазор, существенно меньше паде-

ния электрического потенциала в приповерхностной области объемного заряда полупроводника.

- Приповерхностная область объемного заряда в полупроводниках может быть обеднена (запорный слой) или обогащена основными носителями зарядов (зависит от того, какие носители зарядов приходят (уходят) в полупроводник в процессе установления термодинамического равновесия).

- На участке, соответствующем приповерхностной области объемного заряда в полупроводниках, отображается фактор изгиба энергетических уровней (зон) как реакция на тот факт, что на этом участке существует электрическое поле, способствующее образованию контактной разности потенциалов.

- Суммарная величина изгибов энергетических зон (проводимости и валентной) равна разности работ выхода полупроводников (или полупроводника и металла), образующих контактную систему.

- В случае гетеропереходов (т.е. систем, образованных путем контакта двух полупроводников, имеющих различную ширину запрещенной зоны) изгиб энергетических зон испытывает разрыв в области контакта двух сред.

Учебное издание

Саянн Ашот Геворкович

**ФИЗИКО-ТЕХНИЧЕСКИЕ
ОСНОВЫ НАНОИНЖЕНЕРИИ**

Учебное пособие

Редакторская обработка Т. К. К р е т и н и н а
Доверстка Т. Е. П о л о в н е в а

Подписано в печать 10.05.2011. Формат 60x84 1/16.
Бумага офсетная. Печать офсетная.
Печ. л. 23,5.
Тираж 100 экз. Заказ . Арт. С-16/2011.

Самарский государственный
аэрокосмический университет.
443086 Самара, Московское шоссе, 34.

Изд-во Самарского государственного
аэрокосмического университета.
443086 Самара, Московское шоссе, 34.