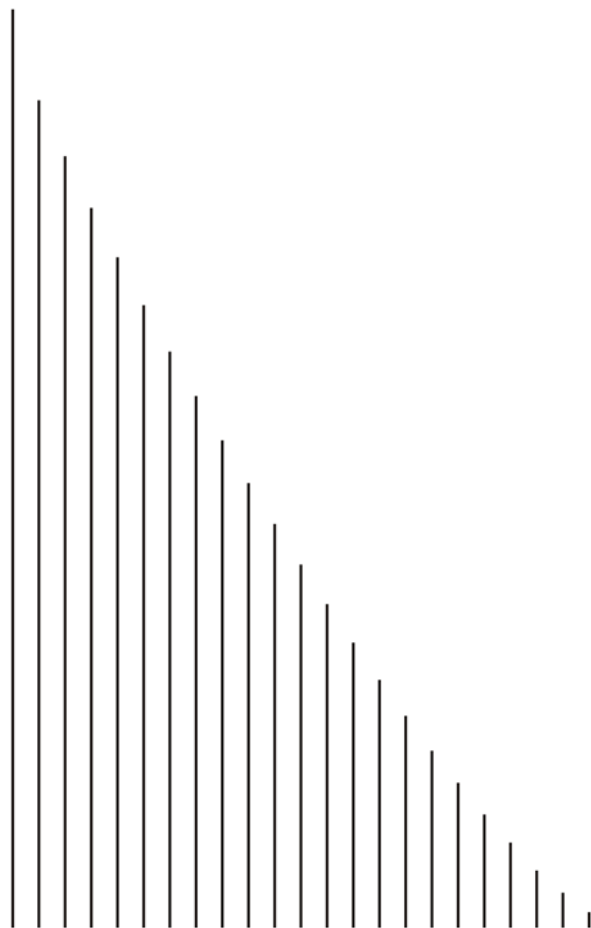
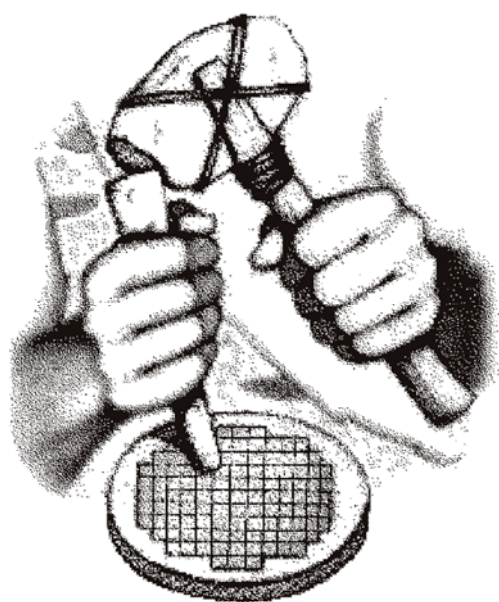


А. Б. Чурилов



ВВЕДЕНИЕ В
наноэлектронику

Учебное пособие

Министерство образования Российской Федерации
Ярославский государственный университет им. П. Г. Демидова

А . Б . Ч у р и л о в

В В Е Д Е Н И Е
В НАНОЭЛЕКТРОНИКУ

Учебное пособие

Ярославль 2002

УДК 621.382.017.7(075.8)

ББК 385я73

Ч93

Рецензенты: кафедра общей физики ЯГПУ им. К.Д. Ушинского;
д-р физ.-мат. наук В.К. Смирнов

Чурилов А.Б. Введение в нанoeлектронику: Учеб. пос. / Яросл. гос.
ун-т. Ярославль, 2002. 132 с.
ISBN 5-8397-0249-8

Учебное пособие является введением в физику и технологию микроэлектроники с субмикронными размерами. Наиболее подробно рассмотрены вопросы, связанные с физическими ограничениями, возникающими при увеличении степени интеграции. Рассмотрены также вопросы физики систем пониженной размерности и приборов нанoeлектроники. Коротко затронуты вопросы, связанные с современной технологией производства микросхем с субмикронными размерами.

Предназначено для студентов старших курсов университетов, специализирующихся в области физики полупроводников и полупроводниковых приборов, микро - и нанoeлектроники. Для усвоения материала необходимо знание общей и теоретической физики в объеме университетского курса, а также физики полупроводников и основ технологии микроэлектроники.

ISBN 5-8397-0249-8

© Ярославский государственный университет, 2002

© Чурилов А. Б., 2002

Учебное издание

Чурилов Анатолий Борисович
Введение в нанoeлектронику

Редактор, корректор А. А. Аладьева
Компьютерный набор, верстка А. Б. Чурилов

Подписано в печать 20.12.02. Формат 60 × 84 ¹/₁₆. Печать офсетная.
Усл. печ. л. 7,67. Уч.-изд. л. 6,07. Тираж 75 экз.

Оригинал макет подготовлен в редакционно-издательском отделе
Ярославского государственного университета.

150000 Ярославль, ул. Советская, 14

Отпечатано на ризографе

ООО "Рио – Гранд"

150000, Ярославль, ул. Свердлова, 18, офис 34.

ВВЕДЕНИЕ

Основные химические элементы, входящие в состав человеческого организма, – это кислород, углерод, водород, азот. Если учитывать элементы, входящие в небольших количествах (такие как сера, фосфор, железо и др.), то их наберется приблизительно еще около 20. Более того, периодическая система элементов Д. И. Менделеева содержит их всего немногим более ста. В связи с этим очень заманчивой кажется решение задачи построения элементов как живой, так и неживой природы, соединяя атом с атомом, молекулу с молекулой и так далее. Решение подобной задачи, или даже части её, имело бы такое же значение для человечества в целом, как, например, использование силы падающей воды, электричество, антибиотики и микроэлектроника [1].

Когда речь заходит об объектах нанометровых размеров, можно привести широкий класс в качестве примеров: от трехатомной молекулы воды до белковых молекул (кислородосодержащих молекул гемоглобина из тысяч атомов) и гигантских молекул ДНК из миллионов атомов.

В последние годы удалось синтезировать наноструктуры различных типов, составляющие основу новых, так называемых высоких технологий. Молекулы различного химического состава и формы удается «собирать» в различные фигуры, придавая вновь образующимся структурам и материалам на их основе новые свойства: электрические, оптические, механические. Контролируя процессы синтеза таких структур, удастся создавать материалы с новыми наборами физико-химических характеристик.

В 1993 году президентом США был создан Национальный научный и технологический совет (National Science and Technology Council - NSTC) [2]. Основной задачей этой организации стала координация действий федерального правительства в области развития науки, технологии и исследования космоса. NSTC возглавляет президент США.

В 1997 году правительство США инвестировало в исследования по нанотехнологиям ~\$116 миллионов, а в 1999

году уже \$260. В 2003 году предполагаемые правительственные ассигнования в США в этой области должны превысить \$600 миллионов. Япония и Европа вкладывают приблизительно такие же средства. Университетские научно – исследовательские центры и лаборатории, а также промышленные компании с помощью NSTC координируют свою деятельность в области исследований по нанотехнологиям.

Нанотехнологии – это новый мир, но не новая область знаний. Характерным примером, иллюстрирующим это, является химический катализ. Это пример «старой нанотехнологии», включающий в себя самый широкий спектр процессов и их продуктов: от получения бензина из нефти до синтеза алмазов из графита. Современные установки катализа широко используют «упорядоченные» металлические и/или керамические наноструктуры, специфическая форма которых позволяет взаимодействовать только с молекулами вполне определенной формы и размера.

Термин «нано» имеет корни в греческом языке и использовался для обозначения карликов. В настоящее время используется как приставка к системным и внесистемным единицам измерения (литр, метр, секунда) для обозначения $1/10^9$ части. В 1977 году американским студентом (в настоящее время - профессором) Дрекслером было придумано слово "нанотехнология" для гипотетической сборки полезных объектов из молекулярных цепочек. Их характерным размером должна была стать одна миллиардная доля метра – нанометр: величина порядка нескольких атомов, построенных в цепочку. А уже в 1990 году за рубежом вышел первый номер толстого журнала "Нанотехнология".

Размеры отдельных элементов в процессоре Intel®Pentium®III достигали 200 нанометров. А современные модели процессоров используют уже технологию со 130 нанометровым разрешением. Постепенно приставка «микро» все чаще и чаще заменяется приставкой «нано», входя в обыденное обращение, используется в названиях фирм, фантастических книгах, фильмах и т. п.

Одним из первых, кто предсказал будущее развитие нанотехнологии, был лауреат Нобелевской премии по физике (1965 г.) Ричард Ф. Фейнман (1918-1988). В 1959 году

представляя в Калифорнийском технологическом институте свои знаменитые лекции по физике, он говорил среди всего прочего и о биологических объектах [3]. Несмотря на свои малые размеры, они активны, могут создавать себе подобных, взаимодействовать с другими объектами строго определенным образом. Более того – биологические объекты хранят информацию. Исключительно привлекательным казалось в то время решение задачи искусственного создания таких объектов. Тогда его слова казались фантастикой. В своей речи, название которой можно перевести как «Там внизу – море места», Фейнман говорил о нанотехнологии тогда, когда этого термина еще не существовало, как, впрочем, еще не существовало и технологии, позволяющей оперировать отдельными атомами на атомарном же уровне (подразумевается возможность опознать отдельный атом, взять его и поставить на другое место). Такая возможность появилась лишь в 81-ом году, когда в швейцарском отделении IBM был разработан сканирующий туннельный микроскоп - прибор, чувствительный к изменениям туннельного тока между поверхностью материала и сверхтонкой иглой. Отечественная наука получила такой инструмент только в 1986 году.

Для того, чтобы стимулировать интерес исследователей к нанотехнологии, Фейнман посулил по 1000 долларов тому, кто «уменьшит» книжную страницу в 25000 раз, и тому, кто поместит электрический моторчик в кубик со стороной в 1/64 дюйма. Широта взгляда одного из последних великих физиков XX столетия на миниатюризацию электроники состоит в том, что было все это в 1959 году, когда термин «чип» еще не был введен в научный лексикон [4]. Основываясь на фундаментальных физических законах, он утверждал, что всю Британскую энциклопедию можно записать на площади порядка булавочной головки. Миллион таких «булавочных головок» займет всего 35 обычных страниц.

Сегодня Foresight Institute обещает \$250 000 тому, кто построит нано - робота («руку», которая сможет оперировать на молекулярном уровне), и тому, кто создаст 8 - битный сумматор, уместящийся в кубике со стороной в 50 нанометров. Цены растут, детали мельчают.

Прошло уже более 40 лет, в течение которых бурно развивалась наука и совершенствовалась микротехнология. Появились новые методы исследований и технологии, такие как сканирующая электронная микроскопия, атомная силовая микроскопия, молекулярно пучковая эпитаксия и многое другое. В настоящее время все более широкое использование в технологиях микро– и наноэлектроники находят процессы самоорганизации. Микроэлектроника позволила создать мощные компьютеры, использующиеся, в свою очередь, для расчетов и моделирования наноструктур и соответствующих физических эффектов и явлений. Это всегда способствовало совершенствованию теоретических представлений и осмыслению экспериментальных результатов, помогало рассчитывать бесконечное количество возможных наноструктур. Экспериментальная возможность манипулирования отдельными атомами создает перспективу для создания квантового компьютера.

Литература

1. *Amato I.* Nanotechnology – Shaping the World Atom by Atom (NSTC report, <http://itri.loyola.edu/nano/IWGN.Public.Brochure/IWGN.Nanotechnology.Brochure.pdf>),
2. http://www.whitehouse.gov/WH/EOP/OSTP/NSTC/html/NSTC_Home.html
3. *Feynman R.* There's Plenty of Room at the Bottom: An Invitation to Enter a New Field of Physics. Talk at the Annual Meeting of the American Physical Society, 29 December 1959. / Reprinted in *Engineering and Science*. 1960. Vol. 23. P. 22-36.
4. В 2000 году Джеку С. Килби была присуждена Нобелевская премия по физике за вклад в открытие и развитие интегральной схемы, получившей название микросхемы или чипа. Благодаря этому открытию стала быстро развиваться микроэлектроника, которая в настоящее время лежит в основе всей современной техники.

1. ФИЗИЧЕСКИЕ ОГРАНИЧЕНИЯ МИНИМАЛЬНЫХ РАЗМЕРОВ ЭЛЕМЕНТОВ МИКРОЭЛЕКТРОНИКИ.

Современная микроэлектроника (МЭ) – это область электроники, объединяющая комплекс физических, радиотехнических и технологических проблем, направленных на создание сложных электронных схем для переработки и передачи информации. Основой современной МЭ являются твердотельные микросхемы, создаваемые на отдельном кристалле – чипе, как правило, кремниевом. Интегральная схема (ИС) может быть представлена как совокупность дискретных элементов, взаимодействующих между собой за счет созданных на чипе межсоединений. Каждый же функциональный элемент ИС обладает различающимися по своим характеристикам однородными областями. Минимально достижимые размеры именно этих областей и являются наиболее интересными с точки зрения физических ограничений [1].

Элементы, из которых состоит ИС, могут быть охарактеризованы следующими параметрами:

N – число активных элементов на чипе;

τ – время переключения элемента между двумя электрическими (логическими) состояниями под действием сигнала;

$\nu = 1/4\tau$ – тактовая частота;

$N\nu$ – функциональное быстродействие (ФБ) схемы;

P – мощность, рассеиваемая в процессе переключения элемента;

$P\tau$ – показатель качества элемента, определяющий количество энергии, рассеиваемой элементом в единичном акте переключения;

d – характерный размер активной области элемента (длина канала, ширина базы), который определяется разрешающей способностью технологического процесса (минимальной шириной линии)

Важнейшими характеристиками ИС, особенно с точки зрения потребителя, являются стоимость бита информации (C), то есть фактически стоимость одного элемента и надежность работы ИС, определяемая вероятностью ее отказа.

Введенные параметры позволяют оценить современное состояние МЭ, сравнить его с прогнозами, сделанными 15-20 лет назад (Табл. 1.1) [1], характеризовать динамику и перспективы развития (Табл. 1.2) [2, 3].

Таблица 1.1

Годы	1980 г.	1985 – 1990 гг.
N , элемент/чип	$10^4 - 10^5$	$3 \cdot 10^5 - 10^6$
τ , с	$10^{-5} - 10^{-6}$	$10^{-8} - 10^{-9}$
ν , Гц	10^5	$2,5 \cdot 10^7$
ФБ, элемент·Гц	$10^{10} - 10^{11}$	10^{13}
P , Вт	10^{-4}	10^{-5}
$P\tau$, Дж	$10^{-9} - 10^{-10}$	$10^{-13} - 10^{-14}$
C , доллар/бит	10^{-4}	10^{-5}
d , мкм	3,5	0,5

Из приведенных таблиц видно, что прогресс в МЭ достигается за счет уменьшения размеров элементов и соответствующего увеличения плотности их упаковки. Улучшение операционных характеристик ИС за счет миниатюризации элементов и увеличения плотности их упаковки в несколько раз более существенно, чем тот же эффект за счет усовершенствования схемотехники и увеличения размеров кристалла. Увеличение числа элементов на кристалле, при практически неизменной его стоимости, сопровождается падением стоимости отдельного элемента.

Таким образом, если развитие микроэлектроники пойдет так, как предсказывают данные, приведенные в Таблице 1.2, к 2014 году мы будем иметь:

- минимальный характерный размер элемента ИС ~ 20-30 нанометров;
- более 10^8 транзисторов/см² для логических ИС;
- более 10^{10} бит/см² для ИС памяти;
- стоимость менее чем $2,7 \cdot 10^{-7}$ доллар/транзистор для логических ИС;
- стоимость менее чем $7 \cdot 10^{-9}$ доллар/бит для ИС памяти;
- рабочую частоту ~ 13 ГГц;
- Рассеиваемая мощность для $1,4 \cdot 10^9$ транзисторов микропроцессорных ИС менее 180 Вт.

Таблица 1.2

Параметры, характеризующие состояние и предполагаемую динамику развития микроэлектроники до 2014 г. (Semiconductor Industry Association – SIA Roadmap за 1999 г.) [3]

Год получения первого результата	2001	2003	2005	2008	2011	2014
Память						
▪ минимальный размер элемента DRAM	150	120	100	70	50	35
▪ бит на чип	2G	4G	8G	24G	68G	194G
▪ размер чипа (мм ²)	438	480	526	603	691	792
▪ плотность (Гбит/см ²)	0,49	0,89	1,63	4,03	9,94	24,50
▪ стоимость одного бита (\$10 ⁻⁸)	21,0	11,0	5,3		0,7	
Логика						
▪ минимальный размер элемента MPU (длина затвора, 10 ⁻⁹ м)	100	80	65	45	30	20
▪ транзисторов на чип	48M	95M	190M	539M	1,5G	4,3G
▪ размер чипа (мм ²)	340	372	408	468	536	615
▪ плотность (млн. транзисторов / см ²)	13	24	44	109	269	664
▪ стоимость одного транзистора (\$10 ⁻⁸)	686	434	217		27	
Диаметр пластины (мм)	300	300	300	300	300	450
Количество слоев межсоединений	7	8	8-9	9	9-10	10
Локальная частота (ГГц)	1,7	2,5	3,5	6,0	10,0	13,5
Частота передачи данных между чипами (ГГц)	1,4	1,7	2,0	2,5	3,0	3,6
Рассеиваемая мощность (Вт)	115	140	160	170	174	183
Стоимость завода для производства микросхем	<div>~ 2 млрд Евро</div> <div>более 5 млрд Евро</div>					

С начала семидесятых годов развитие микроэлектроники происходит в соответствии с законом Мура (удвоение производительности ИС каждые 18 месяцев). Это увеличение было достигнуто главным образом за счет уменьшения характерных размеров элементов, оптимизации и совершенствования существующих и внедрения новых технологий. Специалисты, однако, с потрясающей регулярностью твердят о прекращении действия закона Мура - практически с того момента, как он был сформулирован. Однако теперь уже большинство специалистов соглашается с тем, что в ближайшие десять лет закономерность, описывающая развитие основной технологии, перестанет существовать. Безусловно, уменьшение размеров транзисторов не может продолжаться вечно по той простой причине, что если ситуация будет развиваться в соответствии с законом Мура, то в конечном итоге это приведет к созданию устройств, в которых размер критически важных элементов окажется меньше атома. (Рис.1.1)

В соответствии с этим законом в ближайшие 5-7 лет характерный размер элементов достигнет 100 нм, а заряд, определяющий состояние логических элементов, понижен до величины менее 1000 зарядов свободного электрона. Физические (квантовые эффекты и особенности поведения при малых токах) и технологические ограничения (повышение рассеиваемой мощности, усложнение схемотехнических решений, туннельные токи) могут в недалеком будущем стать серьезным препятствием для дальнейшего развития микроэлектроники на основе обычного уменьшения характерных размеров элементов ИС. С другой стороны, технологические проблемы, вместе с необходимостью сильного увеличения объемов инвестиций для продвижения стандартной CMOS технологии, могут значительно облегчить продвижение принципиально новых, альтернативных технологий, элементов и устройств.

Проблема физических ограничений на минимальные размеры элементов и максимальную степень интеграции в микроэлектронике возникла одновременно с созданием ИС. В последнее время эти вопросы становятся особенно актуальными в связи с прогрессом в технологии МЭ, позволяющим создавать отдельные элементы размером 100 нм и менее. Таким образом,

эти величины становятся сравнимыми с характерными длинами, определяющими работу элемента: ширина области пространственного заряда, длина пробега и т.п.

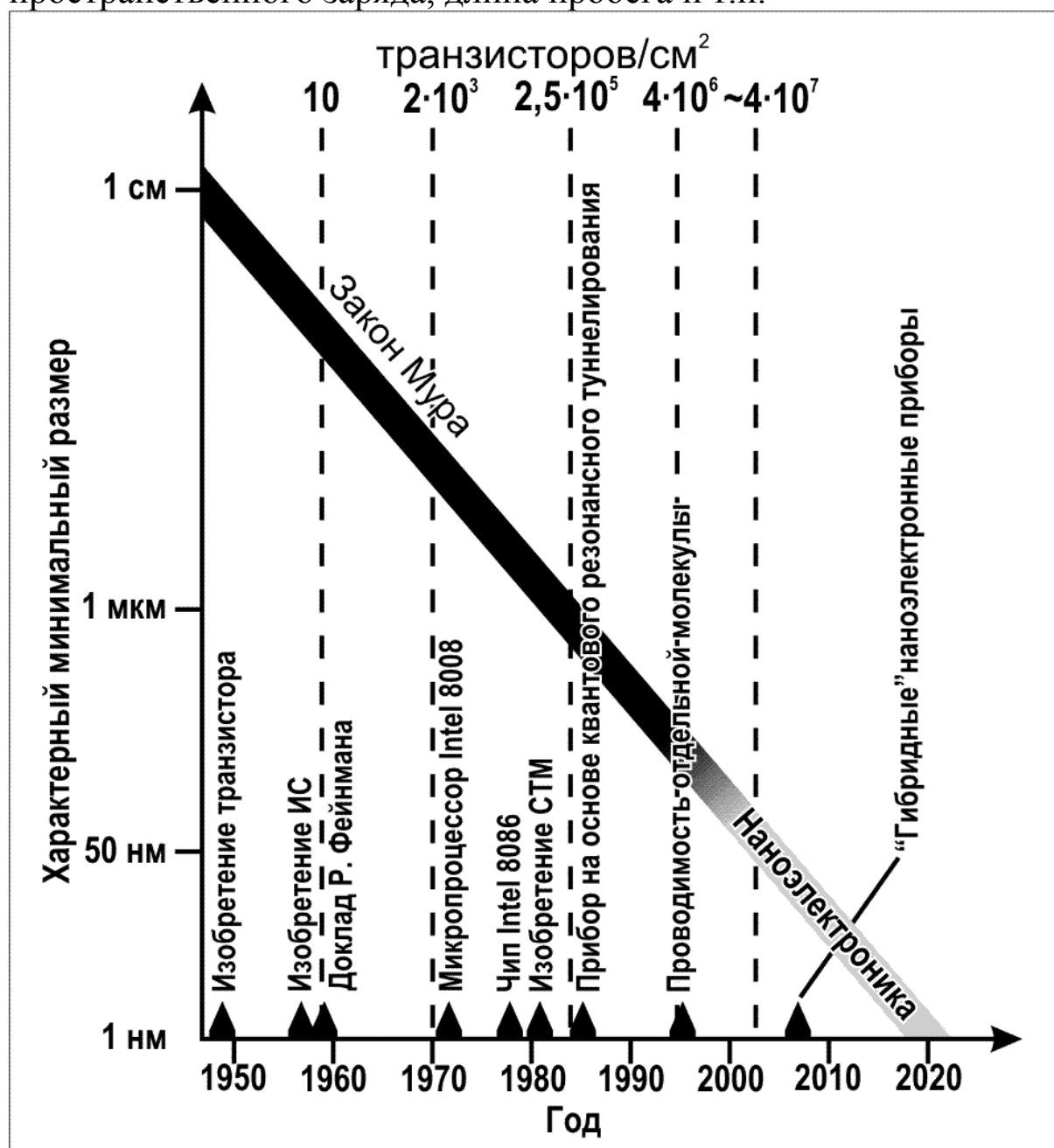


Рис.1.1. Развитие микроэлектроники. Кроме прямой, иллюстрирующей закон Мура, на графике отмечены основные этапы развития микроэлектроники [4].

С другой стороны, появляется возможность использовать новые физические эффекты, проявляющиеся как результат размерного квантования носителей заряда в элементах ИС нанометровых размеров. Именно поэтому вопрос о физических

ограничениях может рассматриваться только при обсуждении конкретных элементов ИС.

Для достижения высокой степени интеграции необходимо:

- уметь создавать элементы достаточно малых размеров;
- добиваться нормального функционирования каждого отдельного элемента;
- добиваться нормального функционирования всех элементов в составе ИС.

Следовательно, возможные ограничения на размеры элементов можно разбить на три группы.

1. Физические ограничения, возникающие в технологии производства ИС (ограничения, связанные с рассеянием экспонирующего луча при литографии, флуктуации легирующих примесей, однородность и минимальные размеры при использовании пучковых технологий и т.п.).

2. Физические ограничения работоспособности отдельных элементов (ограничения, обусловленные смыканием p - n переходов, пробоем сверхтонких диэлектрических слоев, разогревом носителей и т.п.).

3. Физические ограничения на степень интеграции N элементов и размеры элементов, связанные со взаимодействием их друг с другом (джоулев разогрев).

1.1. Физические ограничения в технологии производства ИС

Элементы ИС представляют собой совокупность разнородных и неравновесных областей с резкими границами. Для их формирования в современной технологии используются способы, различающиеся:

- взаимной изоляцией приборов (разделение p - n переходами или диэлектрическими слоями);
- способом создания p - n переходов (диффузия, ионная имплантация, эпитаксия);
- способом формирования межсоединений (металлические соединения, поликристаллические линии, каналы проводимости)
- межуровневой изоляцией и т.д.

Наиболее широко в настоящее время используются различные литографические методы, легирование диффузией по окисным маскам, локальное окисление по нитридным маскам, разнообразные способы эпитаксиального наращивания слоев, ионное легирование, анодирование, металлизация. Литография, например, позволяет создавать «окна» на поверхности пластины, для последующей обработки материала в них. Процесс литографии включает в себя нанесение органического резиста, экспонирование его каким-либо облучением через фотошаблон или непосредственным «рисованием» лучом. Это приводит к полимеризации или деполимеризации фоторезиста. Далее следует удаление неэкспонированных (для негативных резистов) или экспонированных (для позитивных резистов) участков растворением, нагревом или ультрафиолетовым облучением. После литографии производится локальное травление диэлектрических или металлических пленок. Таким образом рисунок фотошаблона переносится на неорганический и нечувствительный материал, ранее нанесенный на поверхность полупроводниковой пластины.

Для создания субмикронных элементов ИС используются следующие методы: электронно-лучевая, рентгеновская и ионная литография, ионное и электронно-лучевое травление, лазерная и ионная обработки, молекулярно-лучевая эпитаксия, ионная имплантация. В последнее время все шире используются процессы самоорганизации для формирования наноструктур. Разрабатываются новые литографические методы, позволяющие получать элементы с характерными размерами вплоть до 30 нм. Это EUVL (extreme ultra violet lithography), DUV (dark ultra violet lithography), nanoimprint lithography. Прогресс в развитии литографических методов показывает, что их разрешающая способность не накладывает существенных ограничений на размеры отдельных элементов, или по крайней мере эти ограничения гораздо менее значительны, чем некоторые другие (например флуктуации концентрации легирующей примеси и её поверхностная диффузия).

Рассмотрим основные технологические ограничения.

1.1.1. Размазка края экспонированной области

Считая, что основное ограничение, связанное с фотолитографией, обусловлено явлениями дифракции света с длиной волны λ , можно записать для размазки края засвечиваемой области (Δx):

$$\Delta x \geq \lambda$$

Минимальный размер определяется длиной волны используемого света. Длина волны видимого света 0,35 – 0,7 мкм; ультрафиолетового излучения (УФ) – 0,1 – 0,3 мкм. Таким образом, использование УФ позволяет получить с помощью фотолитографии минимальный размер $\Delta x = 0,1$ мкм. Однако для достижения таких результатов необходима проекционная фотолитография и тщательно спроектированная и изготовленная УФ оптика. Получить $\Delta x < 1$ мкм с помощью контактной фотолитографии практически невозможно из-за размывания границы тени от фотошаблона.

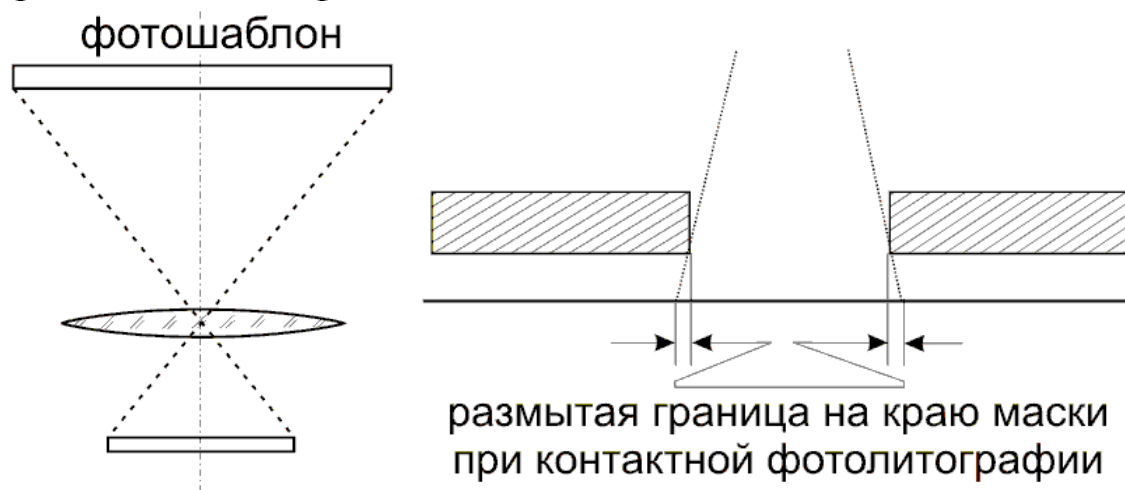


Рис. 1.2. Размазывание границы тени от фотошаблона

Использование других видов фотолитографии (рентгеновской, электронной или так называемых литографических методов нового поколения) позволяет существенно уменьшить λ , а следовательно, и минимальную ширину линии. Оценка этой величины для электронов с энергией $E = 10 - 10^3$ эВ дает значение:

$$\Delta x \geq \lambda = \frac{\hbar}{\sqrt{2mE}} \approx 1 - 0,1 \text{ \AA}$$

1.1.2. Тепловое расширение маски и ИС

В ходе процесса экспонирования фотошаблона может произойти изменение температуры как маски, так и полупроводниковой пластины. Для оценки величины теплового расширения положим коэффициент линейного расширения твердых тел $\xi \cong 2 \cdot 10^{-5} \text{ 1/K}$. Пусть размер фотошаблона $l = 150 \text{ мм}$, $\Delta T = 1 \text{ К}$ (1°C), тогда $\Delta l = l \cdot \xi \cdot \Delta T = 150 \cdot 2 \cdot 10^{-5} \cdot 1 = 3 \cdot 10^{-3} \text{ мм}$, то есть $\Delta l = 3 \text{ мкм}$. Естественно, что субмикронная литография не может допустить таких изменений размеров, вызванных колебаниями температуры. Поэтому необходима стабилизация температуры до $0,01 \text{ К}$ в пределах фотошаблона (маски) и обрабатываемого объекта.

1.1.3. Сферическая абберация электронного луча

Увеличение интенсивности электронного луча (I) приводит к усилению эффектов кулоновского взаимодействия (отталкивания) и, как следствие, росту размера пятна (сферической абберации). Уменьшение I – к увеличению времени экспонирования и, следовательно, к увеличению стоимости литографического процесса.

Радиус пятна электронного луча (R) должен быть больше его поперечной размазки вследствие эффектов абберации:

$$R > \frac{f}{4} (BS)^{3/2} I^{3/2} = kI^{3/2}, \quad (1.1)$$

где f – постоянная сферической абберации, B – яркость источника, S – площадь излучающей поверхности. Пусть N_m – некоторое минимальное число электронов, необходимое для экспонирования. Полное число электронов N_e , попавших на мишень за время экспонирования t_e , должно быть достаточно велико в силу случайного во времени процесса «доставки» электронов до мишени, т. е. $N_e > N_m$. Значение N_m определяется из условия, что при $N_e = N_m$ среднеквадратичное отклонение

(дисперсия) много меньше N_m . Зная N_m , возможно оценить минимальную интенсивность, и, подставляя ее в (1.1) получим:

$$R > R_1(t) = k \left(\frac{eN_m}{t} \right)^{3/2} \quad (1.2)$$

Увеличение времени экспонирования пятна (t) ведет к росту полного времени экспонирования (t_r) чипа площадью S_r :

$$t_r = \frac{S_r}{\pi R^2} t \quad (1.3)$$

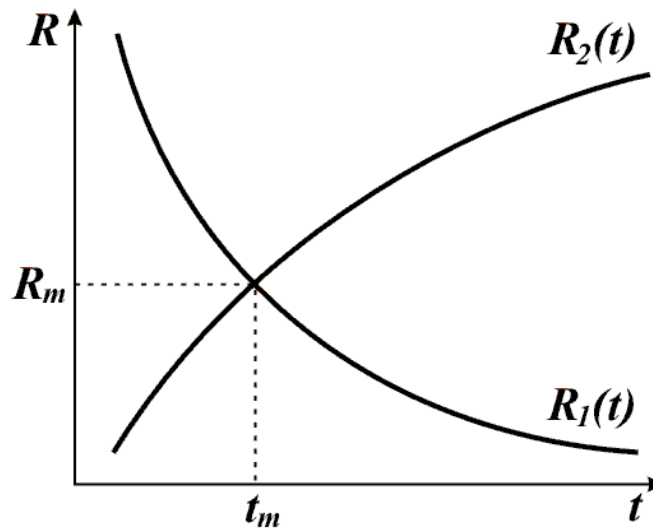


Рис. 1.3. Зависимости $R_1(t)$ и $R_2(t)$, ограничивающие размеры элементов из-за сферических aberrаций и роста стоимости процесса экспонирования соответственно. t – время экспонирования, R_m – минимальный размер элемента, t_m – время экспонирования соответствующее R_m . Область разрешенных размеров пятна: $R > R_1, R_2$

Если стоимость работы экспонирующей установки в единицу времени равна q , то стоимость процесса экспонирования $C_9 = t_r q$ не должна превышать некоторой максимальной величины C_M . Учитывая (1.3) получаем ограничение на радиус пятна:

$$R \geq R_2(t) = k \left(\frac{S_r t q}{\pi C_M} \right)^{1/2} \quad (1.4)$$

Минимальное значение радиуса пятна R_m находим из (1.2) и (1.4):

$$R_m = \left(\frac{f}{4} \right)^{1/4} k \left(\frac{S_r q e N}{\pi C_M B S} \right)^{3/8} \quad (1.5)$$

Кроме aberrаций, возникающих в оптической системе, следует помнить о требованиях, предъявляемых к жесткости конструкции и защите от вибраций.

1.1.4. Рассеяние луча в резисте и в полупроводнике

Используемые в литографии органические резисты, такие как полиметилметакрилат (ПММА) или полиметилизопренилкетон (ПМИПК), имеют минимальную толщину около 50 Å. Рассеяние луча в резисте приводит к размазке линии на величину порядка толщины резиста. Использование других типов резистов позволяет достичь еще большей разрешающей способности. В процессе травления характерным размером, определяющим размазку, является длина молекул резиста.

