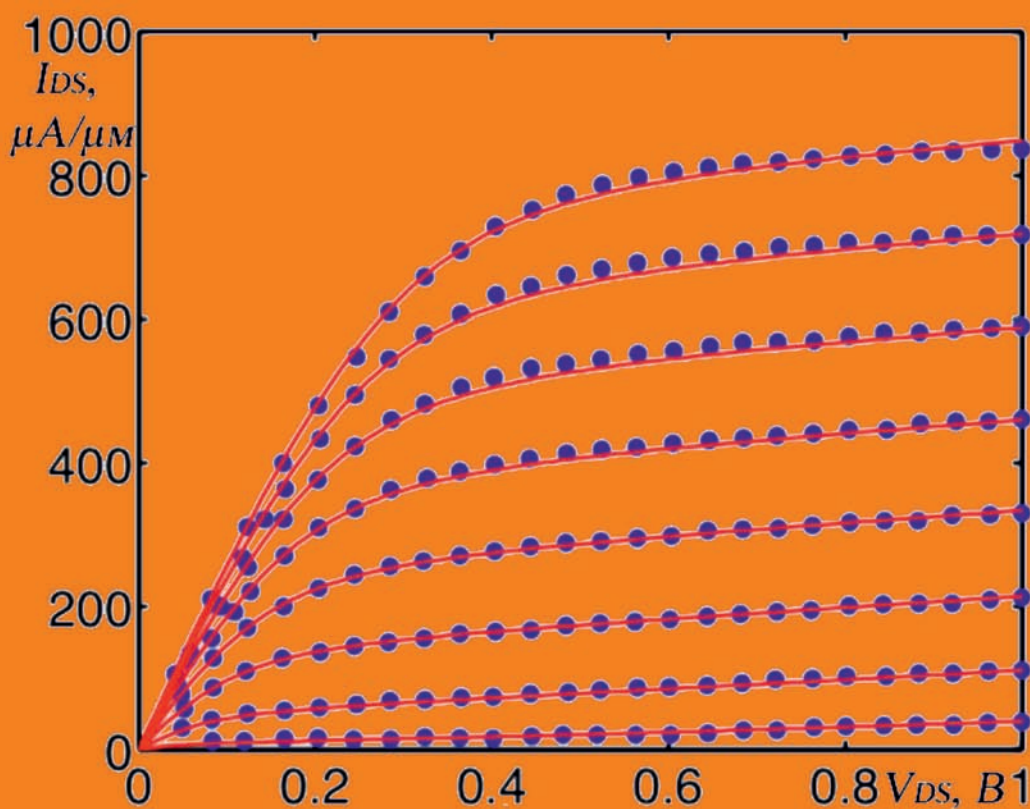


# Ю. А. Кругляк

## Физика и моделирование нанотранзисторов



Одесса  
ТЭС  
2018

**Ю. А. Кругляк**

**Физика и моделирование нанотранзисторов**

Одесса  
ТЭС  
2018

УДК 621.382 PACS: 72.80.Ey, 85.30.-z, 85.30.De, 85.30.Tv, 85.40.-e  
К 840 Кругляк, Юрий Алексеевич

**Физика и моделирование нанотранзисторов/**

**Ю. А. Кругляк. – Одесса: ТЭС, 2018 – 314 стр.; 142 рис.; 210 лит.**

Книга посвящена физике и моделированию полевых транзисторов с акцентом на короткоканальные микро- и современные нанотранзисторы типа МДП-транзистора с изолированным затвором MOSFET, полевого транзистора с высокой подвижностью электронов III-V HEMT, транзисторов на чрезвычайно тонкой КНИ-подложке ETSOI, в том числе с затворами DG и GAA, с каналами NW и CNT, транзисторов FinFET, однако, развитые подходы применимы и к транзисторам типа JFET, MESFET, BJT и им подобным. Рассматривается физическое устройство, метрика и управление FET, теория MOSFET дается сначала в традиционном изложении, затем формулируются начала модели виртуального истока, приближение истощения на фоне уравнение Пуассона, обсуждается подвижный электронный заряд в массивной структуре MOS и в исключительно тонкой SOI, из 2D электростатики MOS подробно рассматриваются DIBL, геометрическое экранирование, емкостная модель, масштабирование транзисторов, их пробой.

Во второй части достаточно подробно излагается обобщенная транспортная модель Ландауэра – Датта – Лундстрема сначала применительно к 2D баллистическим MOSFET, затем баллистическая модель объединяется с моделью виртуального истока, качественно обсуждается рассеяние электронов в канале проводимости и строится модель прохождения MOSFET, которая затем объединяется с моделью виртуального истока в финальный, наиболее современный ныне вариант модели MVS-прохождения, иллюстрируемой примерами обработки экспериментальных выходных характеристик нанотранзисторов ETSOI MOSFET и III-V HEMT.