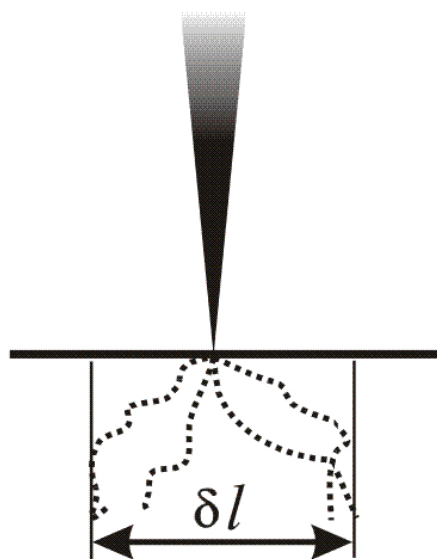


Рис. 1.4. Рассеяние электронов в резисте. δl может достигать 0,5 мкм. При $E > 100$ кэВ рассеяние увеличивается.

В процессе экспонирования электроны, попавшие в полупроводник, могут засветить резист из-за эффектов обратного рассеяния. Это приводит к дополнительной размазке края засвечиваемой области на величину порядка длины свободного пробега электронов в полупроводнике (l). Для электронов с энергией 25 кэВ длина пробега достигает 3 мкм. Соответственно,

величина размазки может быть уменьшена при использовании электронов меньших энергий. Однако электронный пучок с малой энергией $E = 5 - 10$ кэВ гораздо труднее сфокусировать.

Использование ионной литографии, а также электронной в сочетании с рентгеновской позволяет значительно снизить влияние эффектов рассеяния. Например, вторичные электроны, возбужденные рентгеновским излучением, возвращаясь в резист, дают размазку порядка 100 \AA .

1.1.5. Влияние флуктуаций примеси

Как известно, при диффузионном легировании происходит размывание краёв области легирования. Этого недостатка лишен процесс ионной имплантации (ионного легирования). В этом случае получаются резкие края зоны легирования (Рис.1.5.)

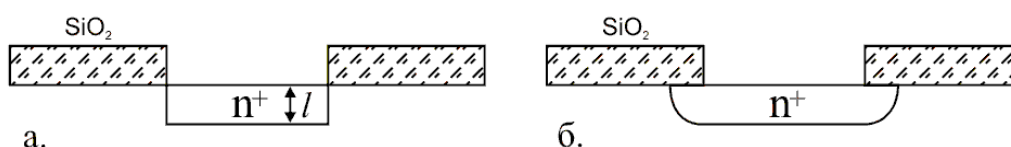


Рис. 1.5. Формирование границы легирования при ионной имплантации (а) и диффузии (б)

С помощью ионной имплантации возможно получение заглубленного легирования требуемого профиля. Для этого подбирается энергия ионного пучка. Окончательное формирование легированной области происходит при последующей диффузионной разгонке.

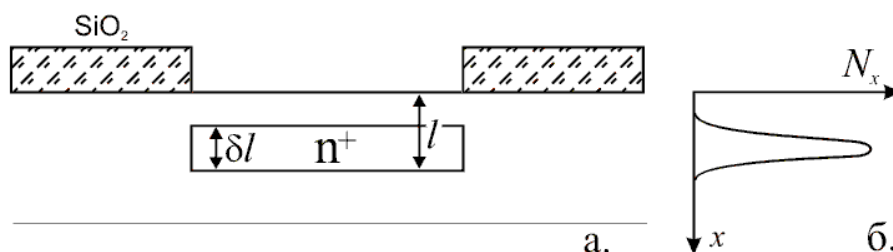


Рис 1.6. Формирование скрытого n^+ - слоя легирования толщиной δl на глубине l при имплантации высокоэнергетичными ионами (а). На рис. (б) представлен профиль легирования

Элементы ИС с p - n переходами чувствительны к уровню легирования, что приводит к ограничениям размеров для таких рабочих областей. Неизбежные в технологических процессах флуктуации примесей оказывают особенно заметное влияние при малых значениях размера активной области (d) и средней концентрации легирующей примеси (\bar{n}). Допустим, что распределение примесей по объему является гауссовым. Пусть \bar{N}_i - среднее количество примеси в активной области элемента ($\bar{N}_i = \bar{n}d^3$), ε_M - максимально допустимое относительное отклонение количества примеси от среднего значения, $\varepsilon = (N_i - \bar{N}_i) / \bar{N}_i$. Вероятность того, что в заданном объеме d^3 будет справедливо $\varepsilon > \varepsilon_M$ равна:

$$P = 1 - \sqrt{\frac{2}{\pi}} \int_0^{\varepsilon_M \bar{N}_i^{1/2}} e^{-y^2/2} dy \quad (1.6)$$

Число элементов на чипе $N = S/d^2$, умноженное на вероятность P , дает число элементов, не удовлетворяющих выбранному критерию распределения примесей (т.е. дефектных). Эта величина для чипа должна быть меньше единицы:

$$\frac{d^2}{S} < 1 - \sqrt{\frac{2}{\pi}} \int_0^{\varepsilon_M \bar{N}_i^{1/2}} e^{-y^2/2} dy \quad (1.7)$$

Подставляя в (1.7) значения $\varepsilon_M = 0,1$, $\bar{n} = 10^{19} \text{ см}^{-3}$, $S = 10^{-2} \text{ см}^2$ находим ограничение на размер элемента, а именно $d > 10^{-5} \text{ см}$.

На практике известен еще ряд существенных ограничений, проявляющихся в технологических процессах. Это, например, влияние поверхностной диффузии на размеры активных областей или накопление ошибок, появляющееся в результате многократного применения масок.

Таким образом, ограничения возникают в литографии при характерных размерах около 10 нм. Влияние легирующей примеси и поверхностной диффузии может быть более существенно. Однако эти ограничения могут быть преодолены переходом к другим типам активных элементов (напр. гетеропереходам), совершенствованием технологических процессов.

1.1.6. Статистическая воспроизводимость технологического процесса

Пусть l – размер конструктивного элемента микросхемы. При $l \gg a$, где a – атомный размер, постоянная решетки ($a \cong 3 \text{ \AA}$), дискретность кристаллической структуры не проявляется, то есть размер l можно считать макроскопическим. При этом материал рассматривается, как непрерывная среда. Если $l \cong a$, размер считается микроскопическим. Каждый атом или небольшая группа атомов должны рассматриваться как самостоятельный объект. При $l = (10 \div 100)a$ объект называется мезоскопическим. Свойства такого объекта статистически неустойчивы. Или, другими словами: мезоскопические эффекты связаны со статистической неопределенностью свойств изучаемых объектов.

Можно оценить размеры конструктивного элемента транзистора, при котором его свойства становятся статистически устойчивы, т.е. элемент можно считать макроскопическим. Пусть это будет любая область ИМС, получаемая путем легирования: база биполярного транзистора или области стока или истока МДП транзистора.

Если N – число атомов легирующей примеси, а n – их концентрация, то для объема V можно записать $N = n \cdot V$. Среднеквадратичное отклонение $\Delta N_{\text{ср.кв}}$ случайной величины N от среднего значения $N_{\text{ср}}$ определяется как

$$\Delta N_{\text{ср.кв}} = \left[\overline{(N - N_{\text{ср}})^2} \right]^{1/2} \quad (1.8)$$

Из теории вероятностей имеем:

$$\frac{\Delta N_{\text{ср.кв.}}}{N_{\text{ср.}}} = \frac{1}{\sqrt{N_{\text{ср.}}}} \quad (1.9)$$

Это справедливо для случайного процесса, подчиняющегося нормальному распределению. При $N_{\text{ср.}} \cong 10^3$, $\frac{\Delta N}{N} \cong 0,03$ (3%). Получается, что при уменьшении количества атомов примеси до

10^3 элемент становится мезоскопическим и его свойства неустойчивы. Проведем далее оценку размеров элемента, содержащего эти примеси.

Типичная концентрация примесей при тегировании базы биполярного транзистора $n = 3 \cdot 10^{16} \text{ см}^{-3}$. Тогда

$$V = \frac{N_{\text{ср.}}}{n} = \frac{10^3}{3 \cdot 10^{16}} \approx 3,3 \cdot 10^{-14} \text{ см}^3.$$

Если предположить, что элемент по форме - куб, то получаем: $V \approx L^3$ и $L \approx 3,2 \cdot 10^{-5} \text{ см} \approx 0,3 \text{ мкм}$.

Таким образом, размеры элемента (база, области стока и истока и т.д.) становятся мезоскопическими, теряют статистически устойчивые характеристики при размерах меньше $\sim 0,3 \text{ мкм}$. Если $n > 3 \cdot 10^{16} \text{ см}^{-3}$, размер элемента L может быть меньше в несколько раз.

1.2. Физические ограничения, накладываемые механизмом работы элементов ИС.

1.2.1. Классификация приборов по размерам

Классификация элементов ИС по размерам (d) может быть проведена на основе сравнения с характерными длинами, определяющими функционирование прибора. Это ширина области пространственного заряда (L), длина свободного пробега носителей (l), длина волны электрона (λ). На практике выполняется условие: $L > l > \lambda$. Исходя из этого можно определить 4 группы приборов:

- Массивные приборы с $d > L, l, \lambda$.
- Приборы промежуточной группы с $L \geq d > l, \lambda$.
- Баллистические приборы с $l > d > \lambda$.
- Квантовые приборы с $d \leq \lambda$.

Элементы ИС широкого применения относятся в основном к первой группе. Существенную долю рынка в настоящее время занимают приборы 2-й и 3-й групп. Приборы 3-й и 4-й групп и составляют элементную базу того, что в настоящее время

принято называть наноэлектроникой. И если в приборах первых двух групп проводимость описывается введением подвижности, то в последних двух явления переноса осуществляются квазибаллистически или туннелированием.

1.2.2. Полевой транзистор. Скейлинг параметров

Работа полевого транзистора.

Полевой n – канальный транзистор с изолированным затвором представляет собой структуру металл – диэлектрик – полупроводник (МДП) с металлическим затвором и полупроводниковой подложкой p – типа, на поверхности которой сформированы две юбласти n^+ – типа – исток и сток. В основе работы полевого транзистора (ПТ) лежит явление модуляции поверхностной проводимости полупроводника между истоком и стоком при изменении потенциала металлического затвора V_g . Увеличение V_g до порогового значения V_n приводит к образованию на поверхности полупроводника инверсионного слоя n – типа. Таким образом области n^+ истока и стока оказываются соединенными проводящим каналом. Значение V_n определяется из условия равенства проводимости инверсионного канала и объемной проводимости полупроводника:

$$V_n = 2\psi_B + \frac{Q_B}{C_i} = 2\psi_B + \frac{2}{C_i} \sqrt{\epsilon_s e N_a \psi_B} \quad (1.10)$$

Q_B – заряд в области обеднения полупроводника, N_a – концентрация акцепторов в полупроводнике, $e\psi_B$ – разность положений уровня Ферми для двух полупроводников – собственного и подложки, $C_i = \epsilon_i/a$ – емкость диэлектрика, a – толщина диэлектрика, ϵ_i и ϵ_s – диэлектрические проницаемости диэлектрика и полупроводника соответственно.

Введем обозначения: V_D – напряжение между стоком и истоком, V_{SS} – напряжение между истоком и подложкой, w – ширина канала, d – длина канала, μ_n – подвижность электронов в канале. Вольт-амперная характеристика (ВАХ) полевого транзистора при $V_D < V_{DS}$ ($V_{DS} = V_G - 2\psi_B$ напряжение насыщения ВАХ) определяется выражением.

$$J_D = \frac{w}{d} \mu_n C_i \left\{ \left(V_g - 2\psi_B - \frac{V_d}{2} \right) V_d - \frac{2}{C_i} \sqrt{\epsilon_s e N_a} \left[(V_d + 2\psi_B)^{3/2} - (2\psi_B)^{3/2} \right] \right\} \quad (1.11)$$

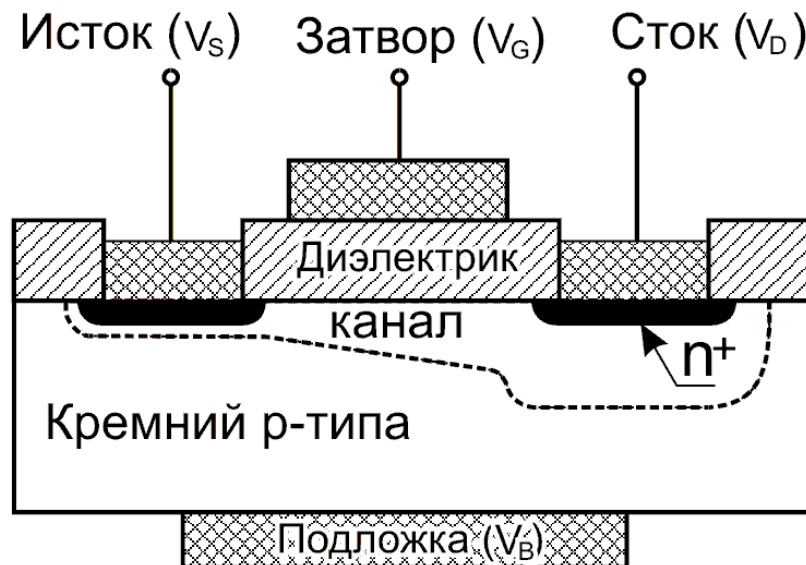


Рис. 1.7. Основные элементы структуры n-канального МОП - транзистора

При малых напряжениях V_D сопротивление канала постоянное и зависимость J_D от V_D линейная. С увеличением V_D уменьшается глубина канала вблизи стока, сопротивление канала растет и происходит насыщение J_D .

Минимальная ширина ОПЗ для КМП.

Проведем оценку минимальной ширины области пространственного заряда (ОПЗ) для контакта металл-полупроводник (КМП) [5]. Начнем с оценки минимального значения ширины ОПЗ L в контакте, для которого размерные эффекты, связанные с d , не проявляются. Используем изменение ВАХ с уменьшением L . Очевидно, что при уменьшении L возможно изменение вида и параметров ВАХ, определяемое изменением физических процессов, которые определяют ток.

Если исходить из критериев диффузионной теории, то минимальная ширина L не может быть меньше, чем

$$L \gg l = \frac{3\pi u_n m_n^* v_n}{8e} \quad (1.12)$$

u_n - подвижность электронов в диэлектрическом слое. Например, для контакта на n-Si при $L = 5l$ ширина L должна быть больше 0,043 мкм. При меньших L выполняется диодная теория до значений, при которых станет существенным туннелирование носителей через ОПЗ. Допустим, что туннелирование отсутствует. Для того чтобы определить ход потенциала $\varphi(x)$ и значения параметров ОПЗ в полупроводнике для слоя Шотки, когда высота потенциального барьера φ_0 изменяется в пределах

$$2,3kT < \varphi_0 < E_g - 2\mu + \frac{3}{2}kT \ln\left(\frac{m_n^*}{m_p^*}\right),$$

необходимо решить уравнение Пуассона для равномерного распределения примесей в полупроводнике. Результатом интегрирования будут выражения

$$\varphi(x) = \frac{e^2 n_0}{2\varepsilon\varepsilon_0} (L-x)^2, \quad E(x) = -\frac{en_0}{\varepsilon\varepsilon_0} (L-x)$$

Здесь: φ_0 - высота потенциального барьера, L - ширина потенциального барьера (ОПЗ), E - напряженность поля, E_g - ширина запрещенной зоны полупроводника, μ - энергия равная разнице между положением уровня Ферми и дном зоны проводимости, m_n^* и m_p^* - эффективные массы электронов и дырок соответственно. При $x = 0$

$$\varphi_0 = \frac{e^2 n_0}{2\varepsilon\varepsilon_0} L_0^2, \quad E_0 = -\frac{en_0}{\varepsilon\varepsilon_0} L_0, \quad L_0 = \sqrt{\frac{2\varepsilon_0\varepsilon\varphi_0}{e^2 n_0}} \quad (1.13)$$

Используя для L приведенные оценки, неравенство (1.12) можно переписать в виде [5]

$$L \gg \frac{\sqrt{\varphi_0 - eU_2}}{kT} \sqrt{\frac{\hbar^2}{2m^*}} \quad (1.14)$$

Значение L можно оценить, если задать $\frac{\sqrt{\varphi_0 - eU_2}}{kT} \cdot \sqrt{\frac{\hbar^2}{2m^*}} \cdot \frac{1}{L} = \delta$.

Например, для КМП на n-Si, если $\varphi_0 - eU_2 = 0,5$ эВ и $\delta = 0,2$, то L

$= 0,05$ мкм. Таким образом, пренебрежение туннелированием приводит к достаточно малым предельным значениям L . Однако следует отметить, что уменьшение L , определяемое из условия пренебрежения туннелированием, до еще меньших размеров не является принципиально невозможным. Действительно, выпрямительные свойства контакта при этом сохраняются [5]. Но существенно изменяется вид ВАХ и может поменяться направление выпрямления для прямого и обратного направлений тока в контакте. Поэтому значения L , полученные из анализа ВАХ, могут быть значительно меньшими, чем заданные выражениями (1.12) и (1.14). При малых L более существенными становятся ограничения, связанные с дискретностью заряда. Так, при решении уравнения Пуассона и получении $\varphi(x)$ используется предположение о непрерывности заряда в ОПЗ. В то же время заряд, связанный с ионизированными примесями, является дискретным. В связи с этим введение L , строго говоря, возможно при выполнении неравенства

$$L \geq (n_{i\delta})^{1/3}, \quad (1.15)$$

где n_{np} - концентрация ионизированных примесей. Если $n_{np} = n_0$, то можно воспользоваться выражением для L (1.13) и переписать последнее неравенство в виде

$$L \geq \frac{l^2}{2\epsilon\epsilon_0(\varphi_0 - eU_2)} \quad (1.16)$$

Если задать значение $\frac{l^2}{2\epsilon\epsilon_0(\varphi_0 - eU_2)L} = \delta$, то можно оценить минимальное значение L . Например, для контакта на n -Si, с $\varphi_0 - eU_2 = 0,5$ эВ при $\delta = 0,2$ $L = 0,2$ мкм.

Рассмотренное ограничение ширины L , связанное с дискретностью заряда в ОПЗ, определяет минимальное значение L как для ВАХ, так и для ВФХ, а также напряженность электрического поля E_2 в ОПЗ. Отметим, что ВАХ и ВФХ не претерпевают принципиальных изменений при изменении L в пределах от $L < l$ до $L > l$, но параметры некоторых приборов с КМП - структурами могут существенно изменяться. Так, лавинно

- пролетные приборы работают лишь в том случае, когда длина свободного пробега электронов $l \ll L$, т. е. при условиях выполнения диффузионной теории (иначе не будет лавинного размножения носителем заряда). Рассмотренные здесь ограничения ширины ОПЗ L определяют минимальный размер КМП в направлении, перпендикулярном его плоскости.

Минимальная толщина пленки металла КМП

Ограничения минимальной толщины пленки металла определяются следующими физическими причинами. Во-первых, минимальная толщина металла определяется требованиями к высоте барьера ϕ_0 . Так как барьер формируется уже при толщинах металла порядка нескольких монослоев, то с этой точки зрения толщина металла может быть $l_m \ll 0,001$ мкм. Во-вторых, минимальная толщина металла определяется из условия достаточно малого сопротивления пленки R_m . Поэтому толщина металла $l_m \geq \frac{R_m}{\rho_m}$ (ρ_m - удельное сопротивление металла). Если $\rho_m = 10^{-5} - 10^{-6} \text{ Ом}^{-1} \cdot \text{см}^{-1}$, то при $R_m = 1$ Ом толщина пленки металла $l_m \leq 0,01 - 0,1$ мкм. Значение l_m в среднем меньше, чем размеры L и d . Таким образом, толщина пленки металла, как и туннельно-прозрачного диэлектрического переходного слоя ($d_n < 0,006$ мкм), не является определяющим при рассмотрении минимальных вертикальных размеров структур КМП.

Изменение параметров приборов при уменьшении их размеров – модель скейлинга ПТ.

Поскольку уменьшение размеров приборов приводит к изменению их рабочих характеристик, то каждый такой шаг требует новых процессов оптимизации рабочих параметров микроэлектронных устройств. Существует, однако, диапазон геометрических размеров, где можно сравнительно просто оценить изменение параметров приборов при уменьшении их геометрических размеров в K раз. Для оценки преимуществ, которые дает миниатюризация, используются простые масштабные соотношения, называемые скейлингом (scale – масштаб).

Таблица 1.3

Скейлинг элементов и межсоединений

Параметры	Расчетные формулы	F(K)
Геометрические размеры d, w, L, a		K^{-1}
Напряжения V_g, V_D	$E = const, V_D = E_D d$	K^{-1}
Концентрация примеси в подложке N_A	$E_s \propto \sqrt{(V_g - E_1 a) N_A} \approx const$	K
Ток в линейной области ВАХ J_D	$J_D = \frac{w}{d} \frac{\varepsilon_s}{a} V_g V_D$	K^{-1}
Время задержки переключения τ	$\tau = \max \left(\frac{d^2}{\mu V_D}; \frac{C_g V_g}{J_D} \right)$	K^{-1}
Джоулева мощность для управления переключением	$P = \frac{1}{2} C_g \Delta V_g^2$	K^{-3}
Функциональное быстродействие ИС, ФБ	$\Phi Б = \overline{N} \tau^{-1}$	K^3
Сопротивление линий межсоединений (ЛМС)	$R_l = \rho_l \frac{l}{w_l h_l}$	K
Время отклика ЛМС, τ_l	$\tau_l = \begin{cases} R_l C_l \\ l/v_0 \end{cases}$ C_l - емкость соединения, v_0 - скорость электромагнитной волны	1
Плотность тока в ЛМС, j_l	$j_l = \frac{J_D}{w_l h_l}$	K
Нормализованное время отклика ЛМС	$\frac{\tau_l}{\tau} \approx \frac{R_l C_l}{\tau}$	K

Таким образом, задачи, решаемые этим способом, состоят в определении некоторых масштабных множителей $F(K)$, с помощью которых параметры приборов, уменьшенных в K раз, могут быть выражены через параметры исходного прибора. Используются различные модели скейлинга, в зависимости от объекта, к которому они применяются или от характера

величины, требование постоянства которых закладывается в теорию.

Пусть, например, геометрические параметры транзистора (w – ширина канала, d – длина канала, a – толщина диэлектрика) уменьшаются в K раз. Напряженность электрического поля E_s в области пространственного заряда (ОПЗ) полупроводника должна оставаться постоянной:

$$E_s \sim \sqrt{(V_g - E_i a) N_a} \quad (1.17)$$

Здесь E_i – напряженность поля в диэлектрике. Таким образом, для неизменности поля в рабочих областях, необходимо концентрацию примесей увеличить в K раз, а все напряжения уменьшить в K раз. Из (1.11) следует, что при $V_g \gg V_n$ выражение для тока:

$$J_D = \frac{w}{d} \mu_n \frac{\varepsilon_i}{a} V_g V_D \quad (1.18)$$

Отсюда следует, что значение J_D обратно пропорционально K . В Таблице 1.3 приведены зависимости некоторых параметров МДП транзисторов и межсоединений при уменьшении их размеров в K раз.

Причины отказа скейлинга

Использование метода масштабирования на первый взгляд позволяет сделать вывод, что размеры элементов можно уменьшать без каких-либо ограничений. Необходимо только вносить соответствующие масштабированию коэффициенты в соответствующие характеристики рабочих областей. Однако практика показывает, что, начиная с некоторых размеров (для кремниевой технологии – это субмикронные масштабы) закономерности скейлинга не работают. Например для SiO_2 – традиционного материала изолирующих слоев минимальная толщина оценивается в 2,3 нм. Этот предел определяется процессами утечки заряда через диэлектрик за счет процессов туннелирования. Уменьшение глубины легирования контактов в областях стока и истока менее 30 нм приводит к росту контактного сопротивления и деградации рабочих характеристик транзистора. В таблице 1.4 приведены значения параметров,

ограничивающие применение метода масштабирования для МДП транзисторов. [6]

Таблица 1.4

Параметр	Прелел	Причина
Толщина окисла	2,3 нм	Утечка
Толщина контактной области	30 нм	Сопротивление
Легирование канала	$V_T=25$ В	Утечка
SDE диффузионный размер	15 нм	Сопротивление
Длина канала	0,06 мкм	Утечка
Длина затвора	0,1 мкм	Утечка

В таблице: SDE – source/drain extension - уширение областей истока и стока легированием этих областей по краю на глубину 30 - 40 нм для усиления эффектов короткого канала.

По мере уменьшения длины канала толщина диэлектрика уменьшалась от 10 мкм до 0,1 мкм. Толщина подзатворного окисла (T_{OX}) линейно зависит от длины канала (L_E). На основе анализа технологии за последние 20-25 лет получено выражение, связывающее эти параметры: $L_E = 45 \times T_{OX}$.

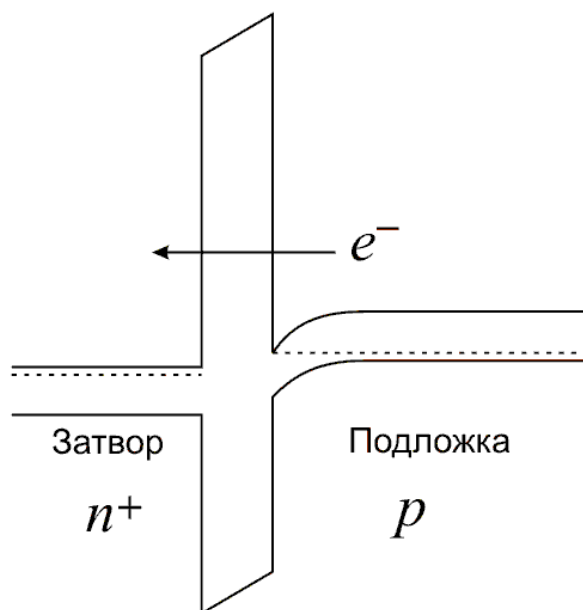


Рис 1.8. Механизм утечки носителей вследствие прямого туннелирования через тонкий диэлектрик.

Современная технология позволяет изготавливать слои оксидов и оксинитридов кремния, применяемых в качестве

подзатворных диэлектриков, толщиной менее 1 нм. Ограничения по толщине никак не ограничены технологией. Основная причина – это утечки вызванные туннелированием носителей через тонкий диэлектрик. На рисунке 1.8 показан механизм туннелирования носителей через потенциальный барьер для n -МДП транзистора в режиме инверсии. Как известно, существует разница в высоте потенциального барьера для электронов и для дырок. Кроме того, вероятность туннелирования в окисел для дырок значительно меньше, чем для электронов. Таким образом, предельное значение толщины подзатворного диэлектрика для n -МДП транзисторов больше, чем для p -МДП.

По поводу возможных причин, приводящих к невыполнению закономерностей масштабирования, авторами [1] приведены следующие соображения:

- Для образования инверсионного канала необходимо выполнение соотношения $V_g > V_n$, т. е. напряжение на затворе V_g больше порогового V_n , которое практически от K не зависит. Тогда начиная с некоторого значения K величина V_g должна оставаться постоянной. Выполнение этого условия невозможно, поскольку напряженность электрического поля E (а именно она должна оставаться постоянной при изменении характерного размера d) и V_g связаны соотношением $V_g = E_g d$
- Тепловые флуктуации тока должны быть значительно меньше собственно тока. Это условие будет выполняться в случае, когда напряжение между стоком и истоком будет существенно больше kT/e ($V_D > 10 kT/e$).
- Длина канала должна быть существенно больше толщины p - n перехода ($d > 2L$). Это условие необходимо для того, чтобы не было смыкания областей пространственных зарядов между областями переходов исток-подложка и сток-подложка.
- Исток, подложка и сток ПТ фактически образуют n - p - n транзистор. Дырки, генерируемые в области обеднения стока, вызывают ток через подложку, являясь для материала подложки основными носителями. В результате увеличивается потенциал подложки вблизи истока, и как следствие – увеличивается инжекция носителей из n - p перехода истока в подложку. Очевидно, что с уменьшением длины канала все эти эффекты начинают проявляться при меньших напряжениях.

Для МДП транзистора ток через подложку является паразитным, а увеличение его с уменьшением длины канала, приводит к аномальному уменьшению напряжения пробоя.

- Если ширина области обеднения истока и стока становится сравнимой с длиной канала (d), происходит уменьшение порогового напряжения V_n (эффект короткого канала). Это связано с тем, что с уменьшением длины канала уменьшается величина заряда Q_B , находящегося в слое обеднения полупроводника. Таким образом, согласно (1.10) V_n также уменьшается.
- Уменьшение ширины канала w , начиная с некоторого значения, приводит к увеличению порогового напряжения V_n (эффект узкого канала). Это связано с эффективным увеличением заряда в слое обеднения полупроводника.
- Уменьшение длины канала МДП транзистора приводит к увеличению предпорогового тока и его зависимости от напряжения между стоком и истоком. Этот эффект может быть существенно подавлен, если производить оптимизацию других. В частности, минимальная длина канала, начиная с которой резко увеличивается предпороговый ток, может быть определена из следующего эмпирического соотношения:

$$d_m = c(X_j a w_1^2)^{1/3} \quad (1.19)$$

где c – постоянная ($c = 0,41 \text{ \AA}^{-1/3}$), a – толщина диэлектрика, w_1 – суммарная ширина областей обеднения стока и истока. Из (1.19) видно, что в отличие от скейлингового подхода, когда $d_m \sim K^{-1}$, оптимизация параметров дает $d_m \sim K^{-4/3}$.

1.2.3. Ограничения электрофизических параметров

При приложении фиксированного напряжения минимальный характерный размер элемента (толщина подзатворного диэлектрика, толщина обедненного слоя, толщина базы и т.п.) определяется пробоем материала. Пробивная напряженность поля для качественного SiO_2 составляет 10^6 В/см. Для того, чтобы снизить вероятность пробоя, необходимо снижать напряжение, прикладываемое к элементу ИС. Дрейф порогового напряжения транзистора может быть, например, уменьшен за счет отладки и совершенствования технологического процесса.

Неустраняемыми являются тепловые флуктуации. Рассмотрим несколько возможных вариантов ограничений электрофизических параметров ИС [7].

Минимальное напряжение питания полупроводниковой ИС.

Пусть E энергия, которую носители заряда должны преодолеть, чтобы переключить цепь (открыть транзистор, переключить состояние триггера и т.п.). Положим, что $E = eU$, где U – напряжение источника питания схемы. Носители заряда находятся в тепловом движении.

Вероятность того, что они наберут в процессе теплового движения энергию E_n , определяется распределением Гиббса :

$$W_n = A_n e^{-E_n/kT} \quad (1.20)$$

Положим, приближенно, что состояние с $E_n = 0$ осуществляется с вероятностью, равной единице. Поэтому $A_n = 1$. Таким образом, вероятность того, что цепь переключится в результате теплового движения:

$$W = e^{-U/\varphi_T}, \quad \varphi_T = \frac{kT}{e},$$

откуда для $T = 300$ К получаем, что $\varphi_T = 0.026$ В.

Пусть в ИС 10^6 элементов, каждый из них может переходить в новые состояния 10^9 раз в секунду. Потребуем теперь, чтобы в течение 10^3 секунд (20 мин) в среднем только в одном элементе при одном переключении может произойти сбой. Тогда :

$$W = \frac{1}{10^6} \cdot \frac{1}{10^9} \cdot \frac{1}{10^3} = 10^{-18}$$

Возвращаясь к (1.20), получим $e^{-U/\varphi_T} = W = 10^{-18}$. Отсюда вытекает требование к напряжению питания: $U = \varphi_T \cdot 2,3 \cdot 18 \cong 41\varphi_T$. При $T = 300$ К, $\varphi_T = 0.026$ В, получаем, что $U_{\text{пит}} \approx 1$ В. Понижение рабочей температуры дает возможность понизить $U_{\text{пит}}$: $U_{\text{пит}} = 0.25$ В для $T=78$ К и $U_{\text{пит}} = 0.013$ В для $T=4$ К.

Учитывая, что напряжённость поля пробоя для хорошего диэлектрика $E_{\text{проб}} = 10^7 - 10^8$ В/м, толщина слоя подзатворного диэлектрика может быть оценена следующим образом:

$$t_{\min} = \frac{U}{E_{\text{проб}}} = \frac{1B}{(10 \div 100) \frac{B}{\text{мкм}}} = 0,1 \div 0,01 \text{ мкм}$$

Средняя толщина диэлектрика, устойчивого к пробоем при напряжениях $U=1$ В составляет 30 – 300 нм (т.е. доли микрометра). При толщине диэлектрика менее 3 нм утечка заряда осуществляется не за счет пробоя, а за счет туннельного эффекта.

Минимально допустимый ток в ИС.

Небольшие беспорядочные отклонения тока как в электровакуумных, так и в полупроводниковых приборах, вызванные неравномерностью инжекции носителей, называются дробовым эффектом. Он сопровождает любые процессы, связанные с протеканием электрического тока через полупроводник. Поэтому минимально допустимый ток в ИС также определяется флуктуациями тока, т.е. дробовым эффектом.

$$I = \frac{n \cdot e}{\tau},$$

где n – число электронов, а τ - длительность импульса тока.

Среднеквадратичное отклонение числа электронов в импульсе от их среднего числа (см. 1.8) :

$$\Delta n_{\text{ср.кв.}} = \left[\overline{(n - n_{\text{ср.кв.}})^2} \right]^{1/2}$$

Учитывая, что для статистически независимого числа частиц :

$$\frac{\Delta n_{\text{ср.кв.}}}{n_{\text{ср}}} = \frac{1}{\sqrt{n_{\text{ср}}}}$$

Получаем, что $\frac{\Delta I}{I} = \frac{\Delta n}{n} \frac{1}{\sqrt{n_{\text{ср}}}}$

Для того, чтобы $\Delta I/I$ не превышало 1 %, нужно иметь $n_{\text{ср}}=10^4$. Пусть $\tau = 2 \cdot 10^{-10}$ с $f_{\text{такт}} = 1000$ МГц (см. рис 1.9). Тогда

$$I_{\text{имп}} = \frac{n_{\text{ср}} \cdot e}{\tau} = \frac{10^4 \cdot 1,6 \cdot 10^{-19}}{2 \cdot 10^{-10}} = 8 \cdot 10^{-6} \text{ А} \cong 10^{-5} \text{ А}$$

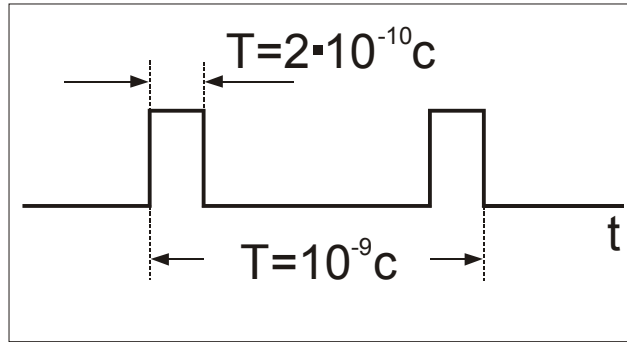


Рис. 1.9. Период и длительность импульсов

Если найденное значение - это ток в цепи базы, то в цепи коллектора может быть в 100 раз больше, т. е. $I_{\text{имп}} = 10^{-3} \text{ A} = 1 \text{ mA}$. Поэтому обычная величина тока импульсов в ИС выражается долями или единицами мА.

Минимальное значение тока, достаточного для перезарядки емкостей (в том числе и паразитных емкостей в схеме) определяется следующим образом:

$$\text{Для } C = 1 \text{ пФ, } U = 1 \text{ В, } \tau = 10^{-10} \text{ с, } I = \frac{CU}{\tau} = \frac{10^{-12}}{10^{-10}} = 0,01 \text{ A} \quad (10 \text{ mA})$$

Допустимые плотности токов в шинах ИС.

Пусть $S = w \times t$ – площадь поперечного сечения шины.



Рис 1.10. Параметры токопроводящей шины

Принятые стандарты размеров : $w = 10; 5; 2.5; 1.25; 0.5 \text{ мкм}$, $t = 1.0; 0.5; 0.2 \text{ мкм}$, $S = 10 \dots 0.1 \text{ мкм}^2$

Пусть ток 0.1 – 1 мА. Тогда плотность тока получается равной:

$$j = \frac{I}{S} = \frac{0,1 \cdot 10^{-3}}{10 \cdot 10^{-12}} = 10^7 \left(\frac{\text{A}}{\text{м}^2} \right) = 10 \left(\frac{\text{A}}{\text{мм}^2} \right)$$

или в самом тяжёлом случае:

$$j = \frac{I}{S} = \frac{1 \cdot 10^{-3}}{0,1 \cdot 10^{-12}} = 10^{10} \left(\frac{A}{m^2} \right) = 10 \left(\frac{\kappa A}{mm^2} \right)$$

Обычно в Al шинах на Si подложке не используется ток плотностью больше $100 \text{ A} \cdot \text{мм}^{-2}$. Напряжённость электрического поля вдоль шины $E = \frac{1}{\sigma} j$. Для Al : $\sigma = 3 \cdot 10^7 \text{ 1/Ом} \cdot \text{м}$. При $j = 10^8 \text{ A/м}^2$ получаем $E = 3 \text{ В/м}$. В масштабах ИС это приведет к ничтожно малому падению напряжения. Плотность мощности на единицу поверхности шины толщиной t :

$$P_{нов} = \frac{j^2}{\sigma} t = \frac{10^{16}}{3 \cdot 10^7} \cdot 10^{-6} = 300 \left(\frac{Bm}{m^2} \right) = 0,3 \left(\frac{Bm}{cm^2} \right)$$

Такой поток мощности легко отводится за счет теплопроводности кремния.

Электромиграция

При плотности тока более 100 A/мм^2 возможна миграция атомов металла в сторону одного из электродов.

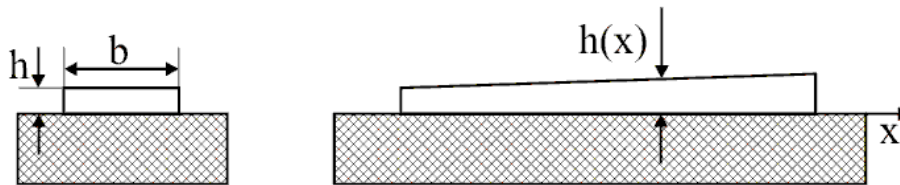


Рис. 1.11. Изменение размеров (толщины) металлической шины при электромиграции

В результате наблюдается изменение размеров металлической шины на изолированной подложке за счет переноса атомов металла.

В процессе теплового движения ионы металла могут занимать нерегулярные положения в кристаллической решетке. В процессе теплового движения происходит движение ионов по междоузлиям, генерация и заполнение вакансий. Это т.н. процесс самодиффузии ионов. При наличии дрейфа электронов, они «подталкивают» ионы. Происходит направленное движение

ионов. Захват ионов дрейфом электронов называют “электронным ветром”.

Направленное движение ионов металла под действием потока электронов называют электромиграцией.

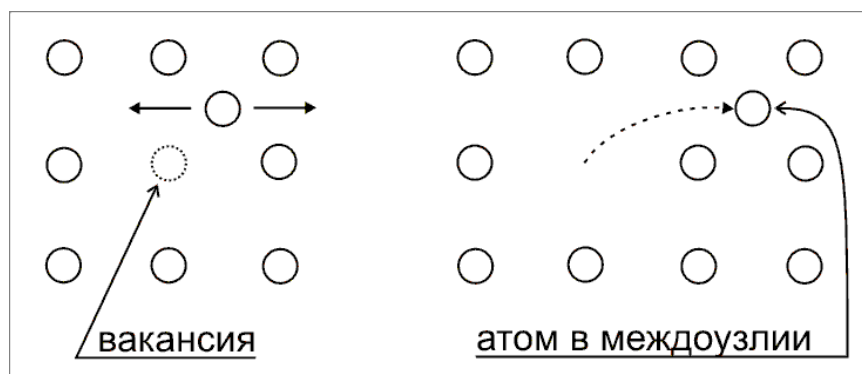


Рис 1.12. Механизмы самодиффузии

Миграция происходит в сторону положительного электрода, а движение ионов уравнивается тепловой диффузией.

$$D \cdot \text{grad } n = \mu \cdot n \cdot E,$$

где, μ - подвижность ионов, увлекаемых электронным ветром, D - коэффициент диффузии, E - напряженность электрического поля.

По формуле Эйнштейна $D = \mu k \frac{T}{e}$. Отсюда для одномерной диффузии имеем :

$$\frac{\mu k T}{e} \cdot \frac{dn}{dx} = \mu n E$$

или:

$$\frac{dn}{n} = \frac{E}{\varphi_T} \cdot dx = \frac{dx}{l_{\text{мigr}}},$$

где использовано обозначение $l_{\text{мigr}} = \frac{\varphi_T}{E}$

Решение записанного уравнения имеет вид

$$n(x) = n(0) \cdot e^{x/l_{\text{мigr}}},$$

где $n(x)$ – число ионов металла на единицу длины токоведущей шины, $l_{\text{мigr}}$ – характерная длина миграции. Необходимо, чтобы размеры $l_{\text{мigr}}$ были много больше длины проводников ИС. Пусть $l_{\text{мigr}} \sim 1$ см – тогда миграция будет мало сказываться.

$$l_{\text{мigr}} = \frac{\varphi_T}{E} = \frac{\varphi_T \sigma}{j}, \text{ откуда } j \leq \frac{\varphi_T \sigma}{l_{\text{мigr}}} = \frac{0,026 \cdot 7 \cdot 10^7}{10^{-2}} \cong 10^8 \left(\frac{A}{\text{м}^2} \right)$$

Процессы миграции существенно усиливаются в случае неравномерной толщины шины.

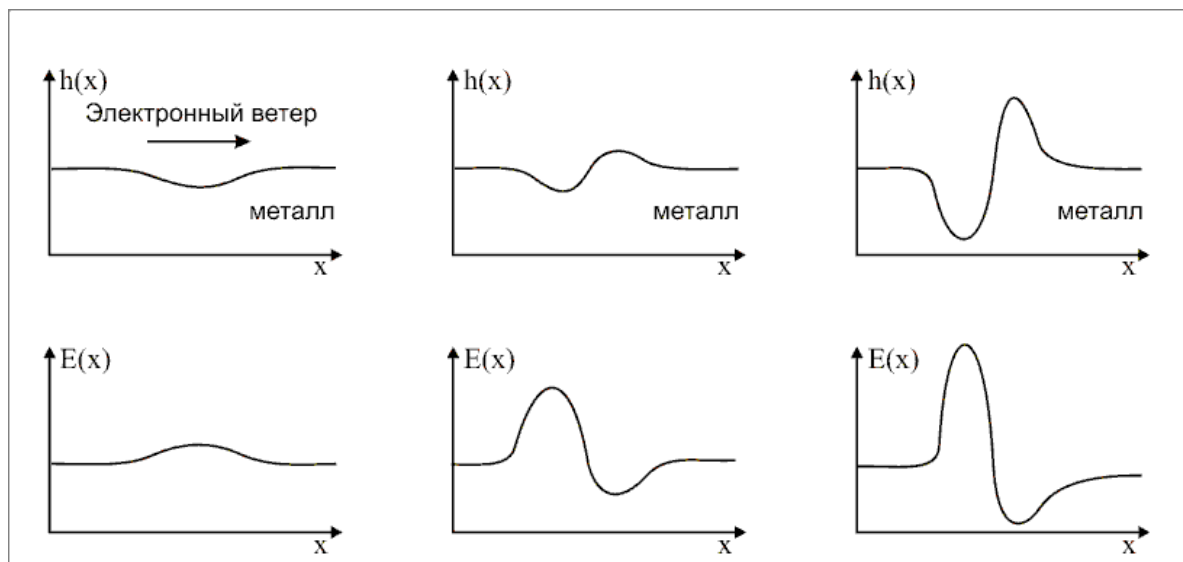


Рис. 1.13. Изменение толщины (h) металлической шины и соответствующая напряженность (E) электрического поля

Здесь показано изменение профиля металлической шины в результате миграции атомов. Параметр $h(x)$ – толщина токоведущей шины. Под действием «электронного ветра» тонкая часть шины утоньшается вплоть до разрушения. Существуют эмпирические формулы, дающие оценку среднего времени жизни шины под током до обрыва (τ)

Для Al на Si :

$$\frac{1}{\tau} = \frac{1}{\tau_0} \cdot \left(\frac{j}{j_0} \right)^n e^{-E_a/kT},$$

E_a – энергия активации образования вакансии в кристаллической решетке металла. Для Al : $E_a = 0.5$ эВ и далее $e^{\frac{-0,5}{0,026}} \approx e^{-20} \cong 10^{-8}$

При $j = j_0$ из приведённой эмпирической формулы получаем $\tau = \tau_0 \cdot 10^8$. Здесь j_0 , τ_0 – характерная плотность тока и характерное время, используемые как параметры модели. Они определяются балансом процессов миграции и диффузии. Для Al

ориентировочная величина характерных параметров: $j_0 = 10^2$ А/мм², $\tau_0 = 10$ час. Значение показателя n изменяется от 2 до 6 в зависимости от качества технологии получения пленки (наличие трещин, пор, неоднородности по толщине и ширине шины).

1.2.4. Задержка и искажение импульсов на связях

Пусть между двумя блоками ИС существует соединительная шина, длина которой может достигать 1 см.

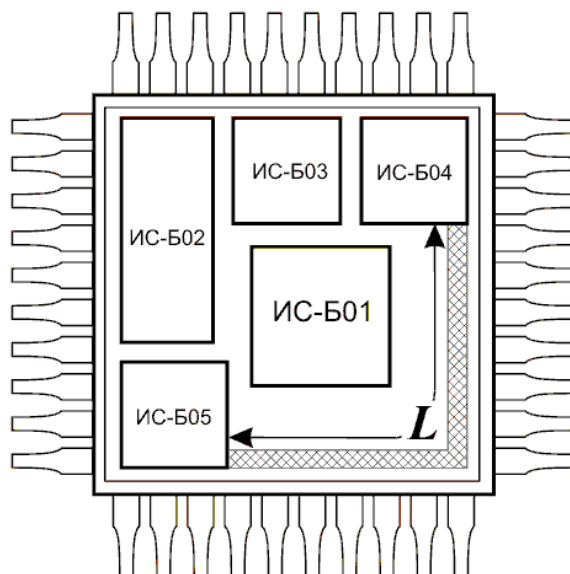


Рис. 1.14. Соединительная шина между двумя блоками ИС.

Важнейшей характеристикой связи является время, в течение которого импульс (или фронт импульса) пробегает вдоль связи. Эта длительность определяется длиной связи и скоростью распространения электромагнитной волны вдоль линии передачи, образованной связью.

$$\tau = \frac{L}{c/\sqrt{\epsilon_{\text{эфф}}}},$$

c — скорость света, $\epsilon_{\text{эфф}}$ — эффективная диэлектрическая проницаемость материала, окружающего линию связи, L — длина связи. Рассмотрим две модели для соединительной шины в составе ИС [7].

Импульс, несущий информацию, длиннее времени пробега по связи.

Это условие может быть записано как: $\tau_{\text{имп}} > \tau$, где $\tau_{\text{имп}}$ - длительность импульса. В этом случае шина может рассматриваться как RC цепь, а задержка сигнала определяется временем заряда ёмкости C через сопротивление R . Очевидно, что это время пропорционально квадрату длины связи, так как и ёмкость связи и её сопротивление пропорциональны длине.

Импульс, несущий информацию, короче времени пробега по связи.

Это условие может быть записано как: $\tau_{\text{имп}} < \tau$, где $\tau_{\text{имп}}$ - длительность импульса. В этом случае шину следует рассматривать как RLC линию передачи, и тогда практически все вопросы задержки и искажения импульсных сигналов смыкаются с вопросами передачи сигналов по линиям связи, которые решаются техникой сверхвысоких частот (СВЧ).

Задержка цифрового сигнала на связях (RC – МОДЕЛЬ)

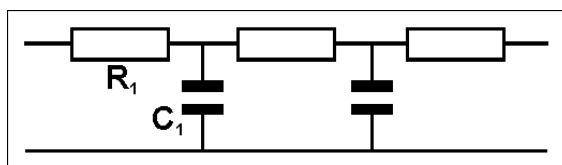
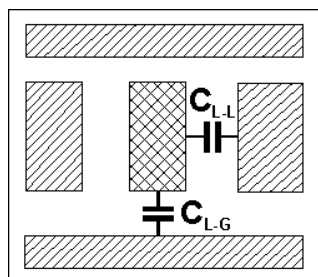


Рис. 1.15. Эквивалентная схема для RC модели

Здесь R_1 – погонное сопротивление шины, C_1 – погонная ёмкость шины.

$$R_1 = \frac{1}{\sigma} \cdot \frac{1}{w \cdot h},$$

где σ - проводимость материала шины (для меди при $T = 300$ К $\sigma = 6 \cdot 10^7$ 1/Ом·м), w , h – ширина и толщина шины. Чем меньше размеры транзисторов, тем меньше приходится делать и размеры



поперечного сечения шин, обеспечивающих связи между транзисторами. Положим, что приближённо $w \cdot h \cong l^2$. Здесь l – характерная длина (длина канала для МДП транзистора). В случае многослойного расположения связей погонная ёмкость определяется как ёмкостью

шины на разделяющий экран (C_{L-G} , ground plane), так и на элементы конструкции, расположенные рядом с шиной (C_{L-L}).

Оценка параметров задержки цифрового сигнала (импульсов) на RC связях представлена в [8].

Расчёт погонной ёмкости для диэлектрической проницаемости материала изолирующих слоёв $\epsilon_d \cong 15$ приводит к зависимости погонной ёмкости от характерного размера элемента интегральной схемы, который представлен на рис. 1.16. Можно ориентировочно считать, что для интегральных схем с достаточно высоким уровнем интеграции ($l = 0,25 - 1,0$ мкм) $C_1^0 = 2,5 \cdot 10^{-16}$ Ф/мкм.

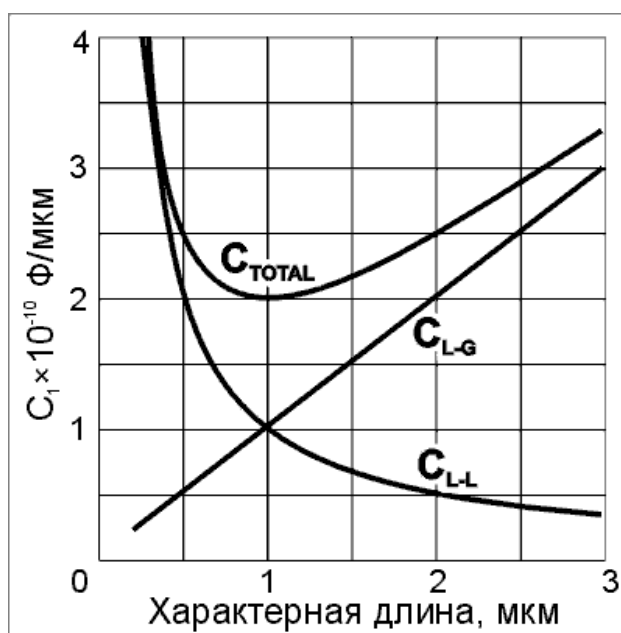


Рис. 1.16. Зависимость погонной емкости от размера характерного элемента ИМС

Теперь можем оценить величину постоянной времени перезарядки шины длиной L .

$$\tau_{св} = RC = R_1 C_1 \cdot L^2,$$

где R и C - полное сопротивление и полная ёмкость шины, R_1 и C_1 - её погонные параметры. Собирая вместе сделанные выше оценки, получим:

$$\tau_{св} = C_1^0 \frac{1}{\sigma} \cdot \frac{1}{l^2} \cdot L^2 \quad (1.21)$$

Для сравнения приведём оценку задержки импульсного сигнала в транзисторе:

$$\tau_{тр.} = \frac{l^2}{\mu U^2}, \quad (1.22)$$

где μ - подвижность, U – рабочее напряжение транзистора.

На рисунке 1.17 приведены зависимости времени задержки в транзисторе (кривая 1) и на связях (кривые 2 и 3) в функции от характерного размера для $\sigma = 6 \cdot 10^7$ 1/Ом·м, $\mu = 0,004$ м²/В·с, $U=1$ В, $L = 5$ мм (кривая 2) и $L = 3$ мм (кривая 3).

Приведённые оценки показывают, что при $L = 3 - 5$ мм для транзисторов с длиной канала меньше 0,5 мкм задержка на связях оказывается большей, чем задержка на одном транзисторе. Заметим, что логические схемы содержат больше одного транзистора на одну логическую операцию, поэтому задержка на одну логическую операцию оказывается больше, чем на один транзистор. Поэтому стремление ещё больше уменьшить задержку на один транзистор оправдано, то есть оправдано сокращение длины канала, и оправдана разработка транзисторов с длиной канала менее четверти микрометра.

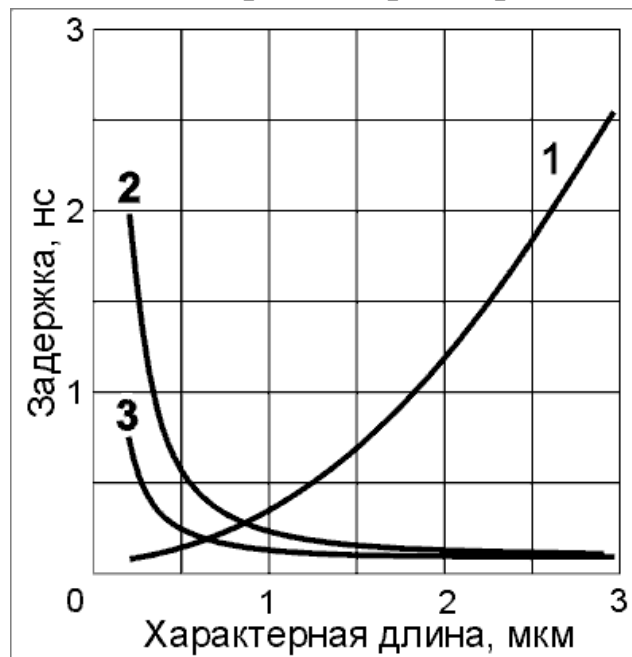


Рис. 1.17. Зависимости времени задержки в транзисторе (кривая 1) и на связях (кривые 2 и 3)

Суммарная задержка определяет допустимую тактовую частоту, с которой производятся вычислительные операции в ИС. Для ориентировочной оценки можно использовать простое

соотношение между тактовой частотой и суммарной задержкой сигнала:

$$f = \frac{1}{\tau_{\text{сумм.}}} \quad (1.23)$$

На рисунке 1.18 приведена зависимость тактовой частоты от характерного размера активного элемента ИС.

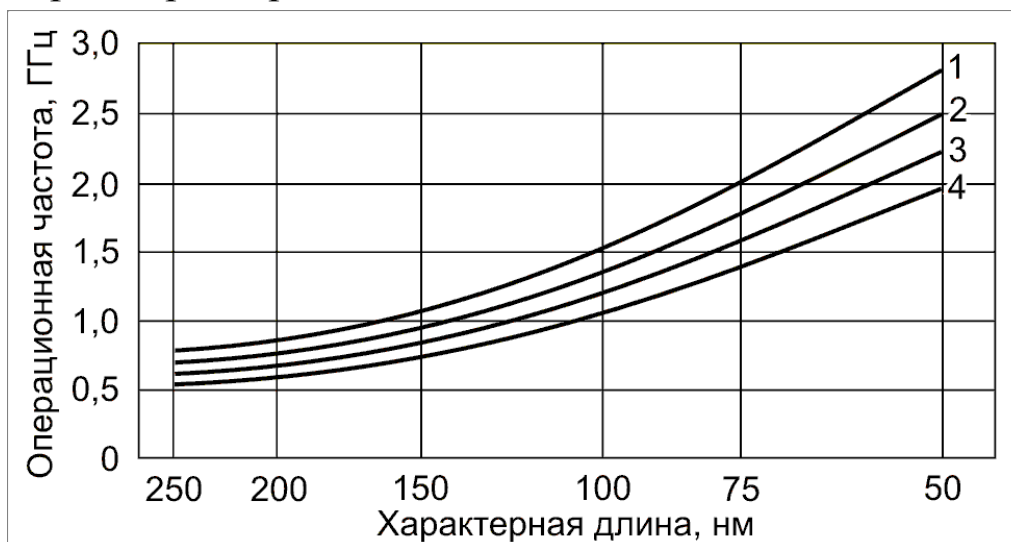


Рис. 1.18 Тактовая частота, как функция характерного размера активного элемента ИС

Построенные на рисунке кривые отвечают следующим параметрам цепей связи в ИС:

- 1) Связи образованы плёночными проводниками из меди, изолирующие прослойки выполнены из полимерных материалов с диэлектрической проницаемостью порядка 2,
- 2) Связи образованы плёночными проводниками из алюминия, изолирующие прослойки выполнены из полимерных материалов с диэлектрической проницаемостью порядка 2,
- 3) Связи образованы плёночными проводниками из меди, изолирующие прослойки выполнены из оксидов с диэлектрической проницаемостью порядка 15,
- 4) Связи образованы плёночными проводниками из алюминия, изолирующие прослойки выполнены из оксидов с диэлектрической проницаемостью порядка 15.

Задержка цифрового сигнала на связях (RLC – модель)

В этом случае шину следует рассматривать как RLC линию передачи, важной характеристикой которой является волновое сопротивление

$$Z_0 = \sqrt{\frac{L_1}{C_1}}$$

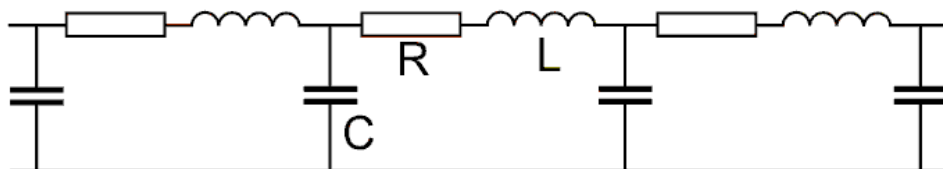


Рис. 1.19. Эквивалентная схема для RLC модели

Здесь C_1 и L_1 – погонная ёмкость и погонная индуктивность линии. Для микрополосковых линий $Z_0 = 50 - 120$ Ом. В этом случае отдельной задачей становится согласование транзисторных схем с таким малым волновым сопротивлением линии связи. Для этого необходимо введение специальных транзисторных каскадов, обеспечивающих согласование импедансов входа и выхода с волновым сопротивлением линии передачи.

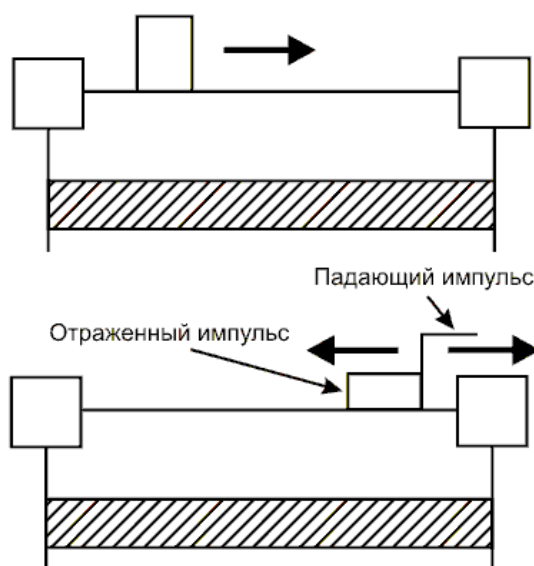


Рис.1.20. Отражение импульса

Скорость распространения сигнала по линии передачи :

$$v_{zp} \cong \frac{c_0}{\sqrt{\epsilon_{эф}}}$$

При $\varepsilon_{\text{эф}} \cong 9$, получим $v_{\text{гр}} = 10^8$ м/с. Пусть L снова обозначает длину линии, τ - время пробега импульса $\tau = L/v_{\text{гр}}$. При $L = 1$ см получаем время пробега: $\tau = 10^{-2}/10^8 = 10^{-10}$ с

Если длительность импульса короче, чем время распространения по линии, то возможно возникновение отражения (Рис 1.20). Многократное отражение импульсов может привести к нарушению работы схемы. Таким образом, в цифровую технику приходят методы техники СВЧ.

При наличии больших полей рассеяния фазовая скорость волны в линии становится функцией частоты. Таким образом, имеет место явление дисперсии фазовой скорости в линии передачи. При этом возникает искажение спектра импульса.

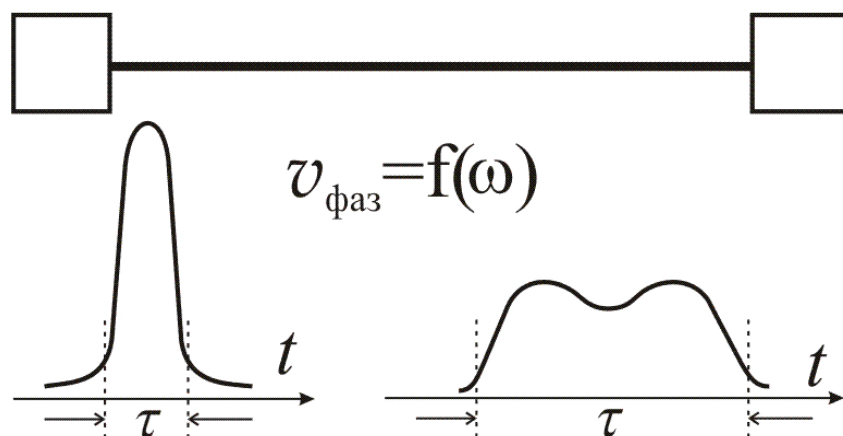


Рис. 1.21. Искаженная форма и длительность импульса

При длительности импульса $\tau < 10^{-10}$ и $f_{\text{такт}} > 3 \cdot 10^9$ Гц возникает проблема отражений и искажения формы импульсов за счёт дисперсии. Уменьшение размеров ИС облегчает борьбу с задержкой и дисперсией, поскольку становятся короче все связи внутри ИС и между различными ИС. Часть внутренних связей даже при весьма высокой тактовой частоте может рассматриваться как РС связи. Другая часть связей, которые имеют большую длину, должна согласовываться как RLC линии передачи. Это создаёт дополнительные трудности при проектировании ИС с высокой тактовой частотой.

1.2.5. Ограничения на размеры элементов, обусловленные сильными электрическими полями

1. Поле в подзатворном окисле ПТ должно быть достаточным для образования инверсионного слоя и одновременно меньше пробивного поля окисла. Кроме того, не должно происходить смыкания ОПЗ $p - n$ переходов истока и стока. Эти требования накладывают существенные ограничения на длину канала ПТ.

Для количественной оценки ограничений считаем, что прокол не происходит, пока длина канала больше трех длин $p-n$ переходов стока и истока ($d > 3L$) [1]. Высота барьеров $p - n$ переходов $\sim 2e\psi_B$. Тогда

$$d \geq 3 \sqrt{\frac{\epsilon_s \psi_B}{\pi e N_a}} \quad (1.24)$$

Для уменьшения d необходимо увеличить N_a . При этом возрастает поле в диэлектрике, которое необходимо для образования $n -$ канала:

$$E_i = \frac{\epsilon_s}{\epsilon_i} E_{sn} = \frac{4}{\epsilon_i} \sqrt{\epsilon_s \pi e \psi_B N_a}, \quad (1.25)$$

которое должно быть меньше пробивного поля в окисле $E_c = 6 \cdot 10^6$ В/см. Приравнявая $E_i = m E_c$, где $m \approx 1/4$, находим максимальную концентрацию примесей N_m :

$$N_m = \frac{\epsilon_i^2 m^2 E_c^2}{16 \pi \epsilon_s \psi_B} \quad (1.26)$$

и, подставляя (1.26) в (1.24) определяем минимальную длину ОПЗ $p-n$ перехода, а следовательно, и минимальную длину канала:

$$d_m \approx \frac{12 \epsilon_s \psi_B}{\epsilon_i m E_c} \quad (1.27)$$

Для кремниевого ПТ $\epsilon_s/\epsilon_i \approx 3$, $\psi_B \approx 0,5$ В, $N_m = 3 \cdot 10^{17}$ см⁻³, $d_m = 0,2$ мкм. Для нормальной работы ПТ минимальная толщина диэлектрика a должна быть больше длины туннелирования электронов. Для кремниевых ПТ эта величина составляет 50 Å. Если принимать во внимание возможность пробоя диэлектрика,

то ограничение выглядит следующим образом: $a > V_g/mE_c$. Если $V_g = 2$ В, то $a > 130$ Å.

2. Минимальное напряжение между стоком и истоком – величина постоянная и равная $V_{Dm} \approx 10kT/e$. Как следствие, с уменьшением длины канала усиливаются эффекты разогрева электронов. При релаксации на фононах разогрева не происходит в том случае, если дрейфовая скорость электронов (v_d) меньше скорости звука (u). В другой форме это: $\frac{\mu V_D}{d} = \frac{10\mu kT}{ed} < u$ или

$$d \geq \frac{10\mu kT}{eu} = d_m \quad (1.28)$$

Возникающий при $d < d_m$ разогрев электронного газа приводит к лавинному пробую, к появлению токов утечки, к насыщению дрейфовой скорости. Возможно также такое явление, как инжекция электронов в окисел из канала. Это вызывает нестабильность порогового напряжения. При $\mu = 10^2$ см²/В·с, $T = 300$ К, $u = 5 \cdot 10^5$ см/с получаем минимальную длину канала $d_m = 0,5$ мкм.

3. Вследствие постоянства минимального напряжения между стоком и истоком (V_{Dm}), уменьшение размеров элементов приводит к росту электрических полей, а следовательно, и к увеличению дрейфовой скорости носителей. Если не происходит пробоя, при сильных полях возможно насыщение дрейфовой скорости, что ограничивает плотность тока. Таким образом возникает ограничение на площадь поперечного сечения канала ПТ.

4. Для того чтобы дробовый шум в канале был мал, необходимо, чтобы число носителей в канале ПТ было достаточно велико, т. е. $n_M w d \gg 1$, где $n_M = 10^{12}$ см⁻² – максимальная поверхностная плотность носителей, ограниченная электрической прочностью SiO₂. Принимая равными длину и ширину канала ($w \approx d$), получим $d_m \approx 10 (n_M)^{-1/2} \approx 10^{-5}$ см. Физически малая величина дробового шума означает, что заряд, протекший через канал ПТ во время переключения, много больше заряда одного электрона.

1.2.6. Ограничения размеров элементов памяти

Кроме ограничений, рассмотренных для ПТ и справедливых также для элементов памяти, для последних существуют и свои специфические ограничения.

Область, в которой находится информационный заряд элемента памяти, должна быть отделена от других областей потенциальным барьером достаточных ширины и высоты. В этом случае не будет происходить утечки заряда вследствие надбарьерного или туннельного переноса. Минимальная ширина потенциального барьера (минимальный размер элемента памяти)

определяется выражением $d_m \approx \frac{10}{\hbar} 2m\Delta\varepsilon$, где m – масса электрона, $\Delta\varepsilon$ – высота барьера, $\Delta\varepsilon \geq kT$. Для $T=300$ К получаем $d_m \approx 0,01$ мкм.

Существенной причиной сбоев в работе элементов памяти может оказаться влияние ионизирующих излучений (радиоактивности). Попадая в полупроводник, α - частицы, например, могут генерировать до 10^6 электронно - дырочных пар в объеме 10 мкм³. Разделение носителей в электрических полях приводит к появлению захваченного заряда, который в малых объемах может восприниматься как информационный заряд. Для того чтобы индуцированный облучением заряд не приводил к инверсии состояния в элементе памяти, необходимо, чтобы он был мал по сравнению с информационным зарядом. С другой стороны, истинное зарядовое состояние элемента памяти должно быть таково, чтобы поле в барьере было много меньше пробивного. Оценки величины минимального размера дают $d_m \approx 0,6$ мкм.

1.3. Ограничения интеграции элементов

Максимальное число невзаимодействующих элементов с линейными размерами d , которое можно разместить на чипе площадью S при использовании планарной технологии, определяется как $N_M = S/d^2$. Естественно, что этот предел является теоретическим, а реально достижимая степень

интеграции N всегда значительно меньше N_M . Основное ограничение состоит в том, что элементы могут взаимодействовать друг с другом. Эти взаимодействия становятся особенно существенными при уменьшении размеров.

1.3.1. Предельная степень интеграции

Решающим фактором, определяющим степень интеграции, является характерный размер ИС - в случае КМДП схем это длина канала. Длина канала определяется разрешающей способностью литографии. В обозримом будущем литография в массовом производстве ИС – это оптическая фотолитография. На рисунке показано совершенствование разрешающей способности трёх основных видов литографии по годам: 1 – фотолитография, 2- электронная литография, 3 – ионная литография.