В заключение обсуждаются фундаментальные пределы и ограничения Si FET, квантовый транспорт в MOSFET с длиной канала менее 10 нм, упрощающие допущения и неясные вопросы в модели MVS-прохождения, а также полевой транзистор рассматривается с позиций нанoeлектроники «снизу – вверх». Книга написана в основном по лекциям Марка Лундстрема «Fundamentals of Nanotransistors», прочитанных в 2016 году: [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).

Книга предназначена для специалистов в области короткоканальной микроэлектроники и нанoeлектроники, а также физики полупроводников.

**ISBN 978-617-7711-06-2**

## Содержание

|  |    |
|--|----|
| Глава 1. Устройство, метрика и управление MOSFET   | 9  |
| 1.1. Введение  | 9  |
| 1.2. Из истории микро- и нанoeлектроники   | 10 |
| 1.3. О физике нанотранзисторов   | 12 |
| 1.4. Физическая структура MOSFET   | 14 |
| 1.5. Вольт-амперные характеристики   | 18 |
| 1.6. Метрика MOSFET  | 20 |
| 1.7. Управление транзистором   | 24 |
| 1.8. Вольт-амперная характеристика MOSFET с позиций модели термоэлектронной эмиссии                        | 31 |
| Литература   | 39 |
| Глава 2. Теория MOSFET в традиционном изложении, начала модели виртуального истока и приближение истощения | 41 |
| 2.1. Введение  | 41 |
| 2.2. Ток, заряд и скорость   | 41 |
| Линейная область   | 42 |
| Область насыщения: скорость насыщения  | 43 |
| Область насыщения: классическая область отсечки  | 44 |
| Вычисление целевой ВАХ: от линейной части до квадратичной области насыщения                                | 46 |
| 2.3. Изменение электрического поля вдоль канала  | 47 |
| 2.4. Насыщение скорости и насыщение тока стока   | 48 |
| 2.5. Метрика нанотранзисторов  | 50 |
| 2.6. Начала модели виртуального истока   | 51 |
| Скорость в канале и напряжение на стоке  | 52 |
| 2.7. Учет терминальных сопротивлений   | 54 |
| 2.8. Приближение истощения и уравнение Пуассона  | 58 |
| Энергия и изгиб зон  | 59 |
| Уравнение Пуассона – Больцмана   | 66 |
| Приближение истощения  | 67 |
| Инверсия   | 70 |
| Влияние подложки   | 71 |
| Литература   | 78 |

|  |     |
|--|-----|
| Глава 3. Напряжение на затворе и поверхностный потенциал,<br>подвижный электронный заряд в массивной структуре MOS и<br>в исключительно тонкой SOI | 79  |
| 3.1. Введение  | 79  |
| 3.2. Напряжение на затворе и поверхностный потенциал   | 80  |
| Пороговое напряжение   | 85  |
| Емкость затвора  | 86  |
| Приближенная зависимость напряжения на затворе от<br>поверхностного потенциала   | 90  |
| 3.3. Подвижный электронный заряд: массивная структура MOS  | 93  |
| Подвижный электронный заряд  | 94  |
| Подвижный заряд ниже порогового напряжения   | 96  |
| Подвижный заряд выше порогового напряжения   | 98  |
| Еще о поверхностном потенциале и напряжении на затворе   | 102 |
| 3.4. Подвижный электронный заряд: исключительно тонкий SOI   | 104 |
| Подвижный электронный заряд  | 104 |
| Подвижный электронный заряд ниже порогового напряжения   | 111 |
| Подвижный электронный заряд выше порогового напряжения   | 114 |
| Ход поверхностного потенциала с изменением напряжения<br>на затворе  | 118 |
| Литература   | 122 |
| Глава 4. 2D электростатика MOS и модель виртуального истока  | 123 |
| 4.1. Введение  | 123 |
| 4.2. 2D электростатика MOS   | 124 |
| 2D Уравнение Пуассона  | 125 |
| Спад порогового напряжения и <i>DIBL</i>   | 126 |
| Геометрическое экранирование   | 128 |
| 2D Электростатика: емкостная модель  | 131 |
| 4.3. Масштабирование транзисторов  | 136 |
| 4.4. Пробой транзисторов   | 141 |
| 4.5. Модель виртуального истока  | 145 |
| Подпороговая область   | 147 |
| Из под порога в область выше порога  | 152 |
| Литература   | 155 |