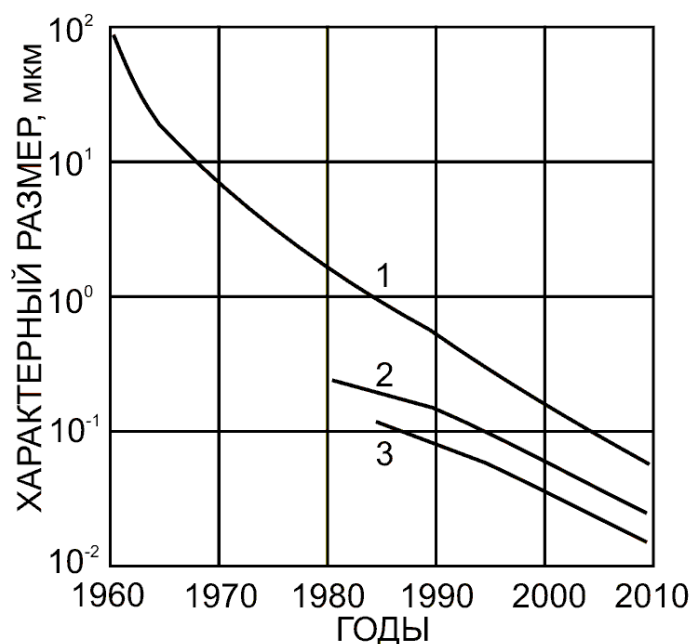


Рис.1.22. Совершенствование литографических методов

Минимальная площадь одного элемента ИС

Рассмотрим МДП-транзистор. Положим : $l_T = 10 l_K$, $w = 20 l_K$
 Кроме того, для площадей транзистора $S_{тр} = l_T \cdot w = 200 \cdot l_K^2$,
 $S_{эл} = 5 \cdot S_{тр}$, где: $S_{тр}$ – площадь, занимаемая непосредственно транзистором, а $S_{эл}$ – площадь, приходящаяся на один транзистор с учетом подводящих шин и необходимого расстояния между ними. Из приведённых двух формул получаем следующее

соотношение, которое связывает площадь, приходящуюся на один транзистор, с его характерным размером: $S_{эл} = 1000 \cdot l_K^2$

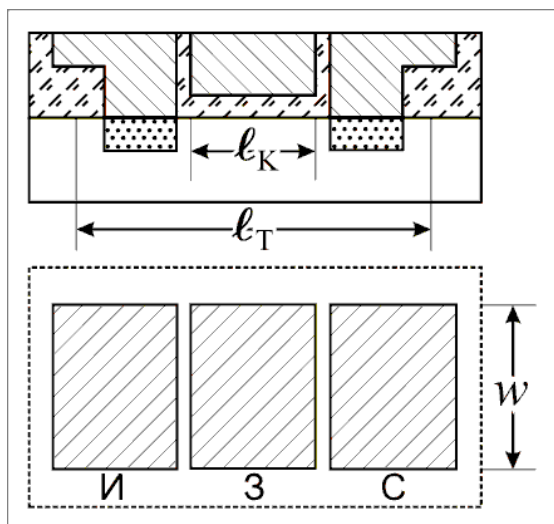


Рис. 1.23. Размеры областей МДП транзистора

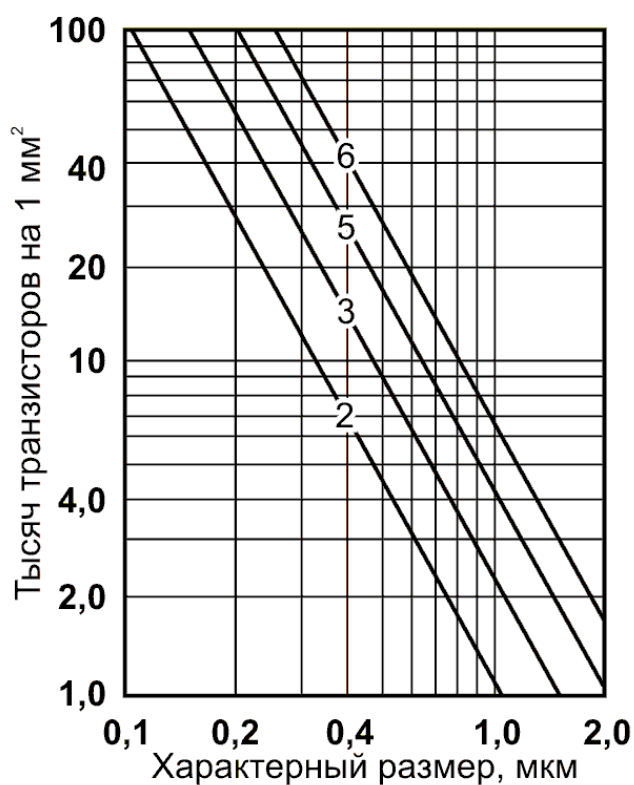


Рис. 1.24. Плотность элементов на чипе при многоуровневой разводке

Полученная оценочная формула справедлива для случая расположения проводников разводки только на одном уровне. Статистические данные производства ИС показывают, что с

достаточной степенью надежности можно положить:
 $S_{эл} = (1000 \cdot l_k^2) / m$, где m – число уровней разводки межсоединений.

Это подтверждается графиками, приведёнными на рис. 1.24, на котором показана плотность размещения элементов ИС в тысячах транзисторов на квадратный миллиметр в функции от характерного размера транзистора (длины канала). Цифры на графиках отвечают числу уровней разводки.

Плотность записи для ЗУ на основе МДП транзисторов.

На один элемент памяти (ЗУ) приходится две пары КМДП – транзисторов. Поскольку конструкция ЗУ такова, что оно имеет регулярную структуру, можно считать, что

$$S_{эл.зу} = 1,2 S_{эл.} = 1200 \frac{l_k^2}{m}$$

Число бит информации, которое может быть записано таким ЗУ, из расчета на единицу площади:

$$N_{bit} = \frac{1}{S_{эл.зу}} = \frac{m}{1200 \cdot l_k^2}$$

При $m = 6$ получаем: $N_{bit} = 5 \cdot 10^{-3} \cdot l_k^{-2}$

При $l_k = 1 \text{ мкм}$ $N_{bit} = 0,5 \text{ Мбит/см}^2$

При $l_k = 0,25 \text{ мкм}$ $N_{bit} = 8 \text{ Мбит/см}^2$

Можно предположить, что площадь ОЗУ на КМДП ИС достигает 10 см^2 . Тогда полная емкость памяти одного ОЗУ на ИС с транзисторами, имеющими длину канала $0,25 \text{ мкм}$ составит 80 Мбит или 10 Мбайт . При дальнейшем развитии технологии объём памяти ОЗУ увеличится ещё больше. При $l_k = 0,07 \text{ мкм}$ $N_{bit} = 100 \text{ Мбит/см}^2$, 10 см^2 площади ИС обеспечат 10^9 бит, т.е. ОЗУ с объёмом памяти до 1 Gbit в одной ИС.

1.3.2. Теплофизические ограничения на рост интеграции

Наиболее существенное влияние на рабочие характеристики ИС может оказывать разогрев чипа. Плотность элементов на чипе определим как $\bar{N} = N / S$ (N – число элементов на чипе), средняя частота обращения к каждому элементу (тактовая частота) равна ν . Каждое переключение сопровождается изменением напряжения на величину $\Delta V = nkT/e$ ($n \geq 10$), и при этом в

межсоединениях и самих элементах расходуется энергия $C(\Delta V)^2$, C – емкость элемента, к которому приложено напряжение. В единицу времени в ИС выделяется количество тепла, равное

$$N\nu \frac{C(\Delta V)^2}{2}$$

Пусть Q_M – максимальное количество тепла, отводимое с единицы площади чипа. Тогда для сохранения теплового баланса на чипе должно выполняться условие: $N\nu \frac{C(\Delta V)^2}{2} \ll Q_M S$.

Ограничение на функциональное быстроедействие (ФБ)

$$N\nu < \frac{2Q_M S}{C(\Delta V)^2} \quad (1.29)$$

Емкость элемента $C = \frac{\varepsilon d^2}{4\pi a} \approx \varepsilon^* d$, где a – толщина диэлектрика,

$\varepsilon^* = \frac{\varepsilon d}{4\pi a}$ – величина, которую можно считать постоянной.

Функциональное быстроедействие единицы площади чипа определяется как $\bar{N}\nu = N\nu/S$. Выражение (1.29) можно переписать в другом виде:

$$\bar{N}\nu < \frac{2Q_M e^2}{\varepsilon^* d n^2 (kT)^2} \quad (1.30)$$

Таким образом, ФБ зависит только от размера элементов, теплоотвода и температуры. Значения величин Q_M для воздушного и водяного способов охлаждения составляют $4 \cdot 10^6$ эрг/см²·с и $2 \cdot 10^8$ эрг/см²·с соответственно. Оценки величины ФБ для этих значений Q_M дают 10^{18} Гц элементов и 10^{20} Гц элементов.

Отвод тепла из ИС происходит за счет теплопроводности материала и описывается выражением

$$q = -\kappa \cdot \text{grad} T$$

Здесь q – плотность потока тепла. Размерность $[q]$ Вт/м², κ – коэффициент теплопроводности. Размерность: $[\kappa]$ Вт/м·К

Коэффициент теплопроводности технической меди $\kappa = 200$ Вт/м·К и слабо зависит от температуры при $T > 100$ К.

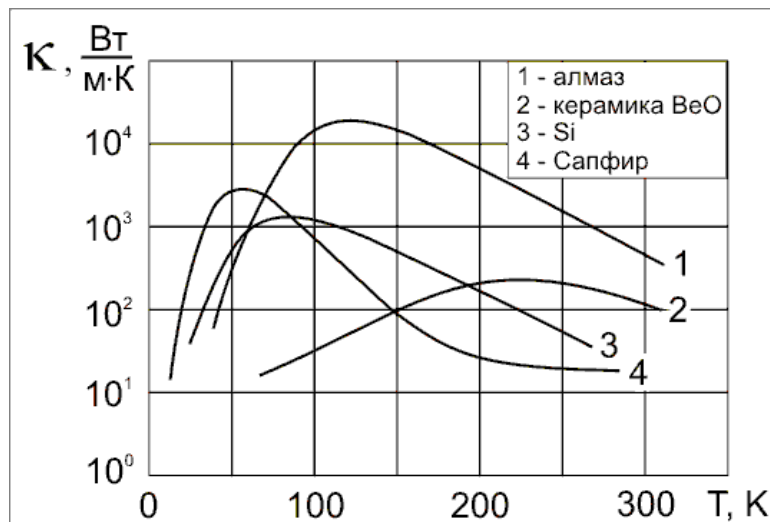


Рис. 1.25. Температурные зависимости коэффициентов теплопроводности материалов микроэлектроники

Рассматривая ИС с точки зрения теплофизики, можно сказать следующее:

- ИС образована из микромощных ключей. Энергия в ней расходуется только на перезарядку емкостей. При перезарядке вся энергия превращается в тепло.
- Тепло удаляется за счет теплопроводности материала.

Вычислим плотности тепловой мощности, рассеиваемой в объеме ИС. Рассмотрим элемент схемы – кубик с объемом L^3 . Емкость электродов, расположенных внутри рассматриваемого элемента:

$$C = \frac{K\varepsilon S}{L} = \varepsilon KL, \text{ где } K = 30 \div 100 \text{ – конструктивный}$$

коэффициент, ε – диэлектрическая проницаемость материала. Пусть $\varepsilon = 10^{-10}$ Ф/м, $L = 30$ мкм, тогда при $K = 30 \div 100$ получаем для ёмкости элемента: $C = 0.1 - 0.3$ пФ.

Энергия, запасенная в элементе $E = \frac{CU^2}{2} = \frac{\varepsilon KLU^2}{2}$. Пусть количество операций в секунду - ν , тогда мощность, поглощаемая одним элементом : $P = E\nu = \frac{\varepsilon \nu KLU^2}{2}$

Если обозначить через α - отношение числа активных элементов к их полному числу, то для ИС, представляющей собой процессор, можно положить $\alpha = 0,01-0,1$. Если ИС

выполняет функцию оперативного запоминающего устройства, то правильнее считать $\alpha = 0,001-0,01$. Удельная плотность мощности в ИС, то есть мощность, приходящаяся на единицу объёма,

$$\Pi = P \frac{\alpha}{V} = P \frac{\alpha}{L^3} = \frac{1}{2} \varepsilon \nu \alpha K \left(\frac{U}{L} \right)^2$$

Решим простую теплофизическую задачу в предположении, что ИС имеет форму шара радиуса r и источник тепла, связанный с перезарядкой емкостей. На поверхности ИС задана фиксированная температура T_0 . Вычислим распределение температуры в такой ИС

Плотность потока тепловой мощности через сферу радиуса r :

$$q(r) = -\kappa \cdot \text{grad}[T(r)]$$

Используя закон сохранения энергии, мы должны приравнять мощность, проходящую через сферу радиуса r , и мощность, которая выделяется источниками тепла в объёме шара радиуса r :

$$4\pi r^2 q(r) = \frac{4}{3} \pi r^3 \Pi$$

Тогда $q(r) = \frac{1}{3} r \Pi$ и имеем дифференциальное уравнение для температуры: $\frac{1}{3} r \Pi = -\kappa \frac{dT(r)}{dr}$ с граничными условиями: $T = T_0$ при $r = R$. В результате получаем решение дифференциального уравнения с заданными граничными условиями:

$$T(r) = T_0 + \frac{P}{6\kappa} (R^2 - r^2)$$

Найдём перегрев в центре ИС при $r = 0$

$$\Delta T = T(r)|_{r=0} - T_0 = \frac{P}{6\kappa} R^2$$

Подставим сюда выражение для P и получим:

$$\Delta T = \frac{1}{12} \cdot \frac{\varepsilon \nu \alpha K}{\kappa} \left(\frac{U}{L} \right)^2 \cdot R^2$$

Если число элементов в ИС : $\frac{4\pi R^3}{3L^3} = N$, то $\frac{R}{L} = \left(\frac{3}{4\pi}\right)^{1/3} \cdot N^{1/3}$ и $\left(\frac{R}{L}\right)^2 = \left(\frac{3}{4\pi}\right)^{2/3} \cdot N^{2/3}$. Поскольку $\left(\frac{3}{4\pi}\right)^{2/3} = 0,3848$, то окончательно получим

$$\Delta T = \frac{1}{32} \frac{\varepsilon \nu K U^2 \alpha}{k} N^{2/3} \quad (1.31)$$

Найдем также полную потребляемую мощность и объём, занимаемый ИС: $P_{полн.} = P \cdot L^3$, $V = N \cdot L^3$

Пусть заданы следующие параметры: $U = 1$ В,
 $\nu = 10^9$ Гц $\alpha = 0.1$ $\varepsilon = 10^{-10}$ Ф/м
 $\kappa = 20$ Вт/м·К (для Si) $\Delta T = 20$ К $K = 100$

Каково максимально допустимое число элементов ИС из расчетов нормального теплового баланса. Из (1.31) имеем :

$$N = \left[\frac{1}{U} \cdot \sqrt{\frac{32\kappa}{K\varepsilon\nu\alpha}} \cdot \Delta T \right]^3$$

$$\sqrt{\frac{32\kappa}{K\varepsilon\nu\alpha}} \cdot \Delta T = \sqrt{\frac{32 \cdot 20 \cdot 20}{10^2 \cdot 10^{-10} \cdot 10^9 \cdot 0,1}} = \sqrt{1,28 \cdot 10^4} = 110$$

при $U = 1$ В получаем число элементов в составе ИС: $N \cong 10^6$

Таким образом, при теплоотводе за счет теплопроводности кремния одна ИС на основе КМДП (микромощная электроника!) может содержать в одном корпусе до 10^6 транзисторов. При этом при $L = 30$ мкм полная потребляемая (и рассеиваемая) мощность и объем корпуса составят:

$$P_{полн} = \frac{1}{2} \varepsilon \nu \alpha K L U^2 N = 15 \text{ Вт},$$

$$V = 10^6 \cdot (3 \cdot 10^{-5})^3 = 27 \cdot 10^{-9} \text{ м}^3 = 27 \text{ мм}^3$$

Поскольку для элементов памяти тактовая частота значительно ниже, чем, например, в процессоре, ограничения, обусловленные нагревом чипа, для них значительно «мягче».

Параметр α может быть снижен до $\alpha = 10^{-3}$. Тогда : $\sqrt{\frac{32\kappa}{K\varepsilon\nu\alpha}} = 1100$, а допустимое число элементов вырастет до $N = 10^9$. При этом

потребляемая мощность вырастет до 150 Вт, а объем ИС составит 27см^2 . О сооружении такого рода ОЗУ можно говорить как о перспективе обозримого, но не очень далекого будущего. Можно представить себе размер стороны кубика, вмещающий один элемент: $L = 10\text{мкм}$. Тогда для ИС ОЗУ при $N = 10^9$ получим: $P = 50\text{ Вт}$, $V = 1\text{ см}^3$

Среди других причин, которые могут ограничивать степень интеграции, можно перечислить паразитные связи между элементами, разогрев межсоединений и выход их из строя, а также паразитные связи между соединениями.

1.4. Заключение

Таким образом, мы рассмотрели основные причины, приводящие к физическим ограничениям при уменьшении размеров элементов ИС.

Ограничения в литографии, обусловленные рассеянием луча в резисте и в полупроводнике, дают минимальную ширину технологических линий $\approx 0,01\text{ мкм}$. Очевидно, что это ограничение не является столь критическим, как например, длина волны света при традиционной оптической литографии. Смягчить это ограничение можно использованием излучений с меньшими длинами волн. В настоящее время интенсивно разрабатываются новые литографические методы, позволяющие получать элементы с характерными размерами вплоть до 30 нм. Это EUVL (extreme ultra violet lithography), DUV (dark ultra violet lithography), nanoimprint lithography. Кроме того, все шире используются процессы самоорганизации для формирования наноструктур.

Флуктуации концентрации легирующей примеси, явление поверхностной диффузии, электрические проколы и пробой диэлектриков и p - n переходов ограничивают размеры биполярных и полевых транзисторов с p - n переходами величиной 0,1 мкм.

Функциональное быстродействие и размеры элементов ограничены также явлением джоулева разогрева, возникающим вследствие недостаточного теплоотвода. Это ограничение менее существенно для элементов памяти, и может быть в значительной

степени снято использованием низкотемпературного охлаждения ИС.

Существует также проблема радиационной стойкости элементов. Особенно существенна она для элементов памяти, поскольку может приводить к инверсии битов и сбоям в работе. В структурах с размерами меньше 0,1 мкм начинают проявляться новые механизмы проводимости – туннельный и баллистический. Кроме того, становится возможным использование новых физических эффектов, проявляющихся как результат размерного квантования носителей заряда в элементах ИС нанометровых размеров. Таким образом, возникает проблема разработки новых сверхминиатюрных элементов, в основу работы которых положены новые физические принципы.

Интерес к проблеме физических ограничений неизбежно возрастает, поскольку это определяется всем развитием микроэлектроники. Развитие технологии, особенно в последнее десятилетие, позволило не только приблизиться к характерным субмикронным размерам, но и сформировать раздел микроэлектроники, который получил общепризнанное название - **наноэлектроника**. Дальнейший прогресс в этой области требует постоянного анализа принципов работы элементов с привлечением фундаментальных физических подходов. Важнейшим элементом совершенствования технологии становятся физические принципы, лежащие в основе нанотехнологии.

Литература

1. Гуляев Ю. В., Сандомирский В. Б., Суханов А. А., Ткач Ю. Я. Физические ограничения минимальных размеров элементов современной микроэлектроники. // УФН. 1984. Т.144. ВЫП. 3. С. 475-495
2. Technology Roadmap for Nanoelectronics, Editors: R.Compano, L.Molenkamp, D.J.Paul. Microelectronics Advanced Research Initiative - MELARI NANO (См. :URL <http://www.cordis.lu/esprit/src/melari.htm>)
3. ITRS – International Technology Roadmap for Semiconductors. 1999 Edition, 2000 Update, 2001 Edition. (См. также: <http://public.itrs.net/Home.htm>)
4. Montemerlo M. S., Love J. C., Opiteck G. J., Goldhaber-Gordon D. and Ellenbogen J. C. “Technologies and Designs for Electronic Nanocomputers” MITRE Technical Report №96W0000044, The MITRE Corporation, McLean, VA, July 1996. (См. также: URL:<http://www.mitre.org/research/nanotech>)
5. Стриха В. И. , Бузанева Е. В. Физические основы надежности контактов металл-полупроводник в интегральной электронике. М.: Радио и связь, 1987.
6. Thompson S., Packan P., Bohr M. MOS Scaling: Transistor challenges for the 21 century // Intel Technology Journal,1998. V. Q3. P. 1-19
7. Вендик О. Г. Микроэлектроника. Конспект лекции. СПб, 2002 г. (См: URL: <http://www.1024.ru/science/micro/index.htm>)
8. Havemann R.Y., Hutchdy J.A., High-Performance Interconnections: An Integration Overview // Proc. IEEE, Vol. 89, No. 5, P. 586-601 (May 2001).
9. Дорфман В. Ф. Микрометаллургия в микроэлектронике. М.: Металлургия, 1978.

2. КВАНТОВАЯ МЕХАНИКА ПРОСТЕЙШИХ СТРУКТУР

Если область локализации носителей заряда сравнима или меньше длины волны электронов (дырок), то в такой структуре возникают эффекты размерного квантования. Эффекты такого типа начали впервые исследоваться теоретически [1], а затем и экспериментально [2] в пленках. Работа [3] была первой, где обсуждалось квантование энергии электронов в приповерхностной области пространственного заряда (ОПЗ) полупроводников. Начало интенсивных исследований квантовых эффектов в приповерхностных областях полупроводников относится к концу 60-х годов. К настоящему времени сформировалась обширная область физики, изучающая квантово - размерные структуры и явления. Наиболее полно в обзорах на эту тему представлены двумерные системы [4, 5, 6]. Важнейшей частью исследований является изучение поверхностных каналов и границ раздела полупроводников при условии квантования в них энергии носителей заряда. Для узких каналов, где есть квантование, удалось получить наиболее определенные результаты о механизмах рассеяния носителей заряда: фононном, заряженными центрами и шероховатостями поверхности.

В настоящее время технология изготовления полупроводниковых наноструктур находится на достаточно высоком уровне и постоянно продолжает совершенствоваться. Гетеропереходы $\text{GaAs-Al}_x\text{Ga}_{1-x}\text{As}$ и структуры металл – окисел – полупроводник (МОП) являются наиболее исследованными типами квантово-размерных структур. В настоящее время совершенствование технологии позволило получать гетеропереходы на основе других полупроводниковых материалов – например германия и кремния – $\text{Ge}_x\text{Si}_{1-x}/\text{Si}$ или $\text{Ge}_x\text{Si}_{1-x}/\text{Ge}$. Во всех вышеперечисленных типах двумерных структур (2D) движение ограничено в одном направлении. Существует ряд методов, позволяющих создавать структуры, в которых движение имеет одномерный (1D структуры или квантовые нити) или нульмерный (0D структуры или квантовые точки) характер. Примерами таких материалов и структур

являются квантовые мостики на поверхности полупроводников, приборы одноэлектроники, лазеры на квантовых точках и многие другие. Понижение размерности, учет и использование в работе приборов новых физических эффектов и явлений стало одним из основных направлений исследований в современной наноэлектронике.

В квантовой механике известен ряд одномерных потенциалов $U(x)$, для которых существует аналитическое решение уравнения Шредингера. Простые квантовые модели обычно используются для описания электронных состояний в реальных структурах. Такой подход позволяет с достаточной точностью найти волновые функции и спектры энергии в некоторых типах квантовых структур: квантовых ямах, нитях, точках, двумерных каналах и т. п. [7] Рассмотрение обычно ведется в рамках приближения эффективной массы (m), в котором электронный спектр определяется этим параметром. Рассмотрим ряд простейших примеров.

2.1. Прямоугольная яма

Для одномерного случая прямоугольной потенциальной ямы конечной глубины U_0 уравнение Шредингера запишется в виде

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi}{dx^2} + U(x)\Psi = \varepsilon\Psi, \quad (2.1)$$

где потенциальная энергия определяется как

$$U(x) = \begin{cases} -U_0, & |x| < L/2 \\ 0, & |x| > L/2 \end{cases} \quad (2.2)$$

Здесь m – эффективная масса, \hbar – постоянная Планка, ε – энергия. В яме при $\varepsilon < 0$ спектр собственных значений энергии будет дискретным, а вне ямы, при $\varepsilon > 0$ – непрерывным. В силу симметрии потенциала $U(x)$ по отношению к замене x на $-x$ собственные функции будут либо четными, либо нечетными.

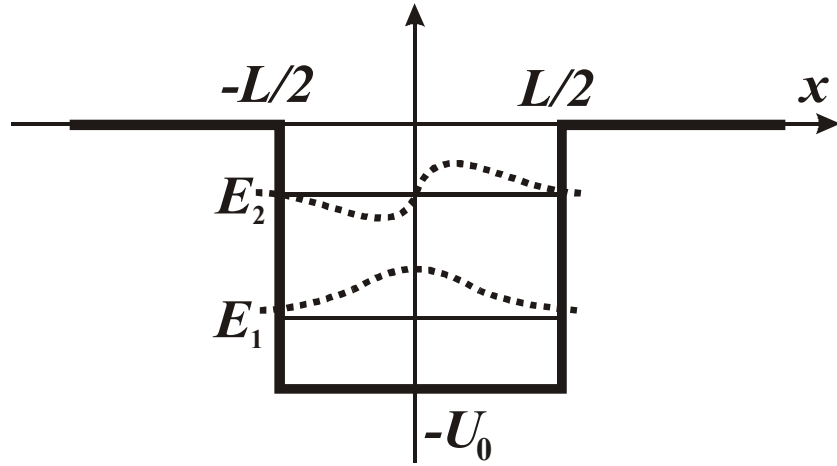


Рис. 2.1. Волновые функции и уровни энергии в прямоугольной квантовой яме.

Для области $|x| < L/2$ получим:

$$\Psi_n(x) \propto \begin{cases} \cos k_n x \\ \sin k_n x \end{cases}, \quad (2.3)$$

где $k_n = \sqrt{2m(U_0 - |\varepsilon_n|)/\hbar^2}$, а при $x \rightarrow \pm\infty$ волновые функции экспоненциально обращаются в нуль:

$$\Psi_n(x) \propto e^{\mp \kappa_n x}, \quad (2.4)$$

где $\kappa_n = \sqrt{2m|\varepsilon_n|/\hbar^2}$ определяет скорость убывания волновой функции. Знаки «+» и «-» относятся к областям $x < -L/2$ и $x > L/2$ соответственно. Собственные значения энергии ε_n находятся из решения трансцендентного уравнения, которое, в свою очередь, следует из условий непрерывности Ψ и $d\Psi/dx$ на границах ямы. Число дискретных уровней в яме (четных и нечетных) определяется условием

$$(n_{\max} - 1) < \sqrt{2mU_0 L^2 / (\pi\hbar)^2} \leq n_{\max} \quad (2.5)$$

Если выполняется неравенство $2mU_0 L^2 / (\pi\hbar)^2 \gg 1$, то яма может считаться бесконечно глубокой и при $|x| < L/2$ для нормированной волновой функции имеем

$$\Psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{\pi x}{L} n \quad (n=1, 2, 3 \dots) \quad (2.6)$$

Собственные значения энергии, отсчитываемой от дна ямы

$$\varepsilon_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2 \quad (2.7)$$

Для мелкой ямы, т. е. когда выполняется неравенство $2mU_0L^2/(\pi\hbar)^2 \ll 1$, существует только одно (четное) решение трансцендентного уравнения

$$\varepsilon_1 = \frac{mU_0^2L^2}{2\hbar^2} \quad (2.8)$$

Единственный существующий в мелкой яме энергетический уровень расположен вблизи её «верха». Соответствующая волновая функция почти постоянна в яме и экспоненциально убывает вне её:

$$\Psi_1(x) = \left(\frac{2m\varepsilon_1}{\hbar^2} \right)^{1/4} e^{\mp \kappa_1 x} \quad (2.9)$$

В одномерной потенциальной яме любой формы всегда имеется по крайней мере один уровень энергии – даже если глубина ямы очень мала. Это свойство, однако, специфично именно для одномерного случая и не имеет места в более реальном случае трехмерной потенциальной ямы. Если глубина ямы $|U|$ такова, что $|U| \ll \hbar^2/mL^2$, где L – порядок величины линейных размеров ямы, то в ней нет ни одного дискретного уровня энергии. Другими словами, если яма недостаточно глубока, то в ней нет связанных состояний – частица не может «захватиться» ямой. Случай трехмерной потенциальной ямы относится к рассмотрению состояний в квантовых точках.

2.2. Треугольная яма

Электроны вблизи границы раздела полупроводник-диэлектрик движутся в поле довольно сложного потенциала. С одной стороны границы раздела на них действует поле периодического потенциала полупроводника. На это поле накладывается медленно меняющееся электрическое поле, создаваемое приложенным напряжением, разностью работ выхода металла и полупроводника и неподвижными заряженными примесями. С другой стороны на электроны действует поле атомов диэлектрика. На границе раздела

формируется, как правило, высокий потенциальный барьер, что приводит к отталкиванию электронов от диэлектрика.

В приближении эффективной массы для одномерного случая волновая функция $\Psi_i(z)$ удовлетворяет уравнению Шредингера:

$$\frac{\hbar^2}{2m_z} \frac{d^2 \Psi_i}{dz^2} + [E_i - V(z)] \Psi_i(z) = 0 \quad (2.10)$$

Потенциальную энергию $V(z)$, можно представить в виде суммы трех слагаемых:

$$V(z) = V_d(z) + V_s(z) + V_l(z) \quad (2.11)$$

В выражении (2.11) разделены вклады неподвижного пространственного заряда (ПЗ), то есть потенциальная энергия в обедненном слое, индуцированных зарядов в слое ПЗ или потенциальная энергия электронов, образующих слой пространственного заряда, и изображений зарядов, обусловленных существованием границы раздела полупроводник – диэлектрик, т. е. с различием диэлектрических проницаемостей полупроводника и диэлектрика. Предполагается, что эта граница раздела резкая и на ней образуется бесконечно высокий потенциальный барьер, существование которого приводит к отталкиванию электронов от диэлектрика.

Так как слагаемое $V_s(z)$ в соотношении для эффективного одноэлектронного потенциала (2.11) выражается через волновые функции, то решение уравнения Шредингера является нелинейной задачей на собственные значения. Такой тип задач обычно решается итерационными методами.

Если в соотношении (2.11) исключить потенциал изображения $V_s(z)$, то оставшиеся слагаемые будут описывать потенциал, который по мере удаления от границы раздела линейно возрастает, а затем изгибается, приближаясь к постоянному значению. Вклад инверсионного слоя $V_s(z)$ становится постоянным уже на расстоянии, составляющем несколько значений z_{cp} (среднего удаления электронов от границы раздела), а вклад обедненного слоя $V_d(z)$ – на расстоянии превышающем z_d (толщину обедненного слоя). Поскольку z_d обычно значительно превышает z_{cp} , кривизной потенциала обедненного слоя $V_d(z)$ можно пренебречь, особенно если заряд

обедненного слоя мал по сравнению с зарядом инверсионного слоя.

В рамках приближения треугольной ямы [8] потенциальная энергия бесконечно велика при $z < 0$ и линейно возрастает при $z > 0$.

$$V(z) = \begin{cases} eFz, & z > 0 \\ \infty & z \leq 0 \end{cases} \quad (2.12)$$

Здесь F – напряженность эффективного электрического поля, которая определяется выражением:

$$F = \frac{4\pi(N_p + fN_s)e}{\kappa_{sc}} \quad (2.13)$$

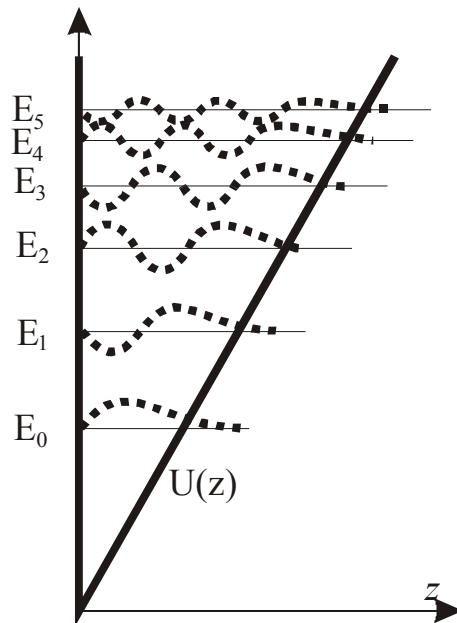


Рис. 2.2 Волновые функции и уровни энергии в треугольной квантовой яме. Пунктиром изображен вид собственных функций, отвечающих шести нижним уровням энергии

где N_p – число неподвижных зарядов на единицу площади в обедненном слое, N_s – концентрация электронов в инверсионном слое, κ_{sc} – диэлектрическая проницаемость полупроводника. Численные значения коэффициента f определяют: поле на границе раздела ($f=1$), поле, создаваемое только обедненным слоем ($f=0$), среднее поле в инверсионном слое ($f=1/2$).

Решением уравнения Шредингера с граничными условиями, согласно которым волновая функция равна нулю при $z = 0$ и на бесконечности, является функция Эйри.

$$\Psi_i(z) = Ai \left[\frac{2m_z eF}{\hbar^2} \left(z - \frac{E_i}{eF} \right) \right], \quad (2.14)$$

где собственные значения E_i имеют асимптотический вид при больших i :

$$E_i = \left(\frac{\hbar^2}{2m_z} \right)^{1/3} \left[\frac{3\pi eF}{2} \left(i + \frac{3}{4} \right) \right]^{2/3}, \quad (2.15)$$

где m_z – эффективная масса электрона в инверсионном слое в направлении оси z , перпендикулярном к поверхности. Корни функции Эйри α_k ($k=1, 2, 3, \dots$) определяют дискретные уровни энергии в треугольной яме.

$$E_i = \left(\frac{e^2 F^2 \hbar^2}{2m_z} \right)^{1/3} \alpha_{i+1} \quad (i = 0, 1, \dots) \quad (2.16)$$

В частности, энергия основного состояния ($\alpha_1=2,34$) равна:

$$E_0 = 1,86 \left(\frac{e^2 F^2 \hbar^2}{m_z} \right)^{1/3} \quad (2.17)$$

Точные собственные значения получаются, если в соотношении (2.15) заменить $i + (3/4)$ при $i=0, 1, 2$ соответственно числами 0,7587, 1,7540, 2,7575. Следует отметить, что средняя величина z в i -й подзоне равна $2E_i/3eF$, а средняя величина z^2 равна $\frac{6}{5}z_i^2$.

2.3. Вариационная волновая функция для нижней подзоны в слое пространственного заряда

Для исследования свойств слоев пространственного заряда кроме численных самосогласованных решений или громоздких аналитических выражений типа функций Эйри часто используются простые аналитические волновые функции, дающие приближенные решения. Простейшее решение было предложено Фэнгом и Ховардом [9] для описания инверсионного слоя.

$$\Psi_0(z) = \left(\frac{b^3}{2}\right)^{1/2} z \exp\left(-\frac{bz}{2}\right) \quad (2.18)$$

Далее необходимо решить уравнение Пуассона, определяющее потенциал ямы, в которой находится электрон, и найти потенциальную энергию электрона в состоянии (2.18). Среднее удаление электронов инверсионного слоя от поверхности полупроводника в этом случае равно $z_0=3/b$.

Параметр b в соотношении (2.18) определяется из условия минимума потенциальной энергии, соответствующей определенным значениям плотности заряда инверсионного и обедненного слоев. Простой вид волновой функции позволяет легко найти средние значения всех членов, составляющих гамильтониан:

$$\langle T \rangle = \frac{\hbar^2 b^2}{8m_z} \quad (2.19.1)$$

$$\langle V_d \rangle = \frac{12\pi e^2 N_p}{\kappa_{sc} b} - \frac{24\pi e^2 N_A}{\kappa_{sc} b^2} \quad (2.19.2)$$

$$\langle V_s \rangle = \frac{33\pi e^2 N_s}{4\kappa_{sc} b} \quad (2.19.3)$$

$$\langle V_I \rangle = \frac{\kappa_{sc} - \kappa_{ins}}{\kappa_{sc} + \kappa_{ins}} \frac{e^2 b}{8\kappa_{sc}} \equiv \delta_\kappa \frac{e^2 b}{8\kappa_{sc}} \quad (2.19.4)$$

Здесь N_A – концентрация заряженных акцепторов, N_p – число неподвижных зарядов на единицу площади в обедненном слое, N_s – концентрация электронов в инверсионном слое, κ_{sc} и κ_{ins} – диэлектрические проницаемости полупроводника и диэлектрика соответственно. Выражения (2.19.1)-(2.19.4) определяют средние значения кинетической энергии электрона, потенциальной энергии электрона, взаимодействующего с зарядами в обедненном слое, потенциальная энергия электрона, взаимодействующего с другими электронами инверсионного слоя, и потенциал изображения. Энергия нижней подзоны определяется суммой:

$$E_0 = \langle T \rangle + \langle V_d \rangle + \langle V_s \rangle + \langle V_I \rangle \quad (2.20)$$

Полная энергия на электрон

$$E / N = \langle T \rangle + \langle V_d \rangle + \frac{1}{2} \langle V_s \rangle + \langle V_I \rangle, \quad (2.21)$$

где множитель $\frac{1}{2}$ компенсирует двукратный учет электрон-электронных взаимодействий.

Если пренебречь потенциалом изображения и вторым членом в $\langle V_d \rangle$, то полная энергия на один электрон минимальна при

$$b = \left(\frac{48\pi m_z e^2 N^*}{\kappa_{sc} \hbar^2} \right)^{1/3}, \quad (2.22)$$

где $N^* = N_p + \frac{11}{32} N_s$. Использование вариационной волновой функции (2.18) позволяет получить значения E_0 и z_0 , превышающие результаты самосогласованных расчетов в приближении Хартри без учета потенциала изображения менее чем на 7%.

2.4. Квантовые состояния в нитях и точках

В квантовой нити движение электрона ограничено в двух направлениях и свободно в третьем направлении. Уравнение Шредингера в этом случае может быть записано как

$$\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi + V(x, y) = E \Psi \quad (2.23)$$

Волновая функция, удовлетворяющая уравнению (2.23), может быть найдена в виде:

$$\Psi(x, y, z) = \frac{1}{\sqrt{L_z}} \exp\left(\frac{ip_z z}{\hbar}\right) \Psi(x, y) \quad (2.24)$$

Для случая, когда потенциальная энергия имеет цилиндрическую симметрию, т.е. $V = V(\rho)$, уравнение Шредингера должно быть записано в соответствующей системе координат. Если рассматривать длинную квантовую нить радиуса ρ_0 , то можно записать

$$V(\rho) = \begin{cases} 0, & \rho < \rho_0 \\ \infty, & \rho > \rho_0 \end{cases} \quad (2.25)$$

Решением такой задачи [8] является функция

$$\Psi_{m,n,p_z} = \frac{1}{(L_z)^{1/2} \rho_0} \left[\frac{J_m \left(j_{m,n} \frac{\rho}{\rho_0} \right)}{J_{m+1}(j_{m,n})} \right] \exp(i\varphi) \exp\left(\frac{ip_z z}{\hbar}\right). \quad (2.26)$$

Здесь J_m - функция Бесселя m -го порядка, а $j_{m,n}$ - n -й корень этой функции.

Спектр собственных значений энергии в цилиндрической квантовой яме определяется выражением

$$E_{n,m,p_z} = \frac{\hbar^2 j_{n,m}^2}{2m\rho_0^2} + \frac{p_z^2}{2m} \quad (2.27)$$

Обязательное существование хотя бы одного уровня энергии в потенциальной яме любой формы и глубины присуще только одномерному движению. В недостаточно глубокой трехмерной потенциальной яме нет связанных состояний – частица не может «захватиться» ямой, то есть в ней нет ни одного дискретного уровня энергии.

В прямоугольном потенциальном ящике с длинами ребер a , b , c , потенциальной энергией $U=0$ внутри этой области и $U=\infty$ вне её свободное движение частицы происходит независимо в трех направлениях. Поэтому уровни энергии даются суммами трех выражений вида (2.7):

$$\varepsilon_{n_1 n_2 n_3} = \frac{\pi^2 \hbar^2}{2m} \left(\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2} \right) \quad (2.28)$$

Волновые функции стационарных состояний

$$\Psi_{n_1 n_2 n_3} = \sqrt{\frac{8}{abc}} \sin \frac{\pi n_1 x}{a} \sin \frac{\pi n_2 y}{b} \sin \frac{\pi n_3 z}{c}, \quad (2.29)$$

где оси x , y , z направлены вдоль трех ребер ящика.

Уравнение Шредингера в случае сферически симметричного потенциала, зависящего только от r :

$$-\frac{\hbar^2}{2m} \left\{ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right] \right\} \Psi + V(r) \Psi = E \Psi \quad (2.30)$$

допускает разделение переменных

$$\Psi(r, \theta, \varphi) = \frac{1}{r} \chi_l(r) Y_{l,m}(\theta, \varphi) \quad (2.31)$$

Функция $\chi_l(r)$ удовлетворяет *радиальному волновому уравнению*

$$-\frac{\hbar^2}{2m}\left(\frac{d^2\chi_l}{dr^2}-\frac{l(l+1)}{r^2}\chi_l\right)+V(r)\chi_l=E\chi_l \quad (2.32)$$

Величина $l(l+1)$ играет роль постоянной разделения переменных r и θ, φ . Решения (2.32) существуют, если l -целое число: $l = 0, 1, 2, \dots$. Для частицы, находящейся в сферически симметричном потенциальном ящике с бесконечными стенками и радиусом R [10], радиальная волновая функция удовлетворяет уравнению

$$\chi_l'' + \left[k^2 - \frac{l(l+1)}{r^2}\right]\chi_l = 0, \quad k^2 = \frac{2mE}{\hbar^2} \quad (2.33)$$

в интервале $0 \leq r \leq R$ и равна нулю всюду вне его. После введения новой переменной $z=kr$ и замены $\chi_l = z^{1/2}\varphi(z)$ уравнение (2.33) приводится к виду

$$\varphi'' + \frac{1}{z}\varphi' + \left[1 - \frac{(l+\frac{1}{2})^2}{z^2}\right]\varphi = 0 \quad (2.34)$$

Решениями уравнения (2.34) являются функции Бесселя $J_{\pm(l+\frac{1}{2})}(z)$ и общее решение уравнения (2.33) запишется в виде

$$\chi_l(r) = \sqrt{\frac{\pi kr}{2}} \left[C_1 J_{l+\frac{1}{2}}(kr) + C_2 J_{-(l+\frac{1}{2})}(kr) \right] \quad (2.35)$$

Для выполнения условия нормировки $\int_0^\infty |\chi_l|^2 dr = 1$ необходимо положить $C_2 = 0$. Собственные функции отбираются путем наложения условия $J_{l+\frac{1}{2}}(kR) = 0$. Каждому фиксированному значению величины $l+\frac{1}{2}$ соответствует бесконечное число нулей функции Бесселя. Таким образом, получается бесконечное число значений $k_{n_r, l}$ и бесконечное число энергетических уровней

$$E_{n_r, l} = \frac{\hbar^2}{2m} k_{n_r, l}^2 \quad (2.36)$$

для каждого значения величины l ($n_r=1, 2, 3 \dots$ – радиальное квантовое число, определяющее число нулей).

Рассмотрим теперь сферически симметричную прямоугольную потенциальную яму, для которой

$$V(r) = \begin{cases} -V_0, & r < R \\ 0, & r > R \end{cases} \quad (2.37)$$

Если ввести обозначения

$$\kappa^2 = \frac{2m|E|}{\hbar^2}, k_0^2 = \frac{2mV_0}{\hbar^2}, k^2 = \frac{2m(V_0 - |E|)}{\hbar^2} = k_0^2 - \kappa^2 \quad (2.38)$$

то радиальное уравнение Шредингера для связанных состояний внутри и вне ямы соответственно будет иметь вид

$$\chi_l'' + \left[k^2 - \frac{l(l+1)}{r^2} \right] \chi_l = 0 \quad (2.39)$$

$$\chi_l'' + \left[-\kappa^2 - \frac{l(l+1)}{r^2} \right] \chi_l = 0 \quad (2.40)$$

Вводя в рассмотрение так называемые сферические функции Бесселя:

$$j_l(z) = \sqrt{\frac{\pi z}{2}} J_{l+1/2}(z) \quad (2.41)$$

и учитывая граничное условие $\chi_l(0) = 0$, уравнение (2.39) имеет решение:

$$\chi_l(r) = A j_l(kr), \quad r < R \quad (2.42)$$

Решением уравнения (2.40) убывающим при больших значениях r как $e^{-\kappa r}$ является сферическая функция Ханкеля мнимого аргумента:

$$\chi_l(r) = B h_l^{(1)}(i\kappa r), \quad r > R, \quad (2.43)$$

где A и B постоянные, определяемые из условий нормировки и непрерывности собственных функций. Условие непрерывности логарифмической производной от $r\chi_l(r)$ на границе ямы ($r = R$) дает:

$$i\kappa R \frac{h_l^{(1)'}(i\kappa R)}{h_l^{(1)}(i\kappa R)} = kR \frac{j_l'(kR)}{j_l(kR)} \quad (2.44)$$

Для $l = 0, 1, 2$ сферические функции Бесселя и Ханкеля имеют вид

$$j_0(z) = \sin z \quad h_0^{(1)}(z) = -ie^{iz}$$

$$j_1(z) = \frac{\sin z}{z} - \cos z \quad h_1^{(1)}(z) = \left(-\frac{i}{z} - 1\right) e^{iz} \quad (2.45)$$

$$j_2(z) = \left(\frac{3}{z^2} - 1\right) \sin z - \frac{3}{z} \cos z \quad h_2^{(1)}(z) = \left(-\frac{3i}{z^2} - \frac{3}{z} + i\right) e^{iz}$$

Если ввести обозначения

$$kR = x, \quad k_0 R = x_0, \quad \frac{k}{k_0} = \xi, \quad \kappa R = x_0 \sqrt{1 - \xi^2} \quad (2.46)$$

то уравнение (2.44) может быть представлено в виде

$$\operatorname{tg}(x_0 \xi) = f_l(x_0 \xi) \quad (2.47)$$

Правая часть уравнения (2.47) будет иметь вид:

$$f_0(x_0, \xi) = -\frac{\xi}{\sqrt{1 - \xi^2}} \quad (2.48)$$

$$f_1(x_0, \xi) = \frac{x_0 \xi}{1 + \frac{\xi^2}{1 - \xi^2} \left(1 + x_0 \sqrt{1 - \xi^2}\right)} \quad (2.49)$$

$$f_2(x_0, \xi) = x_0 \xi \frac{1 + x_0 \sqrt{1 - \xi^2} + \frac{1}{3} x_0^2 \xi^2 (1 - \xi^2)}{1 + x_0 \sqrt{1 - \xi^2} \left[1 - \frac{1}{3} x_0^2 \xi^2 (1 - \xi^2)\right]} \quad (2.50)$$

Уравнением (2.47) определяются, неявным образом, искомые уровни энергии. Взяты должны быть только те корни уравнения, для которых $\operatorname{tg}(x_0 \xi) < 0$, что является следствием выражения (2.48). Первый из этих уровней ($l = 0$) является в то же время самым глубоким из всех уровней энергии, т. е. соответствует нормальному состоянию частицы. Значения переменной ξ , удовлетворяющие уравнениям (2.48-2.50), проще всего находятся графическим методом. При слишком малой глубине V_0 потенциальной ямы уровни отрицательной энергии вообще отсутствуют, т. е. частица не может «удержаться» ямой. [11] Минимальная глубина ямы, при которой появляется первый отрицательный уровень

$$V_{0 \min} = \frac{\pi^2 \hbar^2}{8mR^2} \quad (2.51)$$

Эта величина тем больше, чем меньше радиус ямы R . Величина первого уровня в момент его появления определяется из $kR = \frac{\pi}{2}$ и равна нулю. По мере дальнейшего увеличения глубины ямы нормальный уровень также понижается.

В квантовых точках спектр будет чисто дискретным. Характер и степень вырождения будет определяться симметрией потенциала. Для описания состояний электронов в квантовых точках использовались различные модельные потенциалы – от простейшего потенциала трехмерной параболической ямы до трехмерной ямы, имеющей форму пирамиды.

2.5. Плотность состояний

Для выявления физических характеристик и анализа физических свойств реальных систем необходимо иметь информацию о распределении энергетических состояний. То есть необходимо знать конкретный вид функции, характеризующей число электронных состояний, приходящихся на заданный интервал энергии. Известно, что плотность состояний в n -мерном пространстве волнового вектора равна $(2\pi)^{-n}$. Пусть число энергетических состояний электронов в кристалле объемом Ω с величинами энергий в интервале от E до $E+dE$ будет равно $D(E)dE$. Тогда плотность состояний $D(E)$ записывается как

$$D(E) = \left(\frac{\Omega}{2\pi} \right)^3 \frac{d}{dE} \int d^3k, \quad (2.52)$$

где интегрирование проводится по объему k -пространства, заключенному внутри поверхности S постоянной энергии E . Бесконечно малый объем d^3k равен произведению бесконечно малого элемента указанной поверхности dS на расстояние (по нормали) между поверхностями с энергиями E и $E+dE$.

Исходя из данных определений можно вычислить плотность состояний в низкоразмерных структурах. Проведем этот расчет для 2D электронов. Будем считать, что собственными функциями являются плоские волны, нормированные на площадь образца $S = L_x \times L_y$.

$$\Psi = \frac{1}{\sqrt{S}} \exp \frac{i}{\hbar} (p_x x + p_y y) \quad (2.53)$$

а спектр электрона в приближении эффективной массы (m) изотропный и квадратичный

$$E - E_n = \frac{p^2}{2m} \quad (2.54)$$

Учитывая граничные условия для (2.53) $\Psi(x, y) = \Psi(x + L_x, y)$ и $\Psi(x, y) = \Psi(x, y + L_y)$ разрешенные значения компонент

проекций импульса: $p_x = \frac{2\pi\hbar}{L_x} n_1$, $p_y = \frac{2\pi\hbar}{L_y} n_2$, где n_1 и n_2 –

целые числа. В пространстве импульсов площадь, приходящаяся на одно состояние, будет равна:

$$p_x p_y = \frac{(2\pi\hbar)^2}{L_x L_y} = \frac{(2\pi\hbar)^2}{S} \quad (2.55)$$

Радиус окружности $p = \sqrt{2m(E - E_n)}$ ограничивает те состояния, энергия которых меньше E . Состояния с энергией $E_n > E$ лежат вне этой окружности в импульсном пространстве. Таким образом, если разделить площадь этой окружности на площадь, приходящуюся на одно состояние (2.55), и учесть двукратное спиновое вырождение, можно найти полное число состояний с энергией меньше E :

$$N(E) = 2 \frac{\pi p^2}{p_x p_y} = 2 \frac{\pi p^2 S}{(2\pi\hbar)^2} = \frac{mS(E - E_n)}{\pi\hbar^2} \quad (2.56)$$

Плотность состояний – это число состояний в единичном интервале энергий $\frac{dN(E)}{dE}$ на единицу площади S , то есть:

$$D(E) = \frac{1}{S} \frac{dN(E)}{dE} = \frac{m}{\pi\hbar^2} \mathcal{G}(E - E_n), \quad (2.57)$$

где

$$\mathcal{G}(E - E_n) = \begin{cases} 1, & E > E_n \\ 0, & E < E_n \end{cases} \quad (2.58)$$

Видно, что полученная двумерная плотность состояний при $E > E_n$ возрастает скачком (если пренебрегать разупорядочением или

уширением уровня) до определенного значения и не изменяется при дальнейшем увеличении энергии. Если при $E > E_n$ существуют другие двумерные зоны, то в плотности состояний появляются дополнительные ступени, соответствующие её увеличению.

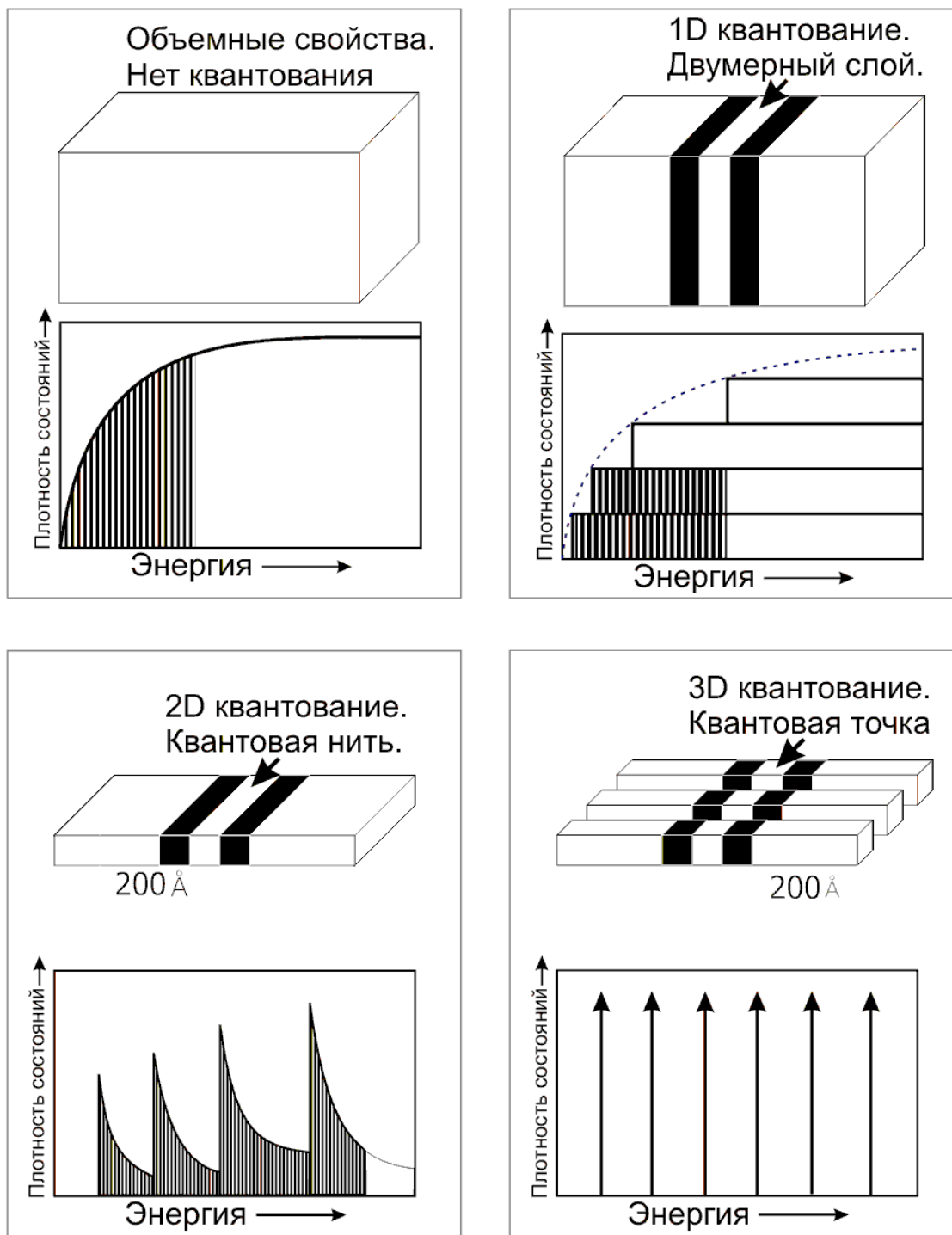


Рис. 2.3. Плотность состояний при размерном квантовании

Если заполнена только нижняя подзона, то число электронов на единицу площади при температуре абсолютного нуля равно

$$N_s = \frac{g_v m}{\pi \hbar^2} (E_F - E_n) \quad (2.59)$$

где: g_v – кратность долинного вырождения, равная числу эквивалентных энергетических зон, E_F – энергия Ферми. Поверхность Ферми в двумерной электронной системе представляет собой кривую, которую также называют линией или контуром Ферми. В простейшем случае изотропной эффективной массы контур Ферми представляет собой окружность с радиусом

$$k_F = \sqrt{\frac{2\pi N_s}{g_v}} \quad (2.60)$$

называемым волновым вектором Ферми.

Полная плотность состояний нескольких подзон – это ступенчатая функция энергии. При больших энергиях $D(E)$ стремится к плотности состояний электронов в объеме (3D), которая пропорциональна \sqrt{E} .

В квантовой нити электрон свободно движется в одном направлении. Если вновь обратиться к рассмотрению в импульсном пространстве (теперь уже одномерном), то необходимо найти полное число состояний с импульсом, не превышающим $|p_z|$. Для этого необходимо разделить $2p_z$ (с учетом возможных $\pm p_z$) на интервал в импульсном пространстве, приходящийся на одно квантовое состояние т. е. на $\frac{2\pi\hbar}{L_z}$. Вновь учитывая двукратное спиновое вырождение, имеем:

$$N(p_z) = 2 \frac{2p_z}{2\pi\hbar} L_z = \frac{2p_z L_z}{\pi\hbar} \quad (2.61)$$

По аналогии с двумерной системой можно записать выражение для энергии в одномерной подзоне:

$$E(p_z) = E_{nm} + \frac{p_z^2}{2m} \quad (2.62)$$

Выражая p_z из (2.62) и подставляя в (2.61), найдем полное число состояний с энергией меньше E :

$$N(E) = \frac{2L_z \sqrt{2m(E - E_{nm})}}{\pi \hbar} \quad (2.63)$$

Плотность состояний на единицу длины равна:

$$D(E) = \frac{dN(E)}{dE} \frac{1}{L_z} = \frac{\sqrt{2m}}{\pi \hbar \sqrt{(E - E_{nm})}} \quad (2.64)$$

Одномерная плотность состояний (2.64) имеет особенности в точках E_{nm} , являющихся нижними границами подзон, и убывает с ростом E .

В системе квантовых точек (0D системе) энергетический спектр является полностью дискретным. Плотность состояний как функция энергии в 0D представляет собой ряд δ -образных пиков. В качестве модели такой системы можно рассмотреть двумерный электронный газ, находящийся в квантующем магнитном поле. Спектр такой системы выражается формулой

$$E_n = \hbar \omega_c (n + 1/2), \quad n = 0, 1, 2, \dots \quad (2.65)$$

где $\omega_c = \frac{eH}{mc}$ - циклотронная частота. С учетом кратности вырождения магнитных уровней выражение для плотности состояний имеет следующий вид:

$$D(E) = \frac{eH}{2\pi c \hbar} \delta \left[E - \hbar \omega_c \left(n + \frac{1}{2} \right) \right] \quad (2.66)$$

Приведенные выражения для плотности состояний были получены для идеальных квантово-размерных систем, то есть без учета электронного рассеяния, приводящего к «размытию» полученных в расчетах идеальных ступеней (2D), пиков (1D) и δ -функций (0D). Исходя из соотношения неопределенностей, энергетический уровень электрона вследствие рассеяния имеет конечную ширину \hbar / τ_p , где τ_p - время релаксации импульса (время жизни электрона в одном квантовом состоянии). Для сохранения особенностей плотности состояний и дискретного характера спектра необходимо, чтобы расстояние между дискретными уровнями было больше уширения, даваемого соотношением неопределенностей:

$$\varepsilon_{n+1} - \varepsilon_n > \frac{\hbar}{\tau_p} \quad (2.67)$$

В экспериментах можно измерить подвижность, связанную с временем релаксации соотношением $\mu = e \frac{\tau_p}{m}$.

2.6. Поляризуемость и экранирование

Пусть идеальный двумерный электронный газ лежит в плоскости $z = 0$. Рассмотрим его отклик на электромагнитное поле, создаваемое слабым статическим потенциалом, медленно меняющимся в пространстве. Допустим, что плоскость помещена в однородную среду, статическая диэлектрическая проницаемость которой при $z < 0$ равна $\epsilon_{\text{диэл}}$, а при $z > 0$ равна $\epsilon_{\text{пп}}$. Дополнительный электростатический потенциал, создаваемый внешним источником, связан с плотностью заряда ρ уравнением Пуассона

$$\nabla(\epsilon \nabla \phi) = -4\pi\rho, \quad (2.68)$$

где $\rho = \rho_{\text{внеш}} + \rho_{\text{инд}}$ – суммарная плотность внешнего и индуцированного заряда, ϵ – статическая диэлектрическая проницаемость. Если потенциал изменяется в пространстве достаточно медленно, то плотность индуцированного заряда в точке \mathbf{r} на плоскости $z=0$ является функцией только локального потенциала, действующего на электроны. Фактически ситуация такая же, как и в трехмерном случае в приближении Томаса-Ферми.

$$\rho_{\text{инд}}(\mathbf{r}) = -e \left[N_s(\bar{\phi}) - N_s(0) \right] \delta(z), \quad (2.69)$$

где $\bar{\phi} = \phi(\mathbf{r}, 0)$ – величина электростатического потенциала в точке \mathbf{r} , усредненная по распределению электронов вдоль оси z . В простейшем случае это распределение описывается δ -функцией. Уравнение (2.69) представляет собой двумерный аналог приближения Томаса – Ферми.

Потенциал $\bar{\phi}$ смещает энергетические уровни электронов на величину $-e\bar{\phi}$ и изменяет расстояние между уровнем Ферми E_F и дном зоны проводимости на $e\bar{\phi}$. Если потенциал считается слабым, то после линеаризации (2.69) имеем:

$$\rho_{\text{инд}}(\mathbf{r}) = -e\bar{\phi}(\mathbf{r}) \frac{dN_s}{d\bar{\phi}} \delta(z) = -e^2 \bar{\phi}(\mathbf{r}) \frac{dN_s}{dE_F} \delta(z) \quad (2.70)$$

Уравнение Пуассона (2.68) тогда преобразуется к виду

$$\nabla(\varkappa \nabla \phi) - 2\bar{\varkappa} \bar{q}_s \bar{\phi}(r) \delta(z) = -4\pi \rho_{\text{внеш}} \quad (2.71)$$

где \bar{q}_s - обратный радиус экранирования, определяемый соотношением

$$\bar{q}_s = \frac{2\pi e^2}{\bar{\varkappa}} \frac{dN_s}{dE_F} \quad (2.72)$$

$$\bar{\varkappa} = \frac{\varkappa_{\text{дизл}} + \varkappa_{\text{пп}}}{2} \quad (2.73)$$

Линейное экранирование в однородной трехмерной системе описывается уравнением

$$\nabla^2 \phi - Q_s^2 \phi = \frac{-4\pi \rho_{\text{внеш}}}{\varkappa}, \quad (2.74)$$

где Q_s - обратный радиус экранирования Томаса-Ферми. Уравнение (2.74) по форме весьма напоминает аналогичное (2.71) для двумерной системы. Если внешний точечный заряд Ze расположен в начале координат, то решением (2.74) будет обычный экранированный кулоновский потенциал:

$$\phi = \frac{Ze}{\varkappa R} \exp(-Q_s R) \quad (2.75)$$

В случае, когда внешний заряд расположен в точке $r = 0, z = z_0 \leq 0$, уравнение (2.71) может быть решено. Для двумерного случая, при больших значениях r когда $q_s r \gg 0$, выражение для усредненного потенциала приобретает асимптотический вид [12,13]

$$\bar{\phi}(r) \approx \frac{Ze(1 + \bar{q}_s z_0)}{\varkappa \bar{q}_s^2 r^3} \quad (2.76)$$

Полученная степенная зависимость потенциала от расстояния значительно слабее экспоненциальной, характерной для объема (3D). Это и есть одно из главных качественных отличий экранирования в двумерной электронной системе от экранирования в трехмерной.

Рассмотрение более реалистичной модели электронного газа, заполняющего слой, толщина которого отлична от нуля, предполагает что учитывается поперечное распределение

электронов, например, используя функцию Фэнга-Ховарда. Радиус экранирования в этом случае выражается как:

$$q_s = \frac{2\pi e^2}{\epsilon_{\text{ин}} dE_F} \frac{dN_s}{dE_F}, \quad (2.77)$$

Сравнивая параметры экранирования для точечного заряда в квазидвумерной и идеально двумерной системах в приближении Томаса - Ферми, легко заметить, что

$$\epsilon_{\text{ин}} q_s = \bar{\epsilon} \bar{q}_s = \frac{2\pi e^2 N_s}{E_d} \quad (2.78)$$

где величину $E_d = N_s \frac{dE_F}{dN_s}$ называют энергией диффузии, входящей в обобщенное соотношение Эйнштейна $D = \frac{E_d}{e} \mu$, связывающее коэффициент диффузии и подвижность.

Литература

1. Лифшиц И. М., Косевич А. М. К теории магнитной восприимчивости тонких слоев металлов при низких температурах // ДАН СССР. 1953. Т. 91. № 4. С.795-798.
2. Огрин Ю. Д., Луцкий В. Н., Елинсон М. И. О наблюдении квантовых размерных эффектов в тонких слоях висмута // Письма в ЖЭТФ. 1966. Т. 3. ВЫП. 3. С. 114 - 118.
3. Shrieffer J. R. Mobility in inversion layers: theory an experiment // Semiconductor Surface Physics Ed. By R. H. Kingston. Philadelphia: Univ. Penn. Press, 1957. P.55 – 69.
4. Стерн Ф. Квантовые свойства поверхностных слоев пространственного заряда // Новое в исследовании поверхности твердого тела. М.: Мир, 1977. С. 243 – 279.
5. Андо Т., Фаулер А., Стерн Ф. Электронные свойств двумерных систем. М.: Мир, 1985. 416.
6. Добровольский В. Н., Литовченко В. Г. Перенос электронов и дырок у поверхности полупроводников. Киев.: Наукова думка, 1985. 192.
7. Демиховский В. Я., Вугальтер Г. А. Физика квантовых низкоразмерных структур. М.: Логос, 2000. 248.

8. *Галицкий В. М., Карнаков Б. М., Коган В. И.* Задачи по квантовой механике. М.: Наука, 1981. 152. (См. задачу 2.15)
9. *Fang F. F., Howard W. E.* Negative field-effect mobility on (100) Si surfaces // *Phys. Rev. Lett.* 1966, V.16. P.797-799.
10. *Флюгге З.* Задачи по квантовой механике. Том 1. М.: Мир, 1974. 344. (См. задачи 62-63)
11. *Ландау Л. Д., Лифшиц Е. М.* Квантовая механика. Нерелятивистская теория. М.: Наука, 1989. 673.
12. *Stern F.* Polarizability of a two-dimensional electron gas // *Phys. Rev. Lett.* 1967, V.18. P. 546-548.
13. *Рытова Н. С.* Экранированный потенциал точечного заряда в тонкой пленке // *Вестник МГУ (физика, астрономия).* 1967. №3. С. 30-37.

3. РЕЗОНАНСНОЕ ТУННЕЛИРОВАНИЕ

В последние 40 лет компьютеры и их возможности растут гораздо быстрее, чем уменьшаются размеры их основного функционального элемента – транзистора. Однако законы квантовой механики, ограничения, связанные с используемыми материалами и технологией, делают невозможным бесконечное уменьшение размеров обычных транзисторов. Уже сегодня понятно, что известные закономерности для объемных полупроводниковых элементов частично не будут выполняться при уменьшении размеров элементов до 100 нанометров и ниже. Альтернативой прямому увеличению плотности элементов микросхем может служить использование в микроэлектронных устройствах новых физических эффектов и закономерностей, проявляющихся при уменьшении характерных размеров этих же элементов (Рис. 3.1). Все приборы наноэлектроники можно разделить на две группы:

- твердотельные квантовые приборы, принципы работы которых основаны на квантовых эффектах и явлениях;
- приборы молекулярной электроники.

Когда мы говорим о наноэлектронике и ее элементной базе, то необходимо всегда помнить о том, что речь идет о размерах менее 0,1 мкм. При этом поведение электронов в полупроводнике не может уже рассматриваться как простое движение частиц - носителей заряда. Необходимо учитывать их волновую природу. Такое поведение обычно называется «квантовым режимом» и требует квантово-механического подхода к описанию физических явлений в функциональных элементах наноэлектронных устройств.

В конечном итоге небольшие «островки» полупроводника или металла, в которых наблюдается размерное квантование носителей, являются той особенностью, которая лежит в основе работы всех приборов наноэлектроники. Островки материала в наноэлектронных приборах выполняют роль аналогичную роли канала, формирующегося под затвором при приложении напряжения в обычном полевом транзисторе.