|   |     |
|---|-----|
| Глава 5. Транспортная модель Ландауэра – Датта – Лундстрома и баллистические MOSFET                             | 157 |
| 5.1. Введение   | 157 |
| 5.2. Обобщенная модель электронного транспорта  | 158 |
| Большие и малые напряжения  | 160 |
| Коэффициент прохождения   | 162 |
| Моды проводимости   | 167 |
| Квант проводимости  | 169 |
| Плотность электронов  | 170 |
| 5.3. Баллистические MOSFET  | 175 |
| MOSFET как устройство с наноканалом проводимости  | 176 |
| Линейная область  | 178 |
| Область насыщения   | 178 |
| От линейной области к насыщению   | 179 |
| Выражения для тока через электронный заряд  | 179 |
| Литература  | 187 |
| Глава 6. Баллистическая скорость впрыскивания и объединение баллистической модели с моделью виртуального истока | 189 |
| 6.1. Введение   | 189 |
| 6.2. Баллистическая скорость впрыскивания   | 189 |
| Зависимость скорости от напряжения на стоке   | 190 |
| Насыщение скорости в баллистических MOSFET  | 191 |
| Баллистическая скорость впрыскивания  | 195 |
| 6.3. Объединение баллистической модели с моделью виртуального истока  | 199 |
| Резюме баллистической модели  | 199 |
| Резюме модели виртуального стока  | 200 |
| Объединенная модель   | 201 |
| Линейная область: баллистическая модель и модель виртуального истока  | 202 |
| Область насыщения: баллистическая модель и модель виртуального истока   | 204 |
| Сравнение с экспериментальными результатами   | 206 |
| Литература  | 212 |

|  |     |
|--|-----|
| Глава 7. Рассеяние электронов и модель прохождения MOSFET                                      | 213 |
| 7.1. Введение  | 213 |
| 7.2. Рассеяние электронов и коэффициент прохождения  | 214 |
| Характеристические времена и длины   | 214 |
| Скорости рассеяния и энергия   | 216 |
| Коэффициент прохождения  | 220 |
| Средняя длина свободного пробега назад   | 224 |
| 7.3. Модель прохождения MOSFET   | 228 |
| Резюме баллистической модели MOSFET  | 229 |
| Линейная область   | 230 |
| Область насыщения  | 230 |
| От линейного тока до насыщения   | 231 |
| Выражение для тока через электронный заряд   | 231 |
| Зависимость коэффициента прохождения от напряжения на стоке                                    | 235 |
| Литература   | 241 |
| Глава 8. Объединение модели прохождения и модели виртуального истока:                          |     |
| модель MVS-прохождения   | 243 |
| 8.1. Введение  | 243 |
| Резюме модели прохождения  | 243 |
| Резюме модели виртуального истока  | 244 |
| 8.2. Объединение модели прохождения и модели виртуального истока                               | 246 |
| Линейная область   | 246 |
| Область насыщения  | 248 |
| 8.3. Транспорт в нанотранзисторах с позиции MVS/прохождения                                    | 252 |
| Обзор модели MVS/прохождения   | 252 |
| 8.4. ETSOI MOSFET и III-V HEMT   | 257 |
| Вписывание MVS/прохождения в экспериментальные данные  | 259 |
| Результаты анализа экспериментальных данных ETSOI MOSFET и III-V HEMT в модели MVS/прохождения | 260 |
| Анализ линейной области  | 262 |
| Анализ области насыщения   | 265 |
| Анализ перехода линейной области к насыщению   | 266 |
| Литература   | 269 |

|  |     |
|--|-----|
| Глава 9. Фундаментальные пределы и ограничения   | 271 |
| 9.1. Введение  | 271 |
| 9.2. Фундаментальные пределы   | 271 |
| 9.3. Квантовый транспорт в MOSFET с длиной канала меньше 10 нм                             | 275 |
| 9.4. Упрощающие допущения в модели прохождения   | 277 |
| Вывод формулы Ландауэра для тока   | 279 |
| Неидеальные контакты   | 282 |
| Критическая длина при рассеянии назад  | 282 |
| Зависимость средней длины свободного пробега и подвижности от<br>длины канала проводимости | 284 |
| Роль рассеяния   | 286 |
| Вырождение носителей тока  | 288 |
| Плотность заряда и транспорт   | 289 |
| Литература   | 294 |
| Глава 10. Полевой транзистор с позиций нанoeлектроники «снизу – вверх»                     | 297 |
| 10.1. Введение   | 297 |
| 10.2. Соотношение между током и напряжением  | 298 |
| 10.3. Почему ток насыщается?   | 299 |
| 10.4. Роль заряжания проводника  | 301 |
| 10.5. Квантовая емкость  | 304 |
| 10.6. Электростатический «выпрямитель»   | 308 |
| 10.7. Расширенная модель проводника  | 310 |
| 10.8. Уравнение диффузии и расширенная модель проводника                                   | 311 |
| Литература   | 313 |





## Глава 1. Устройство, метрика и управление MOSFET

### 1.1. Введение

Изобретение транзистора считают важнейшим открытием 20 века. Транзисторы используются в электронной схеме практического любого электронного прибора. Миллиарды транзисторов содержатся в наших смартфонах и планшетах, в настольных компьютерах и суперкомпьютерах, во всех разнообразных гаджетах, без которых мы уже не можем представить себе 21 век. Нанотранзистор представляет большой интерес и сам по себе (рис. 1).

Электроны в MOSFET движутся от истока  $S$  по проводящему каналу и выходят из стока  $D$ . Соответствующий ток  $I_D$  контролируется затвором  $G$ , изолированным от канала проводимости диэлектрической прослойкой толщиной уже менее 2 нм. Если длина канала проводимости к концу второго тысячелетия составляла около 100 нм, то сегодня – уже приближается к 10 нм.

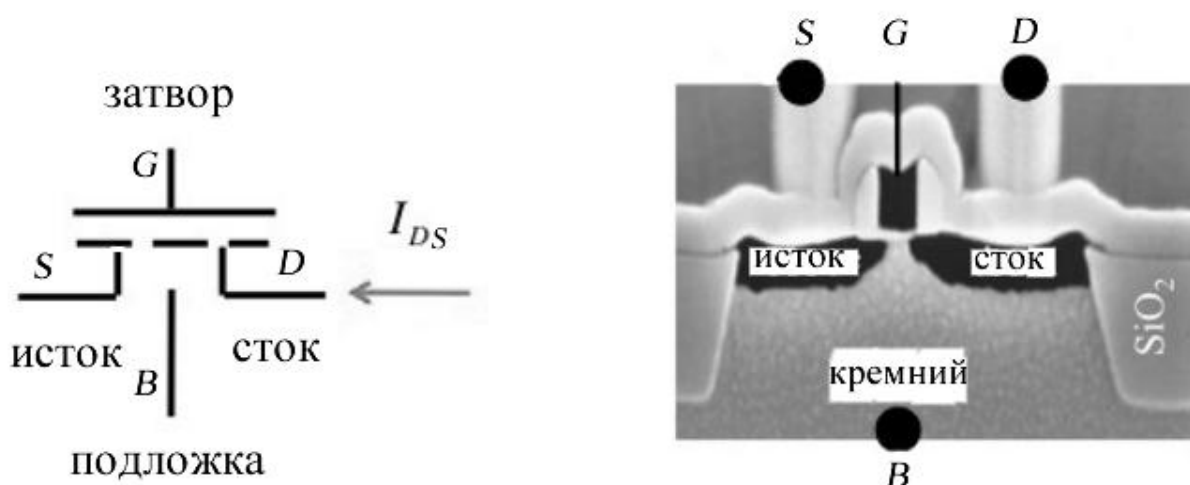


Рис. 1. Si-MOSFET: Слева показано общепринятое в англоязычной литературе обозначение транзистора с расшифровкой контактов – Исток  $S$ /Source, Сток  $D$ /Drain, Затвор  $G$ /Gate и Подложка  $B$ /Body; пунктирной линией слева показан токопроводящий канал. Справа показана микрограмма сечения MOSFET, полученная на сканирующем электронном микроскопе (Texas Instruments, 1997), на которой хорошо видны все контакты и диэлектрическая прослойка, изолирующая затвор  $G$  от канала проводимости.

Функционирование MOSFET интересно с познавательной точки зрения. На рис. 2 показана вольт-амперная характеристика (ВАХ) MOSFET.