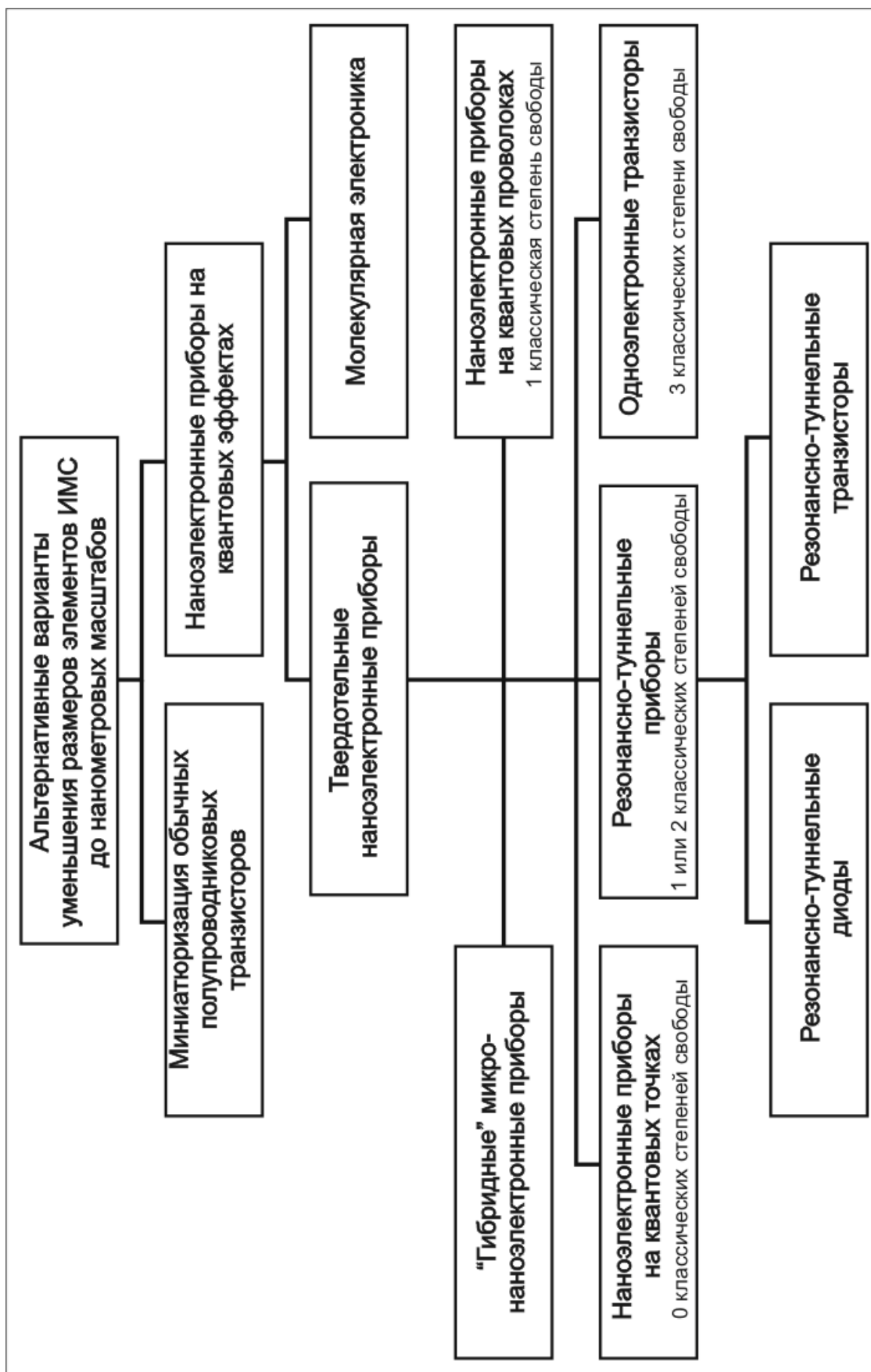


Рис. 3.1. Классификация нанoeлектронных приборов.

Состав, форма и размер этих островков определяют возможности создания совершенно различных приборов с соответствующими характеристиками. Контроль именно этих характеристик позволяет разработчикам наноприборов использовать различные возможности, создаваемые квантовыми режимами, и контролировать перенос носителей заряда через активные области материала.

В основу работы квантовых приборов положены два эффекта – туннелирование и размерное квантование. Примером приборной реализации этих эффектов являются резонансные туннельные приборы. Для понимания квантовых эффектов и их роли в работе элементов наноприборов необходимо помнить, что электрон движется по кристаллу полупроводника как Блоховская волна. Важным свойством является когерентность волнового пакета или волновой функции. В резонансных туннельных приборах (РТП) размеры квантовой ямы меньше длины когерентности Блоховской волны. Структура типа «резонатор» создает возможность интерференции прямой и обратной (отраженной от границы раздела) волн, так, что усиливается туннельный ток. Аналогом этого эффекта может служить, например, усиление звука в трубах органа. Фактически РТП – это электрический аналог резонатора Фабри-Перо. Туннельный ток определяется качеством границ структуры и электрон-фононным рассеянием. Современные технологии, такие как молекулярно-пучковая эпитаксия, позволяют получать структуры с высоким качеством границ раздела.

Хорошо известно, что базовым материалом современной микроэлектроники является кремний (Si) – элемент IV группы периодической системы. Однако большинство современных приборов нанозлектроники изготавливается на основе материалов III и V групп, например AlAs и GaAs. Несомненным преимуществом этих материалов является высокая подвижность носителей и возможность формирования границ раздела высокого качества. Следует однако отметить, что в последнее время все шире начинается использование в нанозлектронике и элементов IV группы: Si и Ge.

3.1. Двухбарьерная структура

Для металлов длина волны де Бройля на порядок меньше, чем в полупроводниках ($\sim 200 - 300$ ангстрем). Именно при таких толщинах пленок следует ожидать проявлений волновых свойств, в частности туннелирования. Кроме того, подобные оценки означают, что эти эффекты легче осуществить в полупроводниках, чем в металлах. Это связано с тем, что возможность создания тонкой пленки с необходимыми для квантования размерами и сохранением монокристаллической структуры тем легче, чем больше ее толщина.

Наибольший интерес для создания новых высокочастотных приборов составляют не одиночные барьеры, а структуры, имеющие двойной потенциальный барьер [1]. Высокий уровень развития современной технологии позволяет создавать слоистые структуры [2,3].

3.1.1. Коэффициенты прохождения и отражения

Пусть на двухбарьерную структуру в положительном направлении оси z падают электроны с энергией E и импульсом $\hbar k = \sqrt{2mE}$ (Рис. 3.2.). Найдем амплитуды отражения R и прохождения D через структуру. Рассмотрим многократные отражения волн от границ барьеров [1].

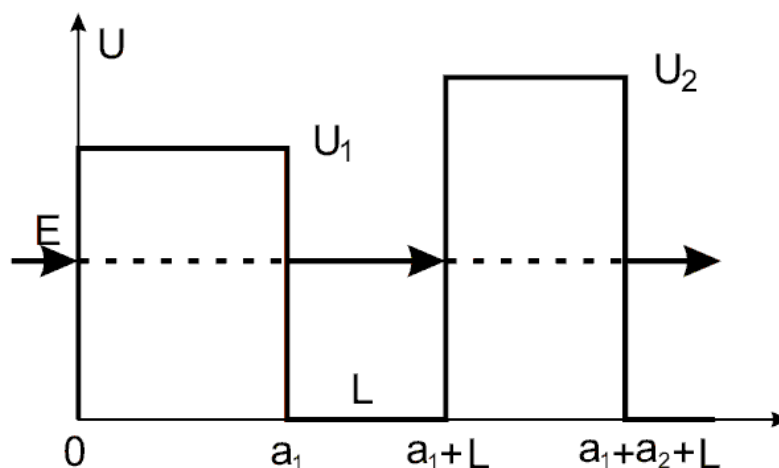


Рис. 3.2. Двухбарьерная квантовая структура

Считая известными амплитуды отражения $R_{1,2}$ и прохождения $D_{1,2}$ для первого и второго барьеров и учитывая фазовый набег волны kL в яме ширины L , можем записать выражения для амплитуд отражения R и прохождения D .

$$R = R_1 + D_1 e^{ikL} R_2 e^{ikL} D_1 + D_1 e^{ikL} R_2 e^{ikL} R_1 e^{ikL} R_2 e^{ikL} D_1 + \dots = R_1 + \frac{D_1^2 R_2 e^{2ikL}}{1 - R_1 R_2 e^{2ikL}} \quad (3.1)$$

$$D = D_1 e^{ikL} D_2 + D_1 e^{ikL} R_2 e^{ikL} R_1 e^{ikL} D_2 + \dots = \frac{D_1 D_2 e^{ikL}}{1 - R_1 R_2 e^{2ikL}} \quad (3.2)$$

Выражения для R и D представляют собой суммы геометрических прогрессий. Обозначим $R_{1,2} = |R_{1,2}| e^{i\varphi_{1,2}}$.

Коэффициент прохождения $|D|^2$ может быть представлен как

$$|D|^2 = \frac{|D_1 D_2|^2}{1 + |R_1 R_2|^2 - 2|R_1 R_2| \cos(2kL + \varphi_1 + \varphi_2)} \quad (3.3)$$

Если для волнового вектора k выполняется условие

$$k_n L + \frac{(\varphi_1 + \varphi_2)}{2} = \pi n, \quad \text{где } n = 0, 1, 2, \dots \quad (3.4)$$

то для коэффициента прохождения получим выражение

$$|D|^2 = \frac{|D_1|^2 |D_2|^2}{(1 - |R_1 R_2|)^2} \quad (3.5)$$

Если считать барьеры достаточно толстыми, то коэффициенты $D_{1,2}$ будут экспоненциально малы (пропорциональны $e^{-\alpha a_1}, e^{-\alpha a_2}$, где $\hbar \alpha_{1,2} = \sqrt{2m(U_{1,2} - E)}$), а коэффициенты отражения близки по модулю к единице:

$$|R_{1,2}| = \sqrt{1 - |D_{1,2}|^2} \approx 1 - \frac{|D_{1,2}|^2}{2} \quad (3.6)$$

Подставляя (3.6) в (3.5), получим

$$|D(E = E_n)|^2 \approx \frac{4|D_1|^2 |D_2|^2}{(|D_1|^2 + |D_2|^2)^2}, \quad (3.7)$$

где $E_n = \frac{\hbar^2 k_n^2}{2m}$. Если барьеры одинаковые, т. е. $|D_1| = |D_2|$, получим $|D(E = E_n)|^2 \approx 1$. Таким образом, при указанных условиях наблюдается явление **резонансного туннелирования**. Поскольку число частиц при прохождении через потенциальный барьер сохраняется, то получим $|R(E = E_n)| = 0$.

Несмотря на то, что барьеры выбраны нами достаточно толстыми, и проницаемости их малы $|D_{1,2}|^2 \ll 1$, при некоторых значениях энергии эти барьеры для частицы являются «прозрачными». Картина аналогична интерферометру Фабри-Перо: из-за интерференции электронных волн при отражении от барьеров снаружи остаются лишь падающая и прошедшая электронные волны. Обратная отраженная волна при этом полностью гасится.

Положив амплитуду падающей волны равной единице, можем записать выражение для волновой функции в яме между барьерами:

$$\Psi(a_1 < z < a_1 + L) = Ae^{ik(z-a_1)} + Be^{ik(z-L-a_1)} \quad (3.8)$$

где

$$A = D_1 + D_1 e^{ikL} R_2 e^{ikL} R_1 + \dots = \frac{D_1}{1 - R_1 R_2 e^{2ikL}} \quad (3.9)$$

$$B = D_1 e^{ikL} R_2 + D_1 e^{ikL} R_2 e^{ikL} R_1 e^{ikL} R_2 + \dots = \frac{D_1 R_2 e^{ikL}}{1 - R_1 R_2 e^{2ikL}} \quad (3.10)$$

Из формул (3.9) и (3.10) при условии, что $k=k_n$ получим

$$A \approx B \approx \frac{|D_1|}{1 - |R_1||R_2|} \approx \frac{2|D_1|}{|D_1|^2 + |D_2|^2} \quad (3.11)$$

Полагая $|D_1| = |D_2|$, получим, что $|A| \approx |B| \approx \frac{1}{|D_1|} \gg 1$. Таким

образом, амплитуда волновой функции внутри ямы (A и B) значительно превышает амплитуду волны вне двухбарьерной структуры (которую мы положили равной единице). Электрон задерживается в яме на большой промежуток времени и

многократно отражается от барьеров. В результате многократно возрастает вероятность туннелирования электрона из ямы.

3.1.2. Квазистационарные состояния в яме между барьерами

Для частицы, находящейся в яме между барьерами, решение уравнение Шредингера

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial z^2} = E \Psi, \quad a_1 < z < a_1 + L \quad (3.12)$$

имеет вид

$$\Psi(z) = C_1 e^{ik(z-a_1)} + C_2 e^{-ik(z-L-a_1)}, \quad (3.13)$$

где C_1 и C_2 – произвольные постоянные, $k = \frac{\sqrt{2mE}}{\hbar}$. Поскольку

решением уравнения (3.11) является функция, описывающая две волны, бегущие в противоположных направлениях, а амплитуды $R_{1,2}$ считаются известными, то для коэффициентов C_1 и C_2 можно записать следующие выражения:

$$C_1 = R_1 e^{ikL} C_2, \quad C_2 = R_2 e^{ikL} C_1 \quad (3.14)$$

Ненулевые значения C_1 и C_2 существуют при условии

$$R_1 R_2 e^{2ikL} = 1 \quad (3.15)$$

Поскольку $|R_1 R_2| < 1$, уравнение (3.14) не имеет действительных корней и в яме не существует стационарных состояний. При $|R_1 R_2| \approx 1$ корни уравнения (3.14) близки к действительным, и можно говорить о квазистационарных состояниях в яме, характеризующихся комплексными значениями энергии:

$$\tilde{E}_n = \text{Re } \tilde{E}_n + i \text{Im } \tilde{E}_n, \quad |\text{Im } \tilde{E}_n| \ll \text{Re } \tilde{E}_n \quad (3.16)$$

или для комплексного значения волнового числа

$$\tilde{k}_n = \sqrt{\frac{2m\tilde{E}_n}{\hbar^2}} = \text{Re } \tilde{k}_n + i \text{Im } \tilde{k}_n, \quad |\text{Im } \tilde{k}_n| \ll \text{Re } \tilde{k}_n \quad (3.17)$$

Учитывая, что

$$\text{Re } \tilde{E}_n \approx \frac{\hbar^2}{2m} (\text{Re } \tilde{k}_n)^2, \quad \text{Im } \tilde{E}_n \approx \frac{\hbar^2}{m} \text{Re } \tilde{k}_n \text{Im } \tilde{k}_n \quad (3.18)$$

Из уравнения (3.15), используя обозначение $R_{1,2} = |R_{1,2}|e^{i\varphi_{1,2}}$ и с учетом (3.17) и (3.6) можно получить

$$\operatorname{Re} \tilde{k}_n L + \frac{1}{2}(\varphi_1 + \varphi_2) \Big|_{k=\operatorname{Re} \tilde{k}_n} \approx \pi n, \quad n = 0, 1, 2, \dots, \quad (3.19)$$

Очевидно, что уравнения (3.17) и (3.4) совпадают. Это позволяет сделать вывод, что $\operatorname{Re} \tilde{E}_n = E_n$, $\operatorname{Re} \tilde{k}_n = k_n$, т.е. резонансное туннелирование идет через квазистационарные уровни электрона в яме.

Если ввести время релаксации квазистационарного состояния (результат соударений о стенки ямы для электронов) с номером n как $\operatorname{Im} \tilde{E}_n = -\frac{\hbar}{\tau_n}$, то оно может быть определено из выражения

$$\tau_n \approx \frac{1}{\nu_n} \frac{4}{|D_1|^2 + |D_2|^2} \Big|_{k=\operatorname{Re} \tilde{k}_n}, \quad (3.20)$$

где $\nu_n = \hbar \frac{\operatorname{Re} \tilde{k}_n}{m \tilde{L}_n}$ - частота ударов о барьеры. При этом $\hbar \frac{\operatorname{Re} \tilde{k}_n}{m}$ может быть интерпретирована как скорость классического движения электрона на уровне $\operatorname{Re} \tilde{E}_n$, а величина

$\tilde{L}_n = L + \frac{1}{2} \frac{d(\varphi_1 + \varphi_2)}{dk} \Big|_{k=\operatorname{Re} \tilde{k}_n}$ имеет смысл расстояния между

плоскостями эффективного отражения электронной волны.

Рассматривая реальные структуры, в которых возможно наблюдение эффекта резонансного туннелирования (например для GaAs/Al_{0,3}Ga_{0,7}As), при $U_1=U_2=0,2$ эВ, $m=0,067m_0$, $L=50$ Å, $a_1=a_2=70$ Å получим $\tau_0 \approx 10^{-10}$ с, а при $a_1=a_2=50$ Å, $\tau_0 \approx 8 \cdot 10^{-12}$ с.

Процессы рассеяния на примесях, фононах, дефектах на границах барьеров нарушают когерентность электронных волн, интерференционную картину и в конечном итоге приводят к уширению энергетических уровней электронов, т.е. к увеличению

$|\operatorname{Im} \tilde{E}_n|$. Если обозначить вклад процессов рассеяния в \tilde{E}_n как $-\frac{\hbar}{\tau_p}$,

то можно констатировать, что время релаксации возрастает с

уменьшением концентрации примесей, дефектов и с понижением температуры. Для чистого GaAs температурная зависимость дает следующие значения:

$$T = 77\text{K} \quad \tau_p \approx 10^{-12} \text{ с}$$

$$T = 200\text{K} \quad \tau_p \approx 10^{-12} \text{ с}$$

$$T = 300\text{K} \quad \tau_p \approx 10^{-13} \text{ с}$$

Значение амплитуды прохождения (3.2) при значениях k близких к $\text{Re}\tilde{k}_n$, т.е. при значениях энергии близких к E_n выражается как

$$|D|^2 = \frac{4|D_1 D_2|^2}{(|D_1|^2 + |D_2|^2)^2} \bigg|_{E=E_n} \frac{\frac{\hbar^2}{\tau_n^2}}{(E - E_n)^2 + \frac{\hbar^2}{\tau_n^2}} \quad (3.21)$$

Если учесть рассеяние электронов, то даже при $|D_1|^2 = |D_2|^2$ максимальное значение $|D|^2 \neq 1$, однако приближается к единице по мере увеличения отношения τ_p / τ_n .

Литература.

1. Демиховский В. Я., Вугальтер Г. А. Физика квантовых низкоразмерных структур. М.: Логос, 2000. 248.
2. Тагер А. С. // Электронная техника. Сер. 1. Электроника СВЧ. 1987. Вып. 9(403), 21.
3. Тагер А. С. // Электронная техника. Сер. 1. Электроника СВЧ. 1988. Вып. 2(406), 17.

4. ОДНОЭЛЕКТРОНИКА

Увеличение степени интеграции в микроэлектронике, сопровождающееся миниатюризацией твердотельных приборов и устройств, приводит к необходимости ответа на вопрос: «Насколько миниатюрными мы можем изготовить транзисторы, резисторы, диоды и т.п., не затрагивая основных физических принципов, лежащих в основе их работы?». Можно поставить вопрос и иначе: «Насколько малы должны быть размеры этих элементов, чтобы коренным образом изменились их рабочие свойства и характеристики?» Под новыми свойствами обычно понимают такие, в которых проявляются законы квантовой механики или квантования заряда в единицах (e) - элементарного заряда электрона. Такого рода эффекты наиболее важны для систем небольших размеров, порядка атомных. О каких же системах или даже устройствах может идти речь? Любой атомный кластер, имеющий истоковый и стоковый контакты, и чьи характеристики могут регулироваться с помощью электрода, выполняющего функцию «затвора». На практике это могут быть молекулы, металлические островки, или электронные приборы, изготовленные по современной микроэлектронной технологии. Это в свою очередь означает, что физические характеристики таких разных структур могут быть объяснены в рамках единого подхода, связанного с изучением явлений переноса в системах пониженной размерности.

Одним из типов такого искусственно изготовленного устройства является квантовая точка. Обычно квантовыми точками называют пространственно - ограниченные области полупроводниковых материалов с размерами порядка 100 нанометров. С конца 1980-х годов исследования в этой области активно развиваются. Квантовые точки оказались системами очень эффективными для изучения большого количества физических явлений: искусственно созданные атомы, квантовый хаос, квантовый эффект Холла, и т.д. В экспериментах, связанных с исследованиями явлений электронного переноса, было показано, что аналогичный подход может быть применен

при изучении молекулярных систем и небольших металлических частиц (зерен).

Хотя под термином «квантовая точка» обычно понимается образование очень малого объема, следует помнить, что полупроводниковая квантовая точка – это тем не менее приблизительно 10^6 атомов с соответствующим числом электронов. Число свободных электронов в точке невелико – от одной до нескольких сотен. Длина волны де Бройля этих электронов сравнима с размерами квантовой точки и энергетический спектр имеет дискретный характер, типичный для квантовых систем. У квантовой точки есть характеристика, аналогичная энергии ионизации в атоме, называемая зарядовой энергией. Это по существу энергия, необходимая для того, чтобы добавить или удалить один электрон из точки. Из-за этой аналогии квантовые точки иногда называют искусственными атомами. В этой связи особенно интересными являются исследования оптических свойств точек, а именно взаимодействие со световыми волнами, спектральные характеристики процессов поглощения и излучения. Кроме того, точки являются уникальным модельным объектом, позволяющим изучать явления токопереноса через «искусственно созданный атом».

4.1. КВАНТОВОЕ ТУННЕЛИРОВАНИЕ

4.1.1. Теория кулоновской блокады

Насколько мал может быть проводник и какова должна быть температура, чтобы изменение числа электронов на единицу привело к эффектам, которые можно измерить в эксперименте? Другими словами, возникает вопрос о величине кулоновского взаимодействия и его влиянии на состояние системы. Для того, чтобы ответить на этот вопрос, рассмотрим электронные свойства небольшого проводника, изображенного на рис. 4.1(а) и представляющего собой своеобразный трехполюсник. Перенос частиц осуществляется только через две границы, как показано на рисунке. Два электрода трехполюсника (исток и сток) соединены в цепь, включающую приборы для измерения тока и/или напряжения. Третий электрод обеспечивает

электростатическую или емкостную связь и может быть использован как затвор

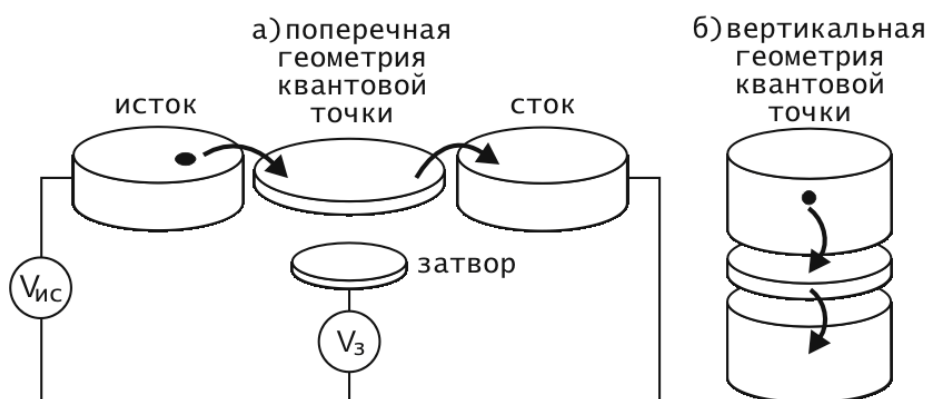


Рис. 4.1. Схематическое изображение квантовой точки, в форме диска, соединенного с истоком и стоком туннельными контактами и емкостным затвором. (а) боковая геометрия и (б) вертикальная геометрия

Если сток и исток не соединены, то есть цепь не замкнута, тогда центральная часть проводника становится своеобразным островком для электронов. Количество электронов на островке определяется целым числом N , то есть заряд квантован и равен eN . Если допустить возможность туннелирования электронов в системе «исток – квантовая точка – сток», то число электронов N регулируется самостоятельно, так чтобы энергия замкнутой системы стала минимальной.

Теория одноэлектронного туннелирования была предложена основоположником науки, получившей название «одноэлектроника», К. К. Лихаревым [1]. Среди наиболее часто цитируемых обзоров на эту тему следует назвать [2, 3].

Рассмотрим систему из одного туннельного перехода, между двумя металлическими контактами. Если емкость такой системы C , то энергия такой системы, представляющей собой конденсатор с зарядом Q на обкладках, выражается как

$$E = \frac{Q^2}{2C} \quad (4.1)$$

Поскольку минимальный квант заряда равен заряду электрона, то минимальная величина изменения энергии ΔE составляет

$$\Delta E = \frac{e^2}{2C} \quad (4.2)$$

Для наблюдения эффектов необходимо, чтобы минимальное изменение энергии превышало температурные флуктуации

$$\Delta E \gg kT, \quad (4.3)$$

где k - постоянная Больцмана, T - температура. Необходимо также, чтобы ΔE превышало квантовые флуктуации

$$\Delta E \gg \frac{hG}{C}, \quad (4.4)$$

где G - это максимальная величина из двух возможных: G_s - проводимости, шунтирующей переход, и G_i - проводимости туннельного перехода. Если $R_Q = h/e^2 = 25,813 \text{ кОм}$ - квантовое сопротивление, то исходя из (4.4) справедливо неравенство

$$G \gg R_Q^{-1} \quad (4.5)$$

Типичное время зарядки или разрядки металлического островка определяется произведением емкости и сопротивления, т.е. $\Delta t = RC$. С другой стороны из соотношения неопределенностей

Гейзенберга $\Delta E \cdot \Delta t = \left(\frac{e^2}{C} \right) \cdot RC \geq h$ следует, что R должно быть

значительно больше квантового сопротивления R_Q . Таким образом, (4.3) и (4.5) являются условиями наблюдения одноэлектронного туннелирования.

Начальный заряд Q_0 на туннельном переходе может быть не только отличен от нуля, но и может принимать значения не кратные заряду электрона. Все дело в том, что Q_0 может быть создан *поляризацией* близлежащих электродов, заряженных примесей и поэтому иметь *любое* значение. Тогда в уравнении

(4.1) $Q = Q_0 - e$. Следовательно, если $-\frac{e}{2} < Q < \frac{e}{2}$, то добавление

или вычитание целого числа электронов увеличивает энергию (4.1), что противоречит высказанному ранее соображению о минимуме энергии. Из рис.4.2 видно, что если заряд меньше

величины $e/2$, то добавление или вычитание одного электрона приводит к увеличению общей энергии. Если заряд больше величины $e/2$, то выгодным становится туннелирование электрона через диэлектрик. Поскольку напряжение на конденсаторе $V = Q \cdot C$, то при напряжениях $-\frac{e}{2C} < V < \frac{e}{2C}$ ток через туннельный переход протекать не будет. Эффект отсутствия тока при приложении напряжения в указанном интервале получил название эффекта кулоновской блокады.

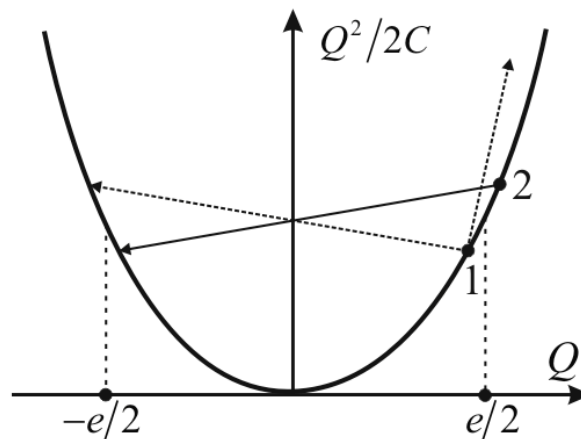


Рис. 4.2. Зависимость зарядовой энергии перехода от заряда. Стрелками показаны добавления или вычитания одного электрона для случаев кулоновской блокады (1) и туннелирования (2). [4]

Явление отсутствия тока при приложении напряжения к туннельному переходу из-за невозможности туннелирования электронов вследствие их кулоновского отталкивания называется **кулоновской блокадой**. Минимальное напряжение

$$V_{CB} = \frac{e}{2C}, \quad (4.6)$$

которое необходимо приложить к туннельному переходу для преодоления кулоновской блокады, называется напряжением отсечки или напряжением кулоновской блокады.

Протекание тока через единичный туннельный переход происходит следующим образом. Поскольку ток является

непрерывной величиной, то заряд на одной стороне перехода постепенно накапливается, пока не достигнет величины $e/2$. Далее происходит туннелирование одного электрона через переход, и процесс повторяется. Аналогия этого процесса с неплотно закрытым краном, из которого по капле падает вода, была предложена К. К. Лихаревым [5]. Заряд одного электрона e , накапливающийся при токе I за время t , выражается как $e = I \cdot t$. После накопления заряд туннелирует и процесс повторяется периодически с частотой

$$f = I/e \quad (4.7)$$

Такие осцилляции получили название одноэлектронных туннельных (SET) осцилляций.

Условие (4.3) накладывает особенно жесткие ограничения на конструкции одноэлектронных приборов. Из (4.2) и (4.3) получается, что значение емкости, необходимое для наблюдения кулоновской блокады при данной температуре T , должно удовлетворять условию:

$$C \gg \frac{e^2}{2kT} \quad (4.8)$$

Численные оценки для различных температур дают следующие значения: $C \gg 2 \times 10^{-16} \Phi$ (для $T = 4,2 \text{ K}$), $C \gg 10^{-17} \Phi$ (для $T = 77 \text{ K}$), $C \gg 3 \times 10^{-18} \Phi$ (для $T = 300 \text{ K}$).

На рис. 4.3(а) приведена эквивалентная схема рассмотренной системы, в которой к переходу приложено напряжение V . Туннельный переход характеризуется емкостью C и сопротивлением R . Емкость подводящих контактов C' . Поскольку емкости C' и C соединены параллельно, то емкость C' будет шунтировать емкость C в том случае, если $C' > C$.

В реальных приборах $C' \geq 10^{-15} \Phi$, то есть $C' \gg C$ даже для гелиевых температур, то есть наблюдение одноэлектронного туннелирования в системе с одним переходом при нынешнем уровне развития технологии является невыполнимой задачей.

Эквивалентная схема конструкции с двумя переходами (рис. 4.1), позволяющей решить задачу экспериментального наблюдения одноэлектронного туннелирования, приведена на рис. 4.3(б).

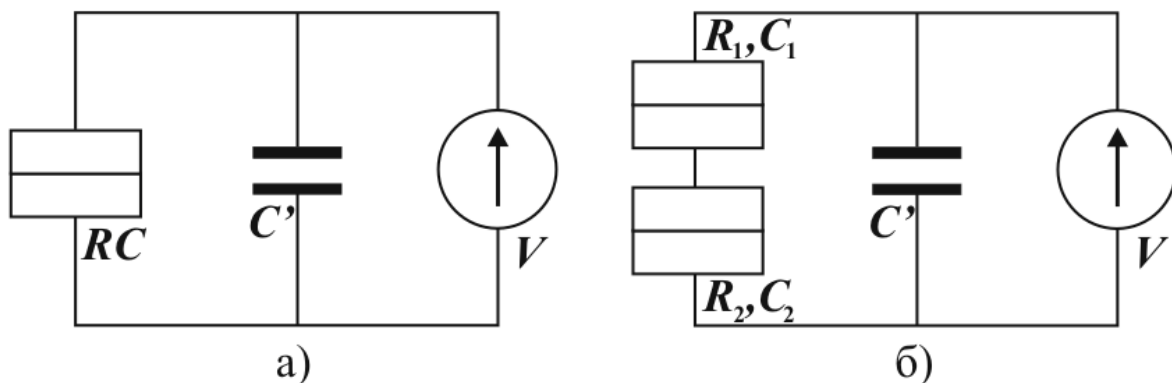


Рис. 4.3. Эквивалентные схемы: а) один туннельный переход; б) два последовательно соединенных туннельных перехода

Энергия такой системы выражается как

$$E = \frac{Q_1^2}{2C_1} + \frac{Q_2^2}{2C_2} \quad (4.9)$$

Емкость контактов не является шунтирующей для каждого туннельного перехода (имеющего индексы 1 и 2). Поскольку центральная частица (квантовая точка) отделена от контактов туннельными переходами, заряд, находящийся на ней, выражается как $Q = Q_1 = Q_2$. Откуда (4.9) приобретает вид

$$E = \frac{Q^2}{2C_\Sigma} \quad (4.10)$$

Формула (4.10) – это полный аналог (4.1), где вместо емкости C фигурирует $C_\Sigma = C_1 + C_2$. Суммарная емкость – это емкость двух переходов, включенных параллельно, если смотреть с частицы. Расширенная эквивалентная схема для структуры с двумя туннельными контактами и затвором (рис. 4.1 (а)) представлена на рис. 4.4.

Полная емкость складывается из суммы емкостей левого (C_l) и правого (C_r) переходов и затворной емкости (C_g). Фактически справедливыми остаются формулы (4.2), (4.4), (4.8), если заменить в них C на C_Σ . Кроме того, в (4.3) и (4.4) надо заменить G на $\max(G_1, G_2)$.

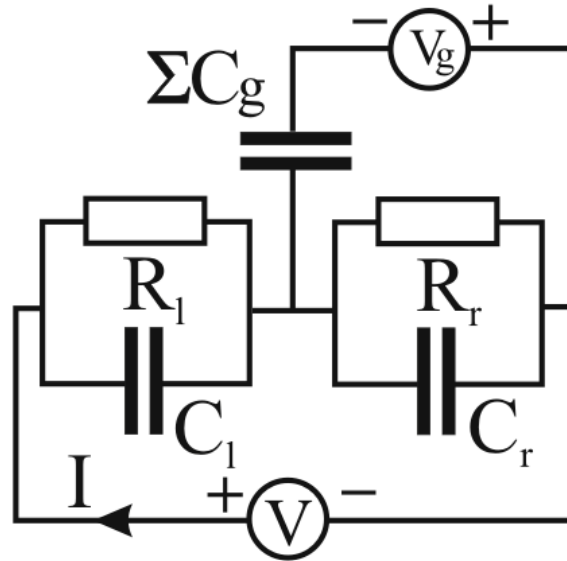


Рис. 4.4. Эквивалентная схема двухбарьерной структуры, на которой туннельные переходы представлены в виде резистора и конденсатора, соединенных параллельно. Различные возможные типы затворов представлены одной эквивалентной емкостью ΣC_g . Зарядовая энергия такой системы: $e^2 / (C_l + C_r + \Sigma C_g)$ [2]

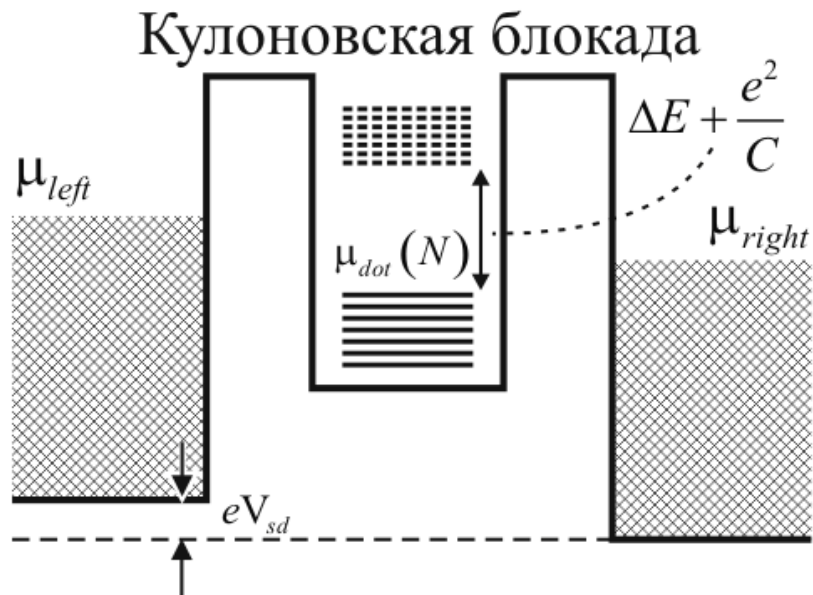


Рис. 4.5. Распределение потенциала в системе «исток – туннельный контакт – квантовая точка – туннельный контакт – сток»

На рис.4.5 показано распределение потенциала в системе «исток – туннельный контакт - квантовая точка – туннельный контакт – сток». Разность потенциалов между стоком и истоком $V_{sd} = (\mu_{left} - \mu_{right})/e$. Состояния в левой и правой частях системы (на истоке и стоке) заполнены до уровней, соответствующих электрохимическим потенциалам $(\mu_{left}, \mu_{right})$. Дискретные 0D (нульмерные), состояния в квантовой точке, заполнены N электронами до уровня $\mu_{dot}(N)$. При нулевой температуре перенос носителей происходит по следующим правилам: ток не равен нулю, если есть незаполненные состояния в интервале энергий между μ_{left} и μ_{right} . Число возможных состояний в квантовой точке определяется из расчетов электрохимического потенциала $\mu_{dot}(N)$. По определению – это минимальная энергия, необходимая для добавления N - го электрона в квантовую точку, т.е. $\mu_{dot} \equiv U(N) - U(N-1)$, где $U(N)$ - полная энергия основного состояния для N электронов в точке при нулевой температуре.