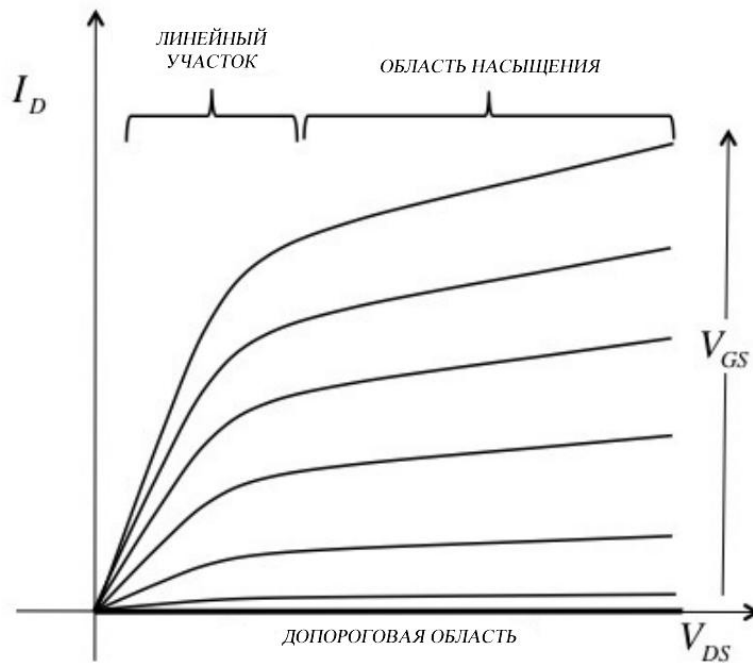


Рис. 2. Вольт-амперные характеристики  $n$ -MOSFET: По вертикальной оси откладывается ток  $I_D$ , текущий от стока  $D$  к истоку  $S$ , а по горизонтальной оси – напряжение между ними  $V_{DS}$ . Разные характеристики соответствуют разным значениям напряжения  $V_{GS}$  на затворе  $G$ . Показаны два важных участка послепороговой области – линейный участок и область насыщения. Максимальное напряжение, прилагаемое между затвором  $G$  и стоком  $D$  есть напряжение питания  $V_{DD}$ . Небольшой ток утечки в допороговой области на линейном участке  $I_D$  не заметен.

Такое электронное устройство как MOSFET полезно в качестве элемента в различных электронных схемах. Они могут выполнять роль цифрового коммутатора (ON/OFF) или, например, аналогового усилителя входных сигналов.

## 1.2. Из истории микро- и нанoeлектроники

Электронная схема любого электронного прибора содержит различные электронные компоненты и устройства. Сопротивления, конденсаторы и дроссели – это очень простые электронные компоненты, однако, электронные схемы нуждаются в нелинейных устройствах, простейшим из которых может служить такой электронный элемент как диод, обладающий различной проводимостью в зависимости от направления электрического тока. Первым применением полупроводниковых диодов в начале 20 века стало обнаружение

радиосигналов. Эти кристаллические детекторы были очень капризны и вскоре были заменены вакуумными ламповыми детекторами, содержащими катод прямого накала и металлический анод, при подаче на который положительного потенциала электроны с катода притягивались к аноду и шел ток.

Вскоре появились вакуумные триоды, а затем и пентоды. При помещении между катодом и анодом металлической сетки ток можно было контролировать путем подачи небольшого потенциала на сетку, усиливая тем самым детектируемый сигнал. Широкое применение вакуумных ламп революционизировало технику связи и привело к созданию первых электронных вычислительных машин, однако, с вакуумными лампами было несколько серьезных проблем – они были громоздкими, хрупкими и слишком энергоемкими.

В 1925 г. львовянин по происхождению Юлий Лилиенфельд [1], позже американский физик запатентовал твердотельную замену вакуумному триоду MESFET (Metal-Semiconductor Field-Effect Transistor – Полевой транзистор с контактом металл-полупроводник (диодом Шотки) в качестве затвора). Его твердотельный MESFET был меньше вакуумного триода по размерам, был более надежный и потреблял меньше энергии. Полупроводниковая технология в то время была еще недостаточно развита чтобы конкурировать с промышленным производством вакуумных ламп, однако, к концу II Мировой войны в Телефонной лаборатории Белл уже начались серьезные технологические исследования по твердотельной замене вакуумных ламп [2]. В результате, в декабре 1947 года появился точечный биполярный транзистор. Со временем все технологические проблемы, связанные с производством полевых транзисторов были преодолены, и сегодня MOSFET является основным полевым транзистором в любых электронных устройствах [3].

К 1960 году освоили технологию производства нескольких транзисторов в виде единой электронной схемы на кремниевой подложке. Гордон Мур в 1965 году заметил, что число транзисторов в таких интегральных схемах удваивается с каждой новой технологической генерацией (закон Мура) [4] – вначале ежегодно, затем каждые полтора года. Он предсказал, что такое удвоение числа транзисторов в чипе сохранится и далее в течение еще некоторого времени, но даже он не мог представить себе, что подобное удвоение продлится более 50 лет [5].