Вычисление $U(N)$ из первых принципов - задача сложная. Упрощенная модель приводит в линейном приближении (т. е. $V_{sd} \ll \Delta E/e$, $V_{sd} \ll e/C$) к выражению для $\mu_{dot}(N)$ в квантовой точке

$$\mu_{dot}(N) = E_N + \frac{(N - N_0 - 1/2)e^2}{C} - e \frac{C_g}{C} V_g \quad (4.11)$$

Электрохимический потенциал определяется как $\mu_{dot}(N) = \mu_{ch}(N) + e\phi_N$, т. е. как сумма химического потенциала $\mu_{ch}(N) = E_N$ и электростатического потенциала $e\phi_N$. Состояние для одной частицы E_N для N - го электрона отсчитывается от дна зоны проводимости и зависит от вида конкретного потенциала, задающего квантово - размерные ограничения. Электростатический потенциал ϕ_N содержит дискретную и непрерывную части. В выражении (4.11) N - это число электронов при напряжении на затворе V_g , N_0 - их число при

нулевом напряжении на затворе. Непрерывная часть φ_N пропорциональна напряжению на затворе. При фиксированном V_g , число электронов N в точке выражается большим целым числом, для которого выполняется условие $\mu_{dot}(N) < \mu_{left}(N) \cong \mu_{right}(N)$. Изменение числа электронов на единицу приводит к изменению электрохимического потенциала

$$\mu_{dot}(N+1) - \mu_{dot}(N) = \Delta E + \frac{e^2}{C} \quad (4.12)$$

4.1.2. Кулоновская лестница

Рассмотрим систему с двумя несимметричными барьерами (рис. 4.6) Темп туннелирования через один из переходов можно записать как

$$\Gamma_1 = \frac{\delta E_1}{e^2 R_1}, \quad (4.13)$$

где $\delta E_1 = eV_1 - \frac{e^2}{2C_1}$ - изменение энергии на первом переходе при падении на нем напряжения V_1 большего напряжения кулоновской блокады. Следовательно, (4.13) может быть представлено как

$$\Gamma_1 = \frac{V_1}{eR_1} - \frac{1}{2R_1C_1} \quad (4.14)$$

Совершенно аналогичное выражение Γ_2 может быть записано и для другого перехода. Из условия (4.14) видно, что темп туннелирования зависит от емкостей и сопротивлений переходов. Если система симметричная, т.е. если параметры переходов одинаковые, то будет наблюдаться плавное нарастание тока на вольт - амперной характеристике, т.к. количество электронов, пришедших и ушедших на «кулоновский остров», будет одинаковым. При несимметричности переходов на островке будет существовать заряд из N электронов. Увеличивая напряжение, мы будем наблюдать резкое увеличение тока. Это будет происходить тогда, когда напряжение будет достаточным

для увеличения числа электронов до значения $N+1$. Работает переход с высоким темпом туннелирования.

Кулоновская лестница

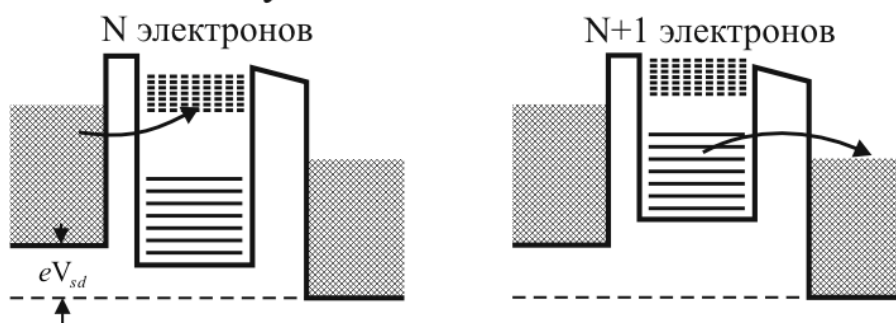


Рис. 4.6. Энергетическая диаграмма, иллюстрирующая появление кулоновской лестницы. При увеличении напряжения между истоком и стоком незаполненные состояния над кулоновской щелью в зоне проводимости могут заполняться.

Дальнейшее увеличение напряжения приводит к медленному росту тока и обуславливается переходом с низким темпом туннелирования. И это будет происходить до тех пор, пока напряжение не достигнет значения, соответствующего переходу $(N+2)$ – го электрона на островок. Ток снова начинает расти быстрее, и ситуация повторяется. В результате непрерывного протекания тока через два перехода, на островке будет существовать определенное число электронов, зависящее от приложенного напряжения.

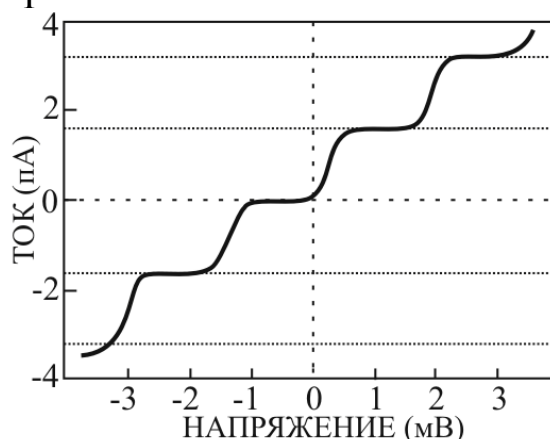


Рис. 4.7. ВАХ системы с двумя несимметричными барьерами – кулоновская лестница

В итоге ВАХ двухпереходной несимметричной системы будет иметь ступенчатый вид так называемой «кулоновской лестницы» (рис. 4.7.)

Заряд Q в уравнении (4.1) складывается из Q_0 и N зарядов электронов, находящихся на кулоновском острове.

$$Q = Q_0 - Ne$$

Поскольку Q_0 имеет поляризационный характер, он может изменяться непрерывно под действием внешнего поля, создаваемого затворным электродом. Таким образом, при непрерывном изменении Q_0 периодически будет выполняться условие кулоновской блокады, проиллюстрированное на рис. 4.2. Следовательно, при последовательном изменении затворного напряжения периодически будет возникать кулоновская блокада, и зависимость тока через кулоновский островок будет носить осцилляционный характер.

Следует отметить, что наряду с классическими эффектами кулоновского взаимодействия всегда присутствуют эффекты квантового туннелирования. В одноэлектронных системах реализуются и квантово - размерные эффекты. В переходных областях могут находиться малые объекты, которые при определенных условиях (температура) могут рассматриваться как нульмерные объекты с дискретным набором энергетического спектра. В отличие от металлов, для полупроводниковых точек «рабочая температура» для квантовых эффектов будет выше из-за низкой плотности состояний. Естественно, что существование дискретного спектра нульмерного объекта предполагает возможность туннелирования электронов через эти уровни. Этот эффект проявляется в экспериментах по измерениям ВАХ в виде тонкой структуры энергетических уровней, накладывающихся на кулоновскую лестницу.

4.2. Классификация и конструкции одноэлектронных приборов

К настоящему времени экспериментально реализовано большое количество приборов одноэлектроники, отличающихся друг от друга своей конструкцией. Характерными признаками,

которые могут быть положены в основу классификации этих приборов, являются [3]:

- направление протекания тока;
- способ формирования квантовых точек;
- количество квантовых точек;
- управляемость параметрами квантовых точек.

По направлению протекания тока приборы делятся на горизонтальные и вертикальные. Направление протекания тока параллельно (горизонтальные) или перпендикулярно (вертикальные) плоскости поверхности структуры.

Квантовые точки в приборах могут быть постоянно сформированными или индуцированными (временными). Постоянная квантовая точка существует все время и представляет собой металлический или полупроводниковый кластер. Индуцированная квантовая точка создается в двумерном электронном газе во время работы прибора путем приложения обедняющих напряжений. Приборы на индуцированных квантовых точках делятся по способу формирования двумерного электронного газа на инверсные и гетероструктурные. В инверсных приборах двумерный электронный газ формируется в инверсионных приповерхностных каналах под действием приложенного напряжения, а в гетероструктурах существует на гетерогранице.

По количеству квантовых точек приборы можно разделить на:

- нульмерные или одноточечные, т.е. такие, в которых рабочий кластер состоит из одной точки;
- одномерные, представляющие собой цепочку точек;
- двумерные, состоящие из массива точек.

По управляемости параметрами квантовых точек приборы разделяются на двухэлектродные, т.е. неуправляемые, и многоэлектродные, управляемые с помощью одного или нескольких затворов.

В обзоре [3] рассмотрены наиболее часто встречающиеся конструкции одноэлектронных приборов.

- Приборы на основе сканирующего туннельного микроскопа.
- Вертикальные одноэлектродные приборы на основе сэндвичевых структур.

- Приборы на основе массивов квантовых точек.
 - Приборы на основе квантовых проволок.
 - Кремниевые одноэлектронные приборы.
 - Приборы на основе двумерного электронного газа в $AlGaAs/GaAs$ гетероструктурах.
 - Приборы на основе структуры $Al/Al_xO_y/Al$
- Остановимся на некоторых из них.

Если между проводящей подложкой и иглой сканирующего туннельного микроскопа расположить металлическую частицу (кластер), изолированную туннельными переходами от подложки и иглы, то такая частица будет играть роль кулоновского островка. По приведенной выше классификации это вертикальный нульмерный неуправляемый прибор на постоянной квантовой точке. (рис. 4.8.)

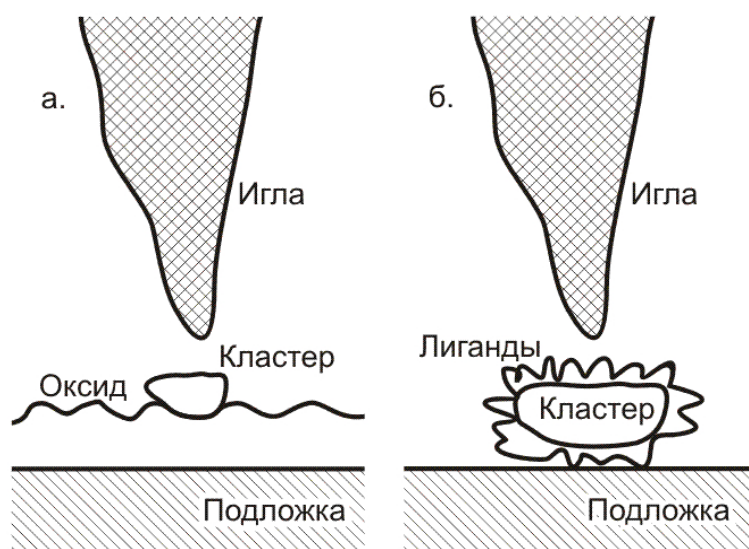


Рис. 4.8. Одноэлектронная система на основе СТМ: а) – металлический кластер на изолирующем слое; б) – металлическая частица окружена изолирующим слоем, играющим роль туннельного барьера [7]

Реализованное таким способом устройство было первым одноэлектронным прибором, работающим при комнатной температуре. Для изоляции металлического кластера от подложки использовались либо непроводящие органические лиганды, либо диэлектрический слой на подложке. В качестве металлических частиц обычно использовалось золото, серебро

или платина, материалов подложки – золото или окисленный алюминий. Толщина диэлектрика обычно составляла ~ 1 нм. Известны работы, где в качестве кулоновского островка использовались органические молекулы жидких кристаллов или молекулы фуллерена [6].

Использование многослойных структур, выращенных методом молекулярно-пучковой эпитаксии (МПЭ), также позволяет создавать одноэлектронные приборы. Поскольку методом МПЭ возможно выращивание структур с точностью до одного монослоя, для формирования квантовой точки достаточно наложить ограничения в двух направлениях, соответствующих плоскости двумерного квантования.

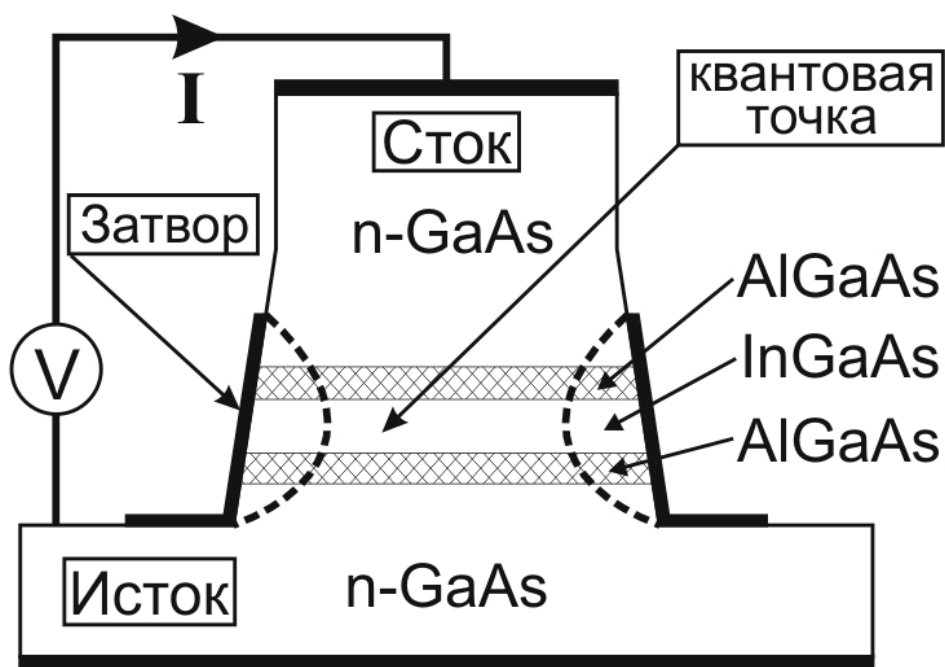


Рис. 4.9. Схематическое изображение субмикронного вертикального одноэлектронного транзистора

Схематическое изображение вертикальной двухбарьерной туннельной структуры, изготовленной в NTT Corp. (Япония), представлено на рис. 4.9. Квантовая точка образуется в области тонкого слоя, между двумя туннельными барьерами. Область локализации электронов в двумерном слое InGaAs изменяется под действием внешнего электрического поля. Таким образом, изменение напряжения на затворе приводит к образованию

квантовой точки, «зажатой» двумя туннельными барьерами (AlGaAs) и внешним электрическим полем.

В работе [8] также исследовалась резонансная двухбарьерная туннельная структура. Верхний металлический контакт (исток) диаметром $0,3 \div 0,7$ мкм, нанесенный на предварительно изготовленную двухбарьерную структуру, использовался в качестве маски для травли 3000 Å. На образовавшуюся ступеньку наносился затворный контакт. Расстояние от затвора до двухбарьерной структуры составляло 500 Å. При подаче отрицательного напряжения на затвор образовывались области обеднения в полупроводнике, образующие квантовую точку между двумя барьерами. Полученная таким образом структура представляет собой вертикальный управляемый прибор на одной индуцированной точке.

Литература

1. *Averin D. V., Likharev K. K.* Coulomb blockade of single electron tunneling, and coherent oscillations in small tunnel junctions // J.Low Temp. Phys 1986. V.62. № 3-4. P. 345-373.
2. *Kouwenhoven L. P., Marcus C. M., McEuen P. L., Tarucha S., Westerwelt R. M., Wingreen N. S.* Electron Transport in Quantum Dots. Mesoscopic Electron Transport // Proc. of the Advanced Study Institute. Ed. by L. L. Sohn, L. P. Kouwenhoven, G Schön. Kluwer, 1997.
3. *Неизвестный И. Г., Соколова О. В., Шамирян Д. Г.* Одноэлектроника. Часть I // Микроэлектроника. 1999. Т. 28. № 2. С. 83-107.
4. *Likharev K. K.* Correlated discrete transfer of single electrons in ultrasmall tunnel junctions // IBM J. Res. Develop. 1988. №1. P. 144 - 158.
5. *Likharev K. K., Claeson T.* Single electronics. // Sci. Am. 1992. V.6. P.80-85.
6. *Nejo H., Aono M., Baksheyev D. S., Tkachenko V. A.* Single-electron charging of molecule observed in scanning tunneling scattering experiments // J. Vac. Sci. Technol. B. 1996. V. 14. Issue 4. P. 2399-2402.

7. *Van Kempen H., Dubois J. G. A., Gerritsen J. W., Schmid G.* Small metallic particles studied by tunneling microscopy // *Physica B*. 1995. V. 204. P. 51-56.
8. *Austing D. G., Honda T., Takura Y., Tarucha S.* Sub-micron vertical AlGaAs/GaAs resonant tunneling Single electron transistor // *Jpn. J. Appl. Phys.* 1995. V. 34. P. 1320-1325

5. МИКРОЛИТОГРАФИЯ

Стремительный прогресс в микроэлектронике – это результат постоянного совершенствования технологических процессов коллективной обработки материалов. Поскольку до настоящего времени основным материалом для производства микросхем является кремний, эта технология носит название *кремниевой планарной технологии* и результатом ее являются готовые интегрированные на кристалле электронные компоненты. Схема производства интегральных МС с высокой степенью интеграции представлена на рис. 5.1. Третья часть производственных затрат в производстве полупроводниковых приборов и ИМС приходится на литографические и высокотемпературные процессы [1].

С другой стороны, литографические методы остаются основным сдерживающим фактором в процессе повышения степени интеграции. Литографические методы, используемые в микроэлектронной технологии, берут свое начало в XIX веке, как процесс изготовления печатных форм для полиграфии. Исторически литографией называется метод печати с плоских металлических или каменных пластин, на которых буквы (или какой-то другой рисунок), удерживающие краску, формируются с помощью фоточувствительного материала (фоторезиста). И только в 1960-х годах эта технология была приспособлена электронной промышленностью для изготовления полупроводниковых приборов, а в дальнейшем и ИМС.

При любом способе переноса изображения методом фотопечати ухудшается резкость края. Этот эффект проявляется при всех способах печати - контактной, с зазором или проекционной. Если есть две щели, размещенные на некотором расстоянии друг от друга, то неэкспонированный участок частично будет экспонирован. Возможные причины этого состоят в следующем:

- дифракция на щели;
- глубина фокуса оптической системы (дефокусировка);
- низкоконтрастный фоторезист;
- отражение от подложки и образование стоячих волн;
- преломление света в фоторезисте.

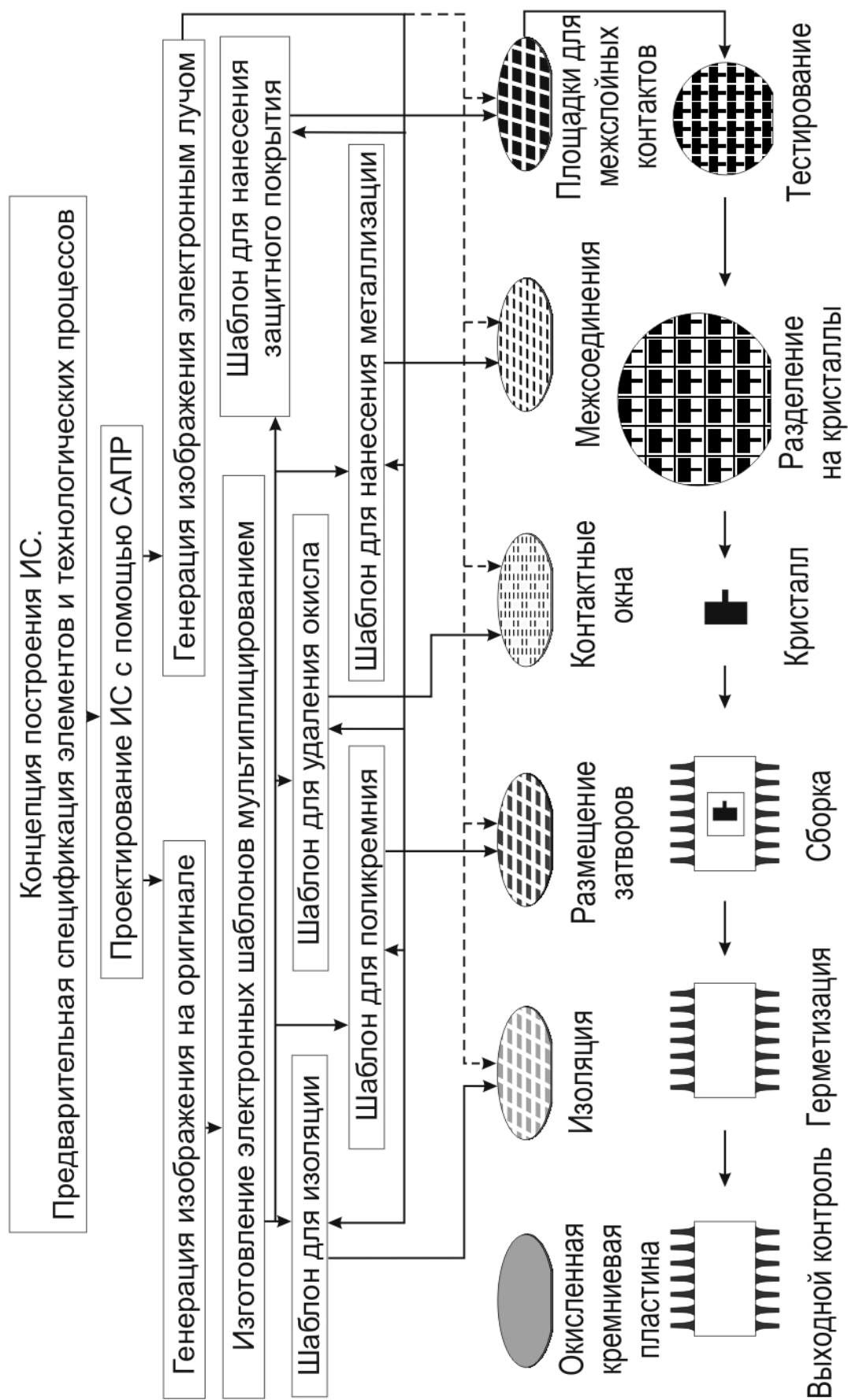


Рис. 5.1. Схема производства ИМС

5.1. Технология и материалы литографии

В конечном итоге задача фотолитографии состоит в том, чтобы воспроизвести двумерный рисунок фотошаблона (ФС) в фоторезисте (ФР) и обеспечить совмещение рисунков с точностью в пределах $\pm 15\%$ от нормального размера элементов ФС [1]. Послойное совмещение должно осуществляться с точностью не менее $\pm 25\%$ от минимального размера элемента. Кроме точного воспроизведения топологии требуется высокая степень воспроизводимости технологического процесса в целом. При этом должны учитываться все проблемы - не только связанные с оптическими явлениями, но и механические ограничения установок экспонирования, термические искривления пластин, свойства ФР и др.

Используемые в обычной оптической литографии источники экспонирующего излучения бывают точечные (лазеры) и протяженные (ртутные лампы). Спектр излучения этих источников можно условно разделить на 3 диапазона:

- Дальний ультрафиолет (ДУФ) – от 100 до 200-300 нанометров.
- Средний ультрафиолет (СУФ) – от 300 до 360 нанометров.
- Ближний ультрафиолет (БУФ) – от 360 до 450 нанометров.

В качестве точечных источников могут быть использованы азотные (N), гелий-кадмиевые (He-Cd), эксимерные лазеры. У ртутной лампы мощность излучения максимальна на длинах волн 365 нм (I – линия), 405 нм (H – линия), 435 нм (G – линия).

Существующие фотолитографические устройства оптической литографии можно разделить на 3 группы:

- установки теневого экспонирования;
- проекционные установки с преломляющей оптикой;
- проекционные установки с отражательной оптикой.

При теновом экспонировании ФС, выполненный в масштабе 1:1, находится в контакте с подложкой (точнее - с фоторезистом) или отделен от нее на некоторое расстояние (несколько микрометров) в случае печати с зазором. Главными недостатками контактной печати являются ограничения по совмещению и возможные повреждения фотошаблонов.

Проекционные системы используют линзы или зеркала, позволяющие проецировать рисунок с фотошаблона на квадратное поле ограниченной площади (напр. 20×20 мм) или полосу шириной несколько миллиметров. Масштаб печати: 1:1, 5:1 или 10:1. Область печати затем сканируется по пластине. Пошаговые системы экспонирования должны удовлетворять следующим условиям:

1. Хорошее совмещение отдельных кристаллов для компенсации искривления шаблонов и платины.
2. Высокоточный механизм перемещения - координатный стол, позволяющий производить точное совмещение рисунков при каждой операции экспонирования, даже тогда, когда различимость меток падает ниже приемлемого уровня.
3. Возможность изменения координатной сетки без снижения производительности.
4. Небольшие размеры установки экспонирования, возможность ее монтажа в сверхчистой комнате.
5. Высокая производительность.

В стандартной проекционной системе фокус объектива (f) является функцией диаметра его входного зрачка (D). Числовая апертура (NA) в среде с показателем преломления n определяется как

$$NA = n \cdot \sin \theta = \frac{D}{2f} \quad (5.1)$$

Разрешение (W) объектива, определяемое для двух непрозрачных объектов в соответствии с критерием Рэлея, равно

$$W = \frac{K \cdot \lambda}{NA} \quad (5.2)$$

Значения коэффициента K различны для разных типов резистов: от $K \geq 0,3$ для резистов, формирующих изображение в верхнем поверхностном слое, до $K \geq 1,1$ для резистов на отражающей поверхности. Практическим разрешением принято считать трехкратное значение разрешения, определенное по Рэлею на длине волны экспонирования.

$$W = 1,83 \frac{\lambda}{NA} \quad (5.3)$$

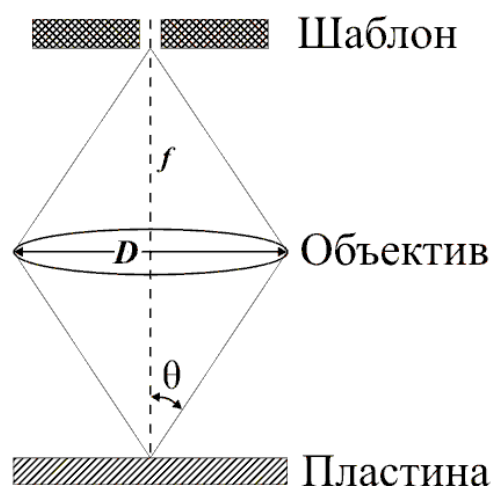


Рис. 5.2. Взаимное расположение элементов оптической системы при определении значения числовой апертуры

Таким образом, разрешение улучшается с использованием более коротковолнового излучения и объектива с большей числовой апертурой (за счет уменьшения размера экспонируемого поля). Глубина фокуса (DF), однако, уменьшается с ростом NA и поэтому определение положения каждого нового кристалла на пластине требует дополнительной фокусировки:

$$DF = \pm \frac{\lambda}{2(NA)^2} \quad (5.4)$$

В таблице 5.1. представлены значения глубины фокуса и числовой апертуры для различных длин волн.

Таблица 5.1.

Значения числовой апертуры (NA) и глубины фокуса (DF) при различных длинах волн излучения.

Числовая апертура (NA)	Глубина фокуса (DF), мкм	
	$\lambda=400$ нм	$\lambda=200$ нм
0,2	5,0	2,5
0,3	2,0	1,0
0,4	1,5	0,8
0,5	0,8	0,4

Например, для объектива с числовой апертурой $NA = 0,35$ при экспонировании на длине волны 300 нм глубина фокуса DF

не более 1,5 мкм. При использовании пластин большого диаметра их неплоскостность, наличие топографического рельефа, неравномерность в толщине резиста могут создать ситуацию, когда невозможно формирование элементов ИМС субмикронных размеров.

Качество и размеры скрытого изображения, сформированного в ФР, величина краевого градиента профилей в резисте ограничивается фундаментальными физическими характеристиками световых волн, такими как дифракция, когерентность, астигматизм, интерференция, хроматические aberrации. Использование лазеров в качестве источника когерентного излучения позволяет избежать влияния ряда этих факторов. Пространственно когерентное лазерное излучение обладает высокой яркостью и низкой угловой расходимостью θ , ограниченной дифракцией

$$\theta = \frac{K \cdot \lambda}{d}, \quad (5.5)$$

где λ - длина волны, d - диаметр выходной диафрагмы лазерного резонатора, $K = 1,2$ для однородного пучка и $K = 2/\pi$ для гауссова пучка. Дифракционный предел разрешения линзы равен

$$\Phi = \frac{K \cdot f \cdot \lambda}{D} \quad (5.6)$$

Здесь f - фокусное расстояние, D - апертура. Основным недостатком использования когерентных источников излучения является серьезное ухудшение изображения за счет дифракционных эффектов.

5.1.1. Контактная печать и печать с зазором

При контакте ФШ и подложки может быть получено сколь угодно высокое разрешение. На практике, однако, многократное использование шаблонов приводит к их повреждениям. Перемещения по пластине, процессы совмещения и как следствие этого «уход» размеров ограничивают предел точности до величины ~ 1 мкм. В контактной литографии обычно применяется узконаправленное когерентное излучение

(расходимость несколько градусов для уменьшения дифракционных эффектов)

Модуляционная передаточная функция (МПФ) процесса выражает связь между объектом и изображением и определяется как

$$МПФ = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (5.7)$$

При контактной печати МПФ достаточно велика ($> 0,8$). Использование более совершенных контактных шаблонов и многослойных резистов позволяет получить изображение $0,1 \text{ мкм}$ и ниже. При печати с зазором Z между шаблоном и пластиной предельное разрешение W составляет $1\text{-}2 \text{ мкм}$ для зазора $5\text{-}10 \text{ мкм}$:

$$W \approx \sqrt{0,7 \cdot \lambda \cdot Z} \quad (5.8)$$

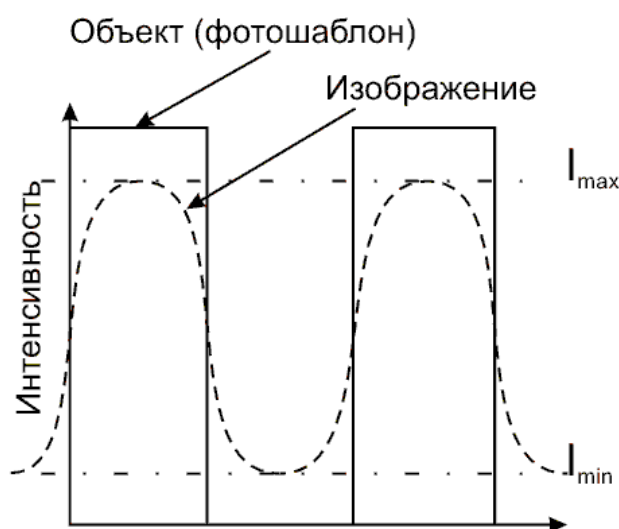


Рис. 5.3. Определение модуляционной передаточной функции (МПФ).

Близко расположенные линии при контактной печати или печати с зазором расплываются из-за конструктивной интерференции между волнами, дифрагировавшими на соответствующих окнах ФШ. Если на одно из соседних окон нанесено покрытие, изменяющее фазу проходящего через него излучения на 180° , то при толщине этого покрытия $t = \lambda(2n - 1)$ между световыми потоками от различных отверстий происходит интерференция, минимизирующая дифракционные эффекты.

Преимущества и недостатки контактной печати и печати с зазором представлены в таблице 5.2.

Таблица 5.2.

	Контактная печать	Печать с зазором
Преимущества	Высокое разрешение, до 0,1 мкм	Удовлетворительное разрешение до 1,0 мкм Большая экспонируемая площадь Высокая производительность Совмещение до 05 мкм и менее
Недостатки	Неидеальный контакт Повреждения шаблона Пропечатка дефектов Трудоемкое совмещение	Малая глубина фокуса Влияние кривизны пластины Жесткие требования к расходимости экспонирующего излучения.

5.1.2. Проекционная печать

При проекционной печати в сканерах и устройствах пошагового совмещения (степперах) используется как отражательная, так и преломляющая оптика. В современных степперах проводится совмещение на каждом поле и достигается согласование по координатам, углу поворота, фокусу и наклону. Метод проекционной печати позволяет воспроизводить элементы в резисте с минимальными размерами ниже 0,1 мкм. Различные методы печати могут быть классифицированы по точности совмещения, производительности, разрешающей способности. В работе [2] был введен параметр предпочтительности M для сравнения методов сканера и степпера.

$$M = \frac{10^6 TY}{CL^2},$$

где T - производительность (пластина/час), $Y = \exp(-ADN)$ – выход годных, C – стоимость, тыс. долларов, L – минимально достижимая ширина линии в микрометрах, A – площадь шаблона

(см²), D - плотность дефектов (дефект/см²), N – число маскирующих слоев.

Кроме рассмотренных литографических методов и присущих им преимуществ и недостатков, следует помнить о других факторах, сказывающихся на конечном результате. Чрезвычайно важен, например, прогресс в технологии изготовления фотошаблонов, совершенствование методов совмещения. Внутренние и взаимные эффекты близости становятся главными проблемами фотолитографии по мере уменьшения характерных размеров элементов.

Учитывая, что оптическая литография является ключевым элементом всей технологии микроэлектроники, для нанoeлектронной технологии роль литографии возрастает еще более. Увеличение степени интеграции требует уменьшения характерных критических размеров, а следовательно, использования все меньших длин волн света. Кроме того, используются все более совершенные фоторезисты, фазосдвигающие маски и коррекция эффектов близости. Современный уровень технологии фактически уменьшил нижнюю границу вплоть до 100 нм. Это подразумевает, однако, использование большого количества сопутствующих технологий, что значительно повышает стоимость технологического процесса. В настоящее время, на ряде ведущих фирм мира, занятых в высокотехнологичном электронном производстве, разрабатываются технологии и изготавливается оборудование для целого класса литографических методов, называемых «литографией нового поколения». Это ультрафиолетовая, рентгеновская, электронно-лучевая и ионная проекционная литографии.

Существует много способов создания структур и элементов с малыми размерами. В основном все технологии связаны с изготовлением шаблонов, представляющих собой реплику того или иного типа структур. Использование этих шаблонов и масок предполагает нанесение рисунка на пластину и дальнейшие технологические операции, предусмотренные технологическим процессом. Новые технологии используют, однако, маски как некоторые промежуточные шаги для изготовления наноструктур. По оценкам ITRS [3] минимальный размер быстро уменьшается,

что требует изменения и других параметров технологических процессов.

Таблица 5.3.

Прогресс в развитии литографии

Год получения	2001	2003	2006	2009
Минимальный размер, нм	150	120	90	65
Толщина верхнего слоя, нм	55	45	35	25
Оптическая литография 193 нм	✓	✓		
Оптическая литография 157 нм		✓	✓	
Дальний ультрафиолет			✓	✓
Рентгеновское излучение				✓
Электронный луч			✓	✓
Ионный пучок			✓	✓
Нанотиснение и печать				✓

Сравнение результатов полученных с помощью различных литографических методов нового поколения для новых технологий и приборов представлено таблицей 5.4. Технология микро и нанoeлектроники может быть условно разделена на две группы. Первая – это технология, использующая традиционные методы оптической литографии, травления, легирования и т.п. Вторая – активно продвигаемые в последние годы технологии, направленные на создание приборов молекулярной электроники и одноэлектронных приборов.

Таблица 5.4.

Проекционные методы		Методы прямого рисования	
Название метода	Критический размер - CD	Название метода	Критический размер - CD
Extreme Ultra Violet - EUV	30 нм	Scanning Probe Method (SPM)	~1 нм
Ion Projection IPL	50 нм	Focused Ion Beam (FIB)	50 нм
Proximity X-Ray 1X X-ray	60 нм	Electron-beam direct write (EBDW)	~20 нм
Electron-beam projection SCALPEL	70 нм	Multi electron-beam direct write (multi-EBDW)	~100 нм
Высокая стоимость		Низкая производительность	

5.2. Литографии для КМОП технологии

5.2.1. Оптическая литография

Методы оптической литографии являются традиционными и в настоящее время им принадлежит доминирующая роль в технологии. Существует ряд путей совершенствования современного фотолитографического оборудования. Это улучшение конструкции объективов и систем совмещения, использование пластин с лучшей плоскостностью, использование планаризирующих слоев, разработка новых резистов. В качестве источников когерентного излучения используются эксимерные лазеры, в частности KrF (248 нм) для процессов со 180 нм разрешением, ArF для 197 нм, а F2 для 157 нм. Ведется поиск новых оптических материалов для изготовления линз с низкой абберацией, высоким апертурным числом, высокой разрешающей способностью.

5.2.2. Литография предельного ультрафиолета

Литография предельного ультрафиолета (Extreme Ultraviolet lithography - EUV) или литография в области мягкого рентгеновского излучения (Soft X-Ray lithography) по существу является развитием методов оптической литографии, использующей источники излучения с меньшими длинами волн. Важнейшим отличием является необходимость использования многослойных зеркал и для отражательной оптики и для масок в интервале длин волн 10-15 нм. Метод быстро и довольно успешно развивается. Если, например, в 2000 году рекордными были результаты: 100 нм разрешение по расстоянию между линиями и 70 нм для отдельного элемента, то к настоящему времени эти цифры снизились до 30 нм. Метод является весьма привлекательным и многообещающим для КМОП технологии, но по оценкам экспертов в промышленности он займет свое место через 4-5 лет.

В качестве источника излучения большой мощности для EUV литографии используется либо синхротронное излучение, либо излучение плазмы, полученной взаимодействием лазерных импульсов со сверхзвуковым потоком газа (см рис. 5.4.).

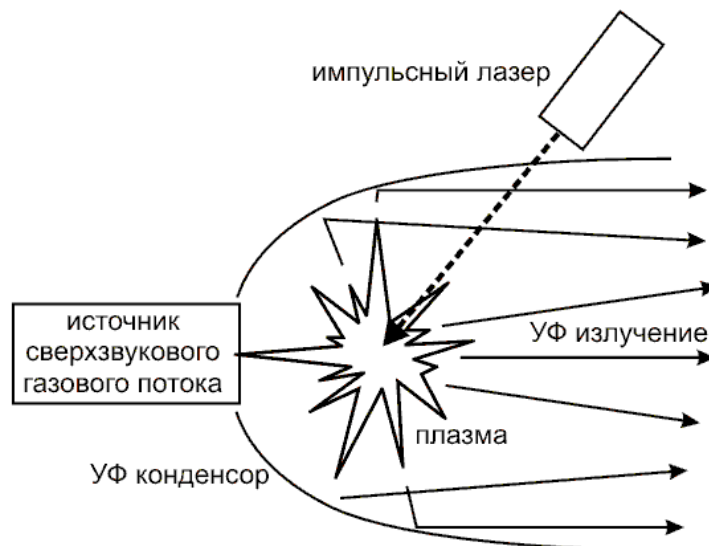


Рис. 5.4. Схема плазменного источника излучения для EUV

В отличие от синхротрона, плазменные источники для EUV литографии значительно дешевле. Однако существует ряд технических проблем, в частности с УФ оптикой. Любое оборудование для этого метода использует зеркала с многослойными металлическими покрытиями. Требования к оптическим характеристикам очень жесткие. В частности поверхности и границы раздела зеркал и других оптических элементов должны иметь неровности не более 0,25 нм. Необходимо также термостатирование рабочих элементов или компенсация «ухода» размеров, вызванного изменением температуры.

В 2002 году фирма Canon объявила о значительном прогрессе в разработке 157 нм сканеров для 70 нм литографии. Главной причиной задержки, вставшей на пути разработки оборудования для производства чипов по новому, более прецизионному техпроцессу, стали проблемы с проекционной оптикой, точнее, проблемы с двойным преломлением лучей (intrinsic birefringence, IBR), присущие линзам из фторида кальция (CaF_2). Подразделение Semiconductor Equipment компании Canon U.S.A. сделало заявление, в котором рапортовало о значительном продвижении на пути разработки 157 нм проекционной оптики, в частности, серьезных модификациях конфигурации оптических систем, позволяющих обойти IBR. И теперь, по словам представителей компании, они будут готовы предоставить первое

поколение систем экспозиции для узлов порядка 70 нм на всестороннее испытание и опробование уже в первой половине 2003 года. Первым потребителем, твердо заявившим о приобретении новой системы, стала германская Infineon Technologies AG, намеренная запустить производство по техпроцессу от Canon уже в начале 2005 года.

Таким образом, разработка катадиоптрических линз прошла успешно, и теперь на основе этой оптики фирма Canon намерена разработать промышленную систему FPA-5800FS1. Пока что компания гарантирует, что новая оптика будет избавлена от эффекта преломления лучей в рамках 100 нм литографии, а работа над более прецизионным процессом будет построена на том же принципе – оптимизации ориентации.

Еще в 2000 году коалиция национальных лабораторий США, таких как Lawrence Livermore National Laboratories и Sandia National Laboratories, а также крупнейших производителей чипов, таких как Intel, AMD и т.д., продемонстрировали прототип новой EUV литографической технологии, которая позволит продолжить миниатюризацию элементов микросхем. Новая технология позволит создать к 2005 году процессор с тактовой частотой 10 ГГц. Тем самым предел миниатюризации, предрекаемый многими экспертами, отодвинут еще как минимум на 8-10 лет. Однако новая технология использует весь потенциал кремния без остатка. И после 2011 года придется найти новый материал для создания чипов или же разработать принципиально новый метод их производства.

5.2.3. Литография ближнего рентгеновского диапазона

Литография ближнего рентгеновского диапазона (X-ray proximity lithography) использует длину волны излучения 1 нм. Этот метод является, по-видимому, последним шагом на пути уменьшения длины волны излучения, используемого в нанолитографии. Метод развивается интенсивно последние 20 лет и почти готов к использованию. В отсутствие необходимой рентгеновской оптики метод основан на проецировании изображения фотошаблона на пластину в масштабе 1:1.

Источники излучения – такие же как и в EUV литографии. Стоимость оборудования ежегодно понижается и в настоящее время оно является вполне доступным для коммерческого использования крупными электронными фирмами.

5.2.4. Электронно-лучевая литография и SCALPEL

Из теории дифракции и практической спектроскопии известно, что разрешающая способность оптических приборов ограничена длиной волны используемого излучения. Именно поэтому для изучения объектов атомных размеров используются более высокоэнергетичные частицы: рентгеновские фотоны, электроны и ионы. Эти аналитические приборы и составили основу литографического оборудования для нанoeлектроники.

Первой областью применения таких экспонирующих систем было производство, контроль и восстановление фотошаблонов. Затем последовало использование сканирующих систем для прямой литографии по кремнию и экспонирование широким пучком. Длина волны экспонирующего излучения высокой энергии измеряется ангстремами, а резистом может служить любой полимер либо неорганический пленкообразующий материал. Энергия излучения такова, что ее достаточно не только для возбуждения атомов, но и для перестройки любых химических связей. Таким образом, использование пучковых методов, с одной стороны, снимает ограничения по разрешению, а с другой – ведет к снижению производительности процесса экспонирования и повышению стоимости технологии.

Пучковое экспонирование (используется также термин «радиационное экспонирование») применяется в тех случаях, когда параметры оптической литографии не удовлетворяют требованиям по точности совмещения и глубине фокуса.

Траектория движения свободного электрона в вакууме может изменяться под действием электрических и магнитных полей. Это явление и лежит в основе метода электронно-лучевой литографии. Фундаментальное ограничение по разрешению определяется соотношением неопределенностей Гейзенберга ($\Delta p_x \Delta x \geq \hbar$) и может достигать значений ниже 10 нм. Как уже отмечалось в первой главе, минимальная ширина

линии $\Delta x \propto 1/\sqrt{E}$. В 1967 году было предложено использовать электронно - лучевую трубку в качестве экономичного и быстродействующего генератора фотошаблонов для специализированных ИМС [4], а в 1968 электронно-лучевое экспонирование было применено для непосредственного «рисования» по резисту [5].

В настоящее время производительность электронно-лучевой (e-beam) литографии ограничена только скоростью сканеров, используемых для нанесения рисунка, и метод является уже серийным в производстве ИМС. Один из предлагаемых способов увеличения производительности – использование матрицы параллельных электронных пучков (рис 5.5.). Однако это накладывает жесткие требования к характеристикам каждого источника электронов в матрице по интенсивности, моноэнергетичности и методам управления пучком.

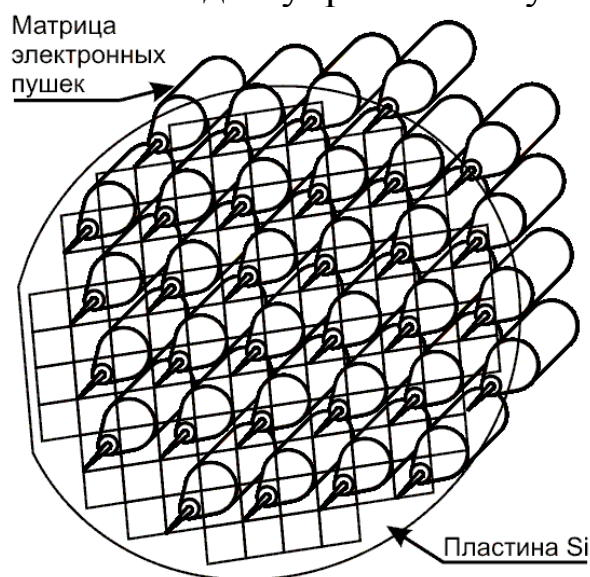


Рис. 5.5. Схема использования матрицы параллельных электронных пучков для прямого нанесения рисунка на пластину кремния

В качестве источников электронов предлагается использовать алмазоподобные соединения углерода - фуллерены или нанотрубки. Идеальный источник должен быть не только механически прочным, но и устойчивым к жестким условиям технологической среды (например устойчивым к процессам

дегазации поверхности пластин), иметь высокие эмиссионные характеристики при низкой энергии пучка.

Высокоэнергетичные пучки, проникая в подложку, нарушают приповерхностный слой, образуют высокую плотность дефектов в тонком приповерхностном слое. Этот эффект особенно чувствителен при повышении степени интеграции и уменьшении всех размеров элементов ИМС. Решение этой проблемы в понижении энергии электронов и уменьшении глубины проникновения в подложку. В частности, электроны с энергией ниже 300 эВ имеют длину свободного пробега в кремнии менее 1 нм и, таким образом, практически не вносят дефектов в приповерхностный слой.

Для формирования рисунка на большой площади пластины, кроме рассмотренной выше схемы микроколонн (или микропучек), используется так называемая технология микропучков. Микроколонны используют «непланарную» конструкцию устройства – размер несколько миллиметров, источник электронов по схеме Шоттки, фильтры, микролинзы и детекторы. При использовании микропучковой технологии конструкция такова, что электроды, отклоняющие системы и детекторы расположены в плоскости кремниевой пластины.

Оценки производительности системы микроколонн в установке для 100 нм литографии следующие:

- разрешение до 50 нм;
- пластины \varnothing 300 мм;
- производительность 10 пластин в час;
- количество микроколонн – 50;
- рабочий ток – 30 нА на колонну.

В случае увеличения числа колонн до 200 ток понижается до 0,8 нА на колонну. Если количество обрабатываемых пластин будет увеличено, например, до 25 штук в час, то все параметры соответственно должны быть изменены в 2,5 раза. Уменьшение времени экспонирования как для микроколонн, так и для микропучков сдерживается ограничениями роста величины тока на пучок. Источник должен работать при низких энергиях, быть стабильным, когерентным, надежным.

В настоящее время работы ведутся в направлении совершенствования электронно-лучевых проекционных литографических установок с целью увеличения разрешающей способности (менее 100 нм). Эксперименты с единичными электронными пучками показывают, что достижимым пределом этой технологии является 30 нм разрешение, а для систем с микроколоннами или микропучками промышленный предел лежит в области 40-50 нм.

Преимущества e-beam проекционной литографии известны несколько десятилетий. Развитие этого метода до некоторого времени сдерживало две причины:

1. Повышение температуры масок как результат поглощения энергии падающего пучка электронов, ограничения ускоряющих напряжений, невозможность создания рисунка в виде замкнутого кольца и использования дополняющих масок.
2. Полноценная оптика, не масштабируемая в связи с увеличением размера пластин и чипов и уменьшением минимальных размеров элементов. Для уменьшения влияния абберрации необходимо малое апертурное число. Зарядовые эффекты, возникающие при приемлемом для достижения производительности токе пучка, разрушают достигаемое разрешение метода.

Работа [6], где эти вопросы были рассмотрены, стала толчком к разработке на фирме Bell Laboratories концепции электронно-лучевой проекционной литографии с ограничением по углу рассеяния - SCALPEL[®] (SCattering with Angular Limitation Projection Electron-beam Lithography) [7]. Эта программа комплексная и включает в себя разработку как проекционного оборудования, так и технологии изготовления масок, резистов, методов обработки материалов. Одно из основных требований – это совместимость со стандартными технологическими процессами.

Принцип работы проекционной электронно - лучевой литографии SCALPEL представлен на рисунке 5.6. Электронный пучок, используемый в процессе, некогерентный. Главным отличием этого метода от всех остальных является конструкция маски (рис. 5.7).

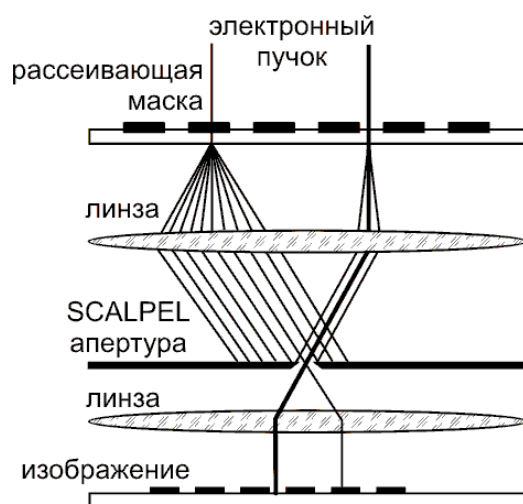


Рис. 5.6. Принцип работы проекционной электронно-лучевой литографии SCALPEL

Мембрана толщиной 75 - 150 нм из SiN_x на решетку вытравленную в Si подложке. На мембрану нанесен рисунок, который собственно и является фотошаблоном. Масштаб рисунка 4:1. То есть элементы с характерным размером 100 нм на рассеивающем масочном покрытии имеют размер 400 нм. Толщина покрытия 50 нм, материал – слои W и Cr. Толщина стенок решетки – 0,2 мм.

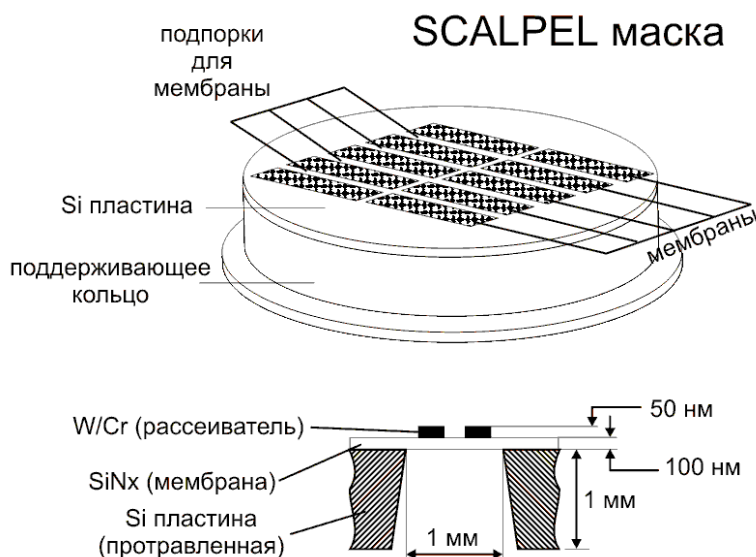


Рис. 5.7. Конструкция маски для SCALPEL технологии.

Электронный пучок почти не рассеивается мембраной и рассеивается на угол в несколько миллирадиан слоями W и Cr. Этого рассеяния достаточно, чтобы апертурное окно, расположенное в фокальной плоскости, задержало рассеянные электроны. Результатом этого процесса является высококонтрастное изображение маски, получаемое на поверхности кремниевой пластины.

Производительность процесса зависит от тока пучка и чувствительности резиста. Взаимодействие электронов в пучке приводит к его существенной размазке, пропорциональной $I^{2/3}$, ограничивая разрешение 35 нм при токе 10 мкА. Следствием некогерентности излучения является отсутствие дифракционных и интерференционных явлений, сохранение линейных зависимостей для всех рабочих параметров при уменьшении размеров элементов до 35 нм. Это в свою очередь означает, что нет необходимости в коррекции погрешностей, обусловленных эффектами близости.

5.2.5. Ионно-пучковая проекционная литография

Ионно-лучевые системы для экспонирования резистов широким пучком созданы на основе промышленных установок ионной имплантации. Источники ионных пучков существенно усовершенствовались за годы развития сканирующих ионных микроскопов и имплантаторов. Главным достоинством ионных пучков является малое обратное рассеяние, а следовательно, минимальный эффект близости.

Фокусированные ионные и электронные пучки используются для рисования непосредственно по пластине элементов с размерами менее 20 нм. Поскольку процесс занимает много времени, использование его в массовом производстве неприемлемо. Ионно-пучковые проекционные установки сочетают в себе высокие разрешающую способность и производительность. Чувствительность резиста высокая, поскольку длина пробега легких ионов (H^+ , He^{++}) при энергиях 50-100 кэВ лежит в том же интервале, что и типично используемая в технологическом процессе толщина резистов. При чувствительности резиста $10^{12} - 10^{13}$ ион/см² необходимо

менее 1 секунды для экспонирования площадки от 10 до 100 мм². Ионы также могут использоваться в безрезистивных процессах, когда происходит прямая модификация поверхности. На подложках различных типов под действием ионного облучения могут образовываться островки магнитных материалов, массивы квантовых точек или волнообразный рельеф.

Таблица 5.5.

Литография	Доза
Ионная проекционная	11,3 мДж/см ²
Электронно-лучевая (SCALPEL)	400 мДж/см ²
Предельный УФ	10 мДж/см ²

Для сравнения скорости рисования электронным и ионным пучками следует помнить, что ионы с энергией 50-100 кэВ всю ее передают резисту, тогда как электроны с такой энергией проникают глубоко в подложку. Это приводит к существенной разнице в чувствительности резистов. В таблице 5.5. приведены данные для различных литографических методов [8].

5.3. Новые нанотехнологии

5.3.1. Электронно-лучевая нанолитография

В настоящее время электронно-лучевая литография составляет существенную часть производства ИМС. Установки с Гауссовым пучком и сканирующие электронные микроскопы работают в интервале энергий от 100 до 200 кэВ и диаметре пучка до 1 - 10 нм. Используя высокую разрешающую способность метода и низкочувствительные резисты (ПММА), минимальный размер элемента до 30 нм удается получать с высокой степенью воспроизводимости. Сообщалось о достижении рекордных результатов в несколько нанометров с использованием неорганических резистов. Основные усилия направлены на получение хорошо воспроизводимых характеристик при уменьшении размеров элементов. В работе [9] сообщалось об одноэлектронном транзисторе, работающем при комнатной температуре и изготовленном методом электронно-лучевой литографии. Размер Si островка в этом транзисторе удалось получить 10 нм, а ширину туннельного барьера - 1 нм.

Главным недостатком всех электронно - лучевых литографических методов остается низкая скорость нанесения рисунка, а следовательно, и низкая производительность, высокая стоимость процесса, неприемлемая для массового производства.

Возможные пути решения этих проблем состоят в следующем:

- параллельное экспонирование и SCALPEL технология;
- рисование электронным пучком сверхнизкой энергии: необходимо увеличение чувствительности резиста, которое может привести к увеличению производительности на порядок. В настоящее время этот метод не позволил получить разрешение лучше 50 нм;
- параллельное рисование с использованием матрицы микроколонн;
- разработка более чувствительных резистов и связанная с этим оптимизация электронно-пучковых процессов.

5.3.2. Методы на основе СТМ

Прогресс в развитии методов, основанных на технике сканирующей туннельной микроскопии тем не менее не решает главной проблемы – низкой производительности. Метод СТМ может быть хорошим дополнением к существующим литографическим методам.

5.3.3. Методы печати

Существует два основных метода печати наноформ, используемых в нанотехнологии. Первый связан с формированием тонкого полимерного слоя, нанесенного на подложку. При этом контролируется температура и давление, используемое при печати. Результат этой технологической операции называется наноотпечатком или тиснением. Второй состоит в перенесении монослоя самоорганизованных молекул с эластичного штампа на подложку, где они и образуют рисунок. Оба типа печати имеют примерно одинаковые проблемы на пути их широкого использования. Это повышение надежности и процесса, увеличение производительности и снижение стоимости.

Тиснение

Рисунок 5.8 демонстрирует основные этапы этого процесса. Форма изготавливается методами оптической или электронно-лучевой литографии и последующим сухим травлением. На подложку наносится тонкий слой полимера. Форма и подложка помещаются в держатели, имеющие температуру несколько выше температуры стеклования полимера. После установления температурного режима для подложки и формы они приводятся в контакт и в полимере образуется отпечаток и следует охлаждение.

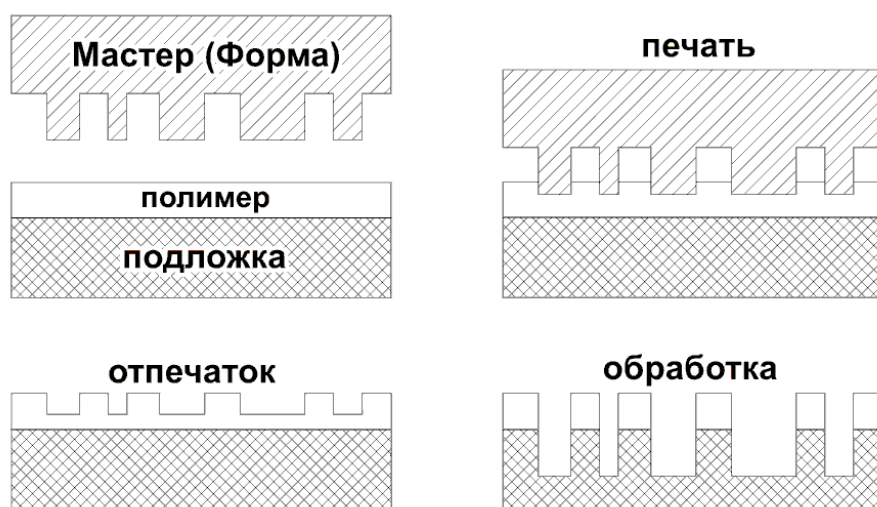


Рис. 5.8. Схема процесса нанотиснения

Процессы нагрева и печати занимают несколько минут. Разделение формы и подложки происходит при $T \approx 50^\circ\text{C}$. После образования рельефа в полимерном покрытии он может быть использован как маска для процесса травления или как этап подготовки какого-либо другого процесса. Возможно использование структуры, образованной в полимере, как самостоятельного элемента, имеющего специфические электрофизические или оптические характеристики. В принципе, используя этот метод могут быть получены размеры менее 10 нм. Однако для получения рекордных результатов необходимо решить ряд серьезных проблем. Это и увеличение размеров формы, повышение плотности элементов на мастер – форме, время ее службы, нежелательное прилипание полимера к поверхности формы и некоторые другие.

Рисование

Схема процесса представлена на рис. 5.9. На штамп с эластичным покрытием нанесен рисунок, который переносится в виде монослоя на подложку. Образующийся отпечаток служит маской для последующих технологических операций, таких как травление.

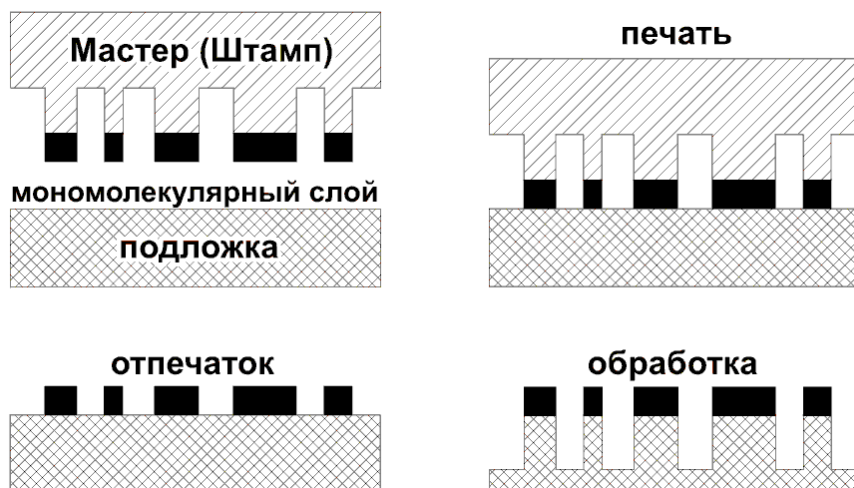


Рис. 5.9. Схема процесса нанорисования

Оценивая производительность метода, следует принимать во внимание время рисования и переноса рисунка, время очистки штампа и т.п. Пути усовершенствования метода с целью повышения производительности и надежности в основном те же, что и для метода тиснения.

Заключение

В настоящее время промышленностью используется высокопроизводительный метод оптической литографии, достигшей уровня 130 нм. В то же время метод электронно-лучевой литографии, имеющий низкую производительность, позволяет достичь размеров 30 нм. Выпуск MPU и DRAM схем вероятнее всего будет опираться на оптическую литографию до 100 нм. Дальнейшее уменьшение размеров потребует разумного компромисса между традиционными оптическими методами и новыми, на основе пучковых технологий. То есть компромисса между высокой производительностью и малыми размерами элементов. Переходный этап, по-видимому, относится к области

100-50 нм и вопрос здесь оказывается наиболее открытым, поскольку развитие новых литографических методов идет крайне неравномерно. Рентгеновская литография и SCALPEL показывают неплохие результаты, судя по литературе. Однако следует учитывать, что момент широкого внедрения новых литографических методов в промышленность сильно зависит от доступности той или иной технологии. И главный вопрос, который возникает на этом этапе – это вопрос стоимости процесса, включая стоимость оборудования.

Следует, однако, заметить, что размеры, которые могут быть достигнуты внедрением новых литографических методов, относятся как раз к той области, где лежат интересы современной нанoeлектронной науки. Это квантово – размерные приборы, резонансно – туннельные диоды, одноэлектронные устройства и многое другое. Необходимость достижения размеров элементов вплоть до 1 нм является очень серьезным стимулом для дальнейшего развития нанотехнологии.

Литература

1. *Моро У.* Микролитография. Принципы, методы, материалы: В двух частях. Часть 1. М.: Мир, 1990. 608.
2. *Allsop B.* Microelectron. Manuf. Test. July, 1981. P. 28.
3. ITRS – International Technology Roadmap for Semiconductors. 1999 Edition, 2000 Update, 2001 Edition. (См. также: <http://public.itrs.net/Home.htm>)
4. US Patent № 3581385, 1971, Texas Instrument.
5. US Patent № 3575588, 1971, IBM.
6. *Berger S.D. and J.M. Gibson* New approach to projection-electron lithography with demonstrated 0.1 μm linewidth // *Appl. Phys. Lett.*, V. 57 153 (1990).
7. *Liddle J.A. , Lloyd R. Harriott, A.E. Novembre and W.K. Waskiewicz* SCALPEL: A Projection Electron-Beam Approach to Sub-Optical Lithography // *Bell Lab. Technology Review*, 1999. (см также <http://www.bell-labs.com/project/SCALPEL>)

8. European Comission, IST Programme: Future and Emerging Technologies/ Technology Roadmap for Nanoelectronics. Ed. by R. Compano, Second Edition, 2000
9. *Kurihara K., Namatsu H., Nagase M., Takahashi Y.* Fabrication and processor room temperature operated single electron transistor using electron beam nanolithography, Silicon Nanoelectronics Workshop, Kyoto, June 1997.

Содержание

<u>ВВЕДЕНИЕ</u>	3
<i><u>Литература</u></i>	6
<u>1. ФИЗИЧЕСКИЕ ОГРАНИЧЕНИЯ МИНИМАЛЬНЫХ РАЗМЕРОВ ЭЛЕМЕНТОВ МИКРОЭЛЕКТРОНИКИ</u>	7
<u>1.1. ФИЗИЧЕСКИЕ ОГРАНИЧЕНИЯ В ТЕХНОЛОГИИ ПРОИЗВОДСТВА ИС</u>	12
<i><u>1.1.1. Размазка края экспонированной области</u></i>	14
<i><u>1.1.2. Тепловое расширение маски и ИС</u></i>	15
<i><u>1.1.3. Сферическая абберация электронного луча</u></i>	15
<i><u>1.1.4. Рассеяние луча в резисте и в полупроводнике</u></i>	17
<i><u>1.1.5. Влияние флуктуаций примеси</u></i>	18
<i><u>1.1.6. Статистическая воспроизводимость технологического процесса</u></i>	20
<u>1.2. ФИЗИЧЕСКИЕ ОГРАНИЧЕНИЯ, НАКЛАДЫВАЕМЫЕ МЕХАНИЗМОМ РАБОТЫ ЭЛЕМЕНТОВ ИС</u>	21
<i><u>1.2.1. Классификация приборов по размерам</u></i>	21
<i><u>1.2.2. Полевой транзистор. Скейлинг параметров</u></i>	22
<i><u>1.2.3. Ограничения электрофизических параметров</u></i>	31
<i><u>1.2.4. Задержка и искажение импульсов на связях</u></i>	38
<i><u>1.2.5. Ограничения на размеры элементов, обусловленные сильными электрическими полями</u></i>	45
<i><u>1.2.6. Ограничения размеров элементов памяти</u></i>	47
<u>1.3. ОГРАНИЧЕНИЯ ИНТЕГРАЦИИ ЭЛЕМЕНТОВ</u>	47
<i><u>1.3.1. Предельная степень интеграции</u></i>	48
<i><u>1.3.2. Теплофизические ограничения на рост интеграции</u></i>	50
<u>1.4. ЗАКЛЮЧЕНИЕ</u>	55
<i><u>Литература</u></i>	57
<u>2. КВАНТОВАЯ МЕХАНИКА ПРОСТЕЙШИХ СТРУКТУР</u>	58
<u>2.1. ПРЯМОУГОЛЬНАЯ ЯМА</u>	59
<u>2.2. ТРЕУГОЛЬНАЯ ЯМА</u>	61
<u>2.3. ВАРИАЦИОННАЯ ВОЛНОВАЯ ФУНКЦИЯ ДЛЯ НИЖНЕЙ ПОДЗОНЫ В СЛОЕ ПРОСТРАНСТВЕННОГО ЗАРЯДА</u>	64
<u>2.4. КВАНТОВЫЕ СОСТОЯНИЯ В НИТЯХ И ТОЧКАХ</u>	66
<u>2.5. ПЛОТНОСТЬ СОСТОЯНИЙ</u>	71

<u>2.6. ПОЛЯРИЗУЕМОСТЬ И ЭКРАНИРОВАНИЕ</u>	76
<u>Литература</u>	78
<u>3. РЕЗОНАНСНОЕ ТУННЕЛИРОВАНИЕ</u>	80
<u>3.1. ДВУХБАРЬЕРНАЯ СТРУКТУРА</u>	83
<u>3.1.1. Коэффициенты прохождения и отражения</u>	83
<u>3.1.2. Квазистационарные состояния в яме между барьерами</u>	86
<u>Литература</u>	88
<u>4. ОДНОЭЛЕКТРОНИКА</u>	89
<u>4.1. КВАНТОВОЕ ТУННЕЛИРОВАНИЕ</u>	90
<u>4.1.1. Теория кулоновской блокады</u>	90
<u>4.1.2. Кулоновская лестница</u>	98
<u>4.2. КЛАССИФИКАЦИЯ И КОНСТРУКЦИИ ОДНОЭЛЕКТРОННЫХ ПРИБОРОВ</u>	100
<u>Литература</u>	104
<u>5. МИКРОЛИТОГРАФИЯ</u>	106
<u>5.1. ТЕХНОЛОГИЯ И МАТЕРИАЛЫ ЛИТОГРАФИИ</u>	108
<u>5.1.1. Контактная печать и печать с зазором</u>	111
<u>5.1.2. Проекционная печать</u>	113
<u>5.2. ЛИТОГРАФИИ ДЛЯ КМОП ТЕХНОЛОГИИ</u>	116
<u>5.2.1. Оптическая литография</u>	116
<u>5.2.2. Литография предельного ультрафиолета</u>	116
<u>5.2.3. Литография ближнего рентгеновского диапазона</u>	118
<u>5.2.4. Электронно-лучевая литография и SCALPEL</u>	119
<u>5.2.5. Ионно-пучковая проекционная литография</u>	124
<u>5.3. НОВЫЕ НАНОТЕХНОЛОГИИ</u>	125
<u>5.3.1. Электронно-лучевая нанолитография</u>	125
<u>5.3.2. Методы на основе СТМ</u>	126
<u>5.3.3. Методы печати</u>	126
<u>Тиснение</u>	127
<u>Рисование</u>	128
<u>Заключение</u>	128
<u>Литература</u>	129