Удвоение числа транзисторов на интегральной схеме сопровождалось неуклонным уменьшением их размеров. Вначале размеры транзисторов

замерялись в микрометрах, и электроника стала именоваться микроэлектроникой. Физики разработали простые математические модели транзисторов [6 – 9], которые надежно описывали все функции транзистора и были достаточны для схемных и системных дизайнеров. Уточнение и дальнейшее развитие этих моделей происходило по мере совершенствования микросистемных технологий [10, 11]. С каждой технологической генерацией размеры транзисторов уменьшались в  $\sqrt{2}$  раз, так что площадь, занимаемая транзистором, уменьшалась в 2 раза, и число транзисторов в чипе удваивалось. В районе 2000 года длина канала проводимости в транзисторах достигла 100 нм, микроэлектроника превратилась в наноэлектронику, и нанотранзистор стал знаменем нанотехнологической революции. К сегодняшнему дню длина канала проводимости уменьшилась в 10 раз и вопрос сейчас звучит следующим образом: «Насколько современная технология позволит канал проводимости сделать менее 10 нм?».

После преодоления нанометрового порога замеры всех электрофизических характеристик нанотранзисторов существенно не изменились. Однако, кое-что существенное таки изменилось – другой стала физика переноса электронов от истока к стоку. Изложению этой новой физики транспортных явлений в нанотранзисторах и посвящается вашему вниманию книга.

### 1.3. О физике нанотранзисторов

Работа подавляющего большинства разнообразных транзисторов контролируется высотой энергетического барьера путем изменения прикладываемого напряжения на затворе. Этот барьер препятствует движению электронов от истока к стоку. Путем изменения напряжения  $V_{GD}$  между затвором  $G$  и стоком  $D$  можно менять высоту этого энергетического барьера и таким образом контролировать ток по каналу проводимости транзистора.

Анализ работы MOSFET обычно начинается с уравнения

$$I_{DS} = W |Q(V_{GS}, V_{DS})| \langle v \rangle, \quad (1)$$

где  $W$  – ширина канала проводимости в направлении, перпендикулярном движению электронов,  $Q$  это подвижный электронный поверхностный заряд,  $\langle v \rangle$  есть средняя скорость движения электронов. Построение математической модели работы транзистора сводится к физической трактовке электростатики ( $Q$ ) и транспорта ( $v$ ). В то время как дизайн и трактовка электростатики

транзистора мало изменились за последние несколько десятилетий, физическая природа и характер переноса электронов существенно менялись по мере того, как длина канала проводимости транзистора становилась все меньше и меньше.

Краеугольным камнем традиционной физики полупроводников является уравнение дрейфа-диффузии

$$J_x = n_s q \mu E_x + q \bar{D} \frac{dn_s}{dx}, \quad (2)$$

которое утверждает, что плотность тока в однородном полупроводнике в направлении движения электронов пропорциональна электрическому полю, а в его отсутствие ток обеспечивается электронами, диффундирующими навстречу градиенту концентрации электронов, где  $n_s$  есть плотность электронов,  $q$  – величина заряда электрона,  $\mu$  – подвижность электронов,  $\bar{D}$  – коэффициент диффузии. Хотя большинство учебников по физике полупроводников начинается с уравнения (2), нет уверенности в том, что предположения, которые делаются при выводе этого уравнения, остаются в силе применительно к нанопроводникам. Действительно, полуклассическое [12] и квантовое [13] моделирование показали, что транспорт электронов в наноразмерных проводниках хотя и оказывается весьма сложным явлением, но вместе с тем и более простым по сравнению с транспортом в массивных проводниках.

Новый подход к описанию транспорта электронов в виде модели упругого резистора был предложен Ландауэром еще в 1957 году [14], задолго до её экспериментального подтверждения в нанопроводниках. В этой модели ток вместо (2) вычисляется по формуле

$$I = \frac{2q}{h} \int T(E) M(E) (f_1(E) - f_2(E)) dE, \quad (3)$$

где  $M(E)$  – число мод (каналов) при значении энергии  $E$ , доступных для проводимости,  $f_1(E)$  и  $f_2(E)$  – равновесные функции Ферми на истоке и стоке,  $T(E)$  – коэффициент прохождения. Если допустимо сравнивать канал проводимости в нанотранзисторе с широкополосным автомобильным шоссе, то число мод аналогично числу полос на шоссе, а коэффициент прохождения лежит в области между нулем и единицей, т. е. является мерой вероятности того, что электрон, впрыснутый истоком, окажется на стоке. Для массивных проводников уравнение (3) сводится к (2) и оно пригодно для

нанопроводников, применительно к которым недостаточно ясно как пользоваться уравнением (2).

Протекание тока рассматривалось до сих пор только полуклассически. Другими словами, электроны рассматривались лишь как частицы, а их квантово-механическая природа учитывалась опосредовано через зонную структуру проводящего канала введением понятия эффективной массы электрона. Однако, по мере дальнейшего уменьшения длины канала проводимости все настоятельнее требуется учесть квантовую природу электронов в явном виде. Можно ожидать, что квантово-механические эффекты станут играть важную роль когда потенциальная энергия начнет меняться быстро в масштабе де-бройлевской длины волны электрона. Простая оценка де-бройлевской длины волны электронов в термически равновесном образце Si дает приблизительно 10 нм, что не намного меньше реально достигнутой на практике длины канала проводимости при массовом производстве нанотранзисторов. В течение последних двух десятилетий были разработаны надежные и достаточно мощные модели квантового транспорта электронов применительно к нанопроводникам [13, 15 – 17]. Чем больше удаляется длина канала проводимости от 10 нм в сторону уменьшения, тем все важнее становится учитывать квантово-механическую природу транспорта электронов, однако, при длине канала проводимости более 10 нм, как мы убедимся позже, полуклассическая модель все еще работает достаточно надежно [18].

#### 1.4. Физическая структура MOSFET

Наша конечная цель – объяснить и математически описать физику процессов, происходящих в транзисторе, т. е. увязать токи, протекающие в транзисторе, с электрическими потенциалами, подаваемыми на контакты (терминалы) транзистора (рис. 1). Вначале полезно рассмотреть транзистор как «черный ящик», формализовано, не вникая в физику происходящих в транзисторе процессов.

Ток, текущий между терминалами  $S$  и  $D$ , контролируется электрическим потенциалом, подаваемым на терминал  $G$  (рис. 1). В некоторых вариантах транзисторов этот терминал может использоваться для подачи тока. Иногда используется и терминал  $B$ . Разработано много вариантов транзисторов [19], но все они имеют три или четыре терминала. Терминалы могут называться по-разному в зависимости от типа транзистора.

Мы сосредоточимся на кремниевых полевых транзисторах Si-MOSFET и на полевых транзисторах с высокой подвижностью электронов III-V HEMT (High Electron Mobility Transistor), однако, развиваемые далее подходы применимы ко всем типам полевых транзисторов.

Сначала рассмотрим устройство  $n$ -канального Si MOSFET. В кремниевой пластине  $p$ -типа формируется канал проводимости  $n$ -типа с двумя омическими контактами истока  $S$  и стока  $D$ . Контактные области  $S$  и  $D$  сильно  $n$ -допированы. По проводящему каналу  $n$ -типа движется поток электронов от истока к стоку. В верхней части пластины изготавливается омический контакт затвора  $G$ . Аналогичный контакт изготавливают в нижней части подложки  $B$ .

В полевом транзисторе  $n$ -Si MOSFET канальный ток управляется электрическим полем, возникающим при приложении электрического напряжения между затвором и стоком.

На рис. 3 показано сечение и вид сверху  $n$ -Si MOSFET.

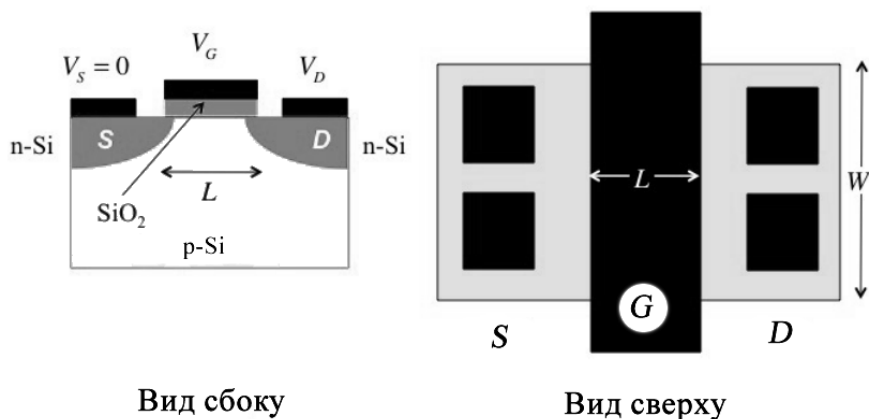


Рис. 3.  $n$ -Канальный Si MOSFET: вид сбоку и сверху. Схематически показаны контакты истока и стока (черные квадраты) и затвора (черный прямоугольник).

Под металлическим контактом затвора находится тонкий слой SiO<sub>2</sub>, а под ним – узкая зона  $p$ -типа. Длина канала проводимости  $L$  является критическим параметром транзистора, она же определяет фактический размер транзистора на чипе и скорость срабатывания транзистора: чем меньше  $L$ , тем больше эта скорость. Ширина канала проводимости  $W$  определяет величину тока по каналу. Для каждой технологической генерации производства транзисторов значения  $L$  и  $W$  проектируются таким образом, чтобы функционирование транзисторов было максимально эффективным для  $L \geq L_{min}$ , некоторого минимального значения длины канала проводимости.



В  $n$ -канальном MOSFET ток обеспечивается электронами в зоне проводимости. Как показано на рис. 4, возможно создать аналогичный транзистор  $p$ -MOSFET, проводимость которого также обеспечивается электронами, но в валентной зоне, что мысленно визуализируется с помощью «дырок» в валентной зоне.  $p$ -Канальный MOSFET конструируется на  $n$ -Si подложке. Контактные области  $S$  и  $D$  сильно  $p$ -допированы. Обратим внимание, что  $V_{DS} < 0$  для  $p$ -MOSFET и что требуется  $V_{GS} < 0$  для перевода  $p$ -MOSFET в режим «ON». Ток течет из стока, а не в сторону истока как в  $n$ -MOSFET. Современная электроника конструируется по технологии CMOS (Complementary MOS – КМОП / комплементарная МОП-структура), в которой каждый  $n$ -транзистор спарен с  $p$ -транзистором.

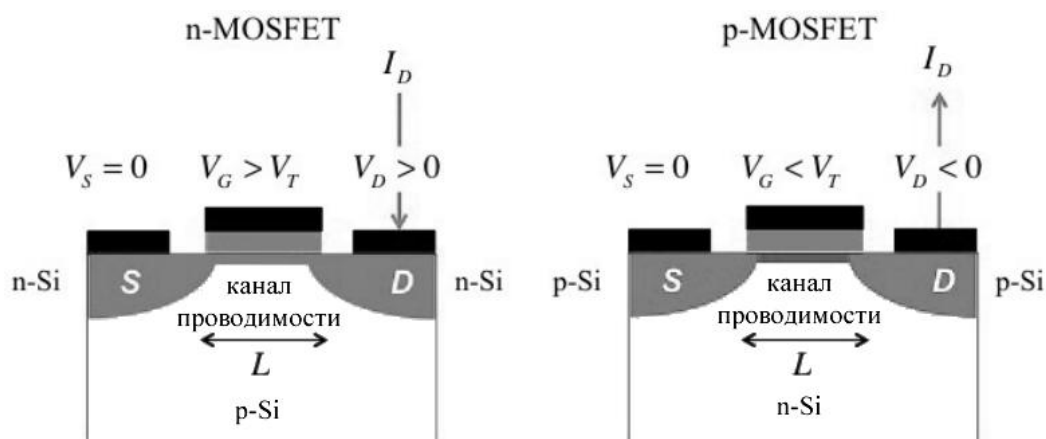


Рис. 4. Сравнение  $n$ -MOSFET и  $p$ -MOSFET.

Обратим внимание, что для  $n$ -MOSFET  $V_{DS}, V_{GS} > 0$ , а для  $p$ -MOSFET  $V_{DS}, V_{GS} < 0$ .

Ток  $I_D$  течет от  $D$  к  $S$  в случае  $n$ -MOSFET и от  $S$  к  $D$  в случае  $p$ -MOSFET.

В прикладной схемотехнике транзисторы обычно конфигурируются воспринимать входное напряжение и работать при определенном выходном напряжении. Входное напряжение измеряется между двумя входными терминалами, а выходное напряжение – между двумя выходными терминалами. Входной ток это ток, который входит в один из двух входных терминалов и выходит из другого входного терминала, а выходной ток это ток, который входит в один из двух выходных терминалов и выходит из другого выходного терминала. В схемотехнике, по договоренности, ток считается положительным, если он входит в терминал, так что ток истока  $I_D$  в  $n$ -MOSFET (рис. 4)

положительный, а ток истока  $I_D$  в  $p$ -MOSFET отрицательный. Поскольку мы располагаем только тремя терминалами (терминал подложки  $B$  особый, он позволяет настраивать операционные характеристики транзистора), то один из терминалов должен быть общим либо со входом, либо с выходом. Есть три возможных схемы включения транзистора – с общим истоком, с общим стоком и с общим затвором.

На рис. 5 показан  $n$ -MOSFET, включенный по схеме с общим истоком.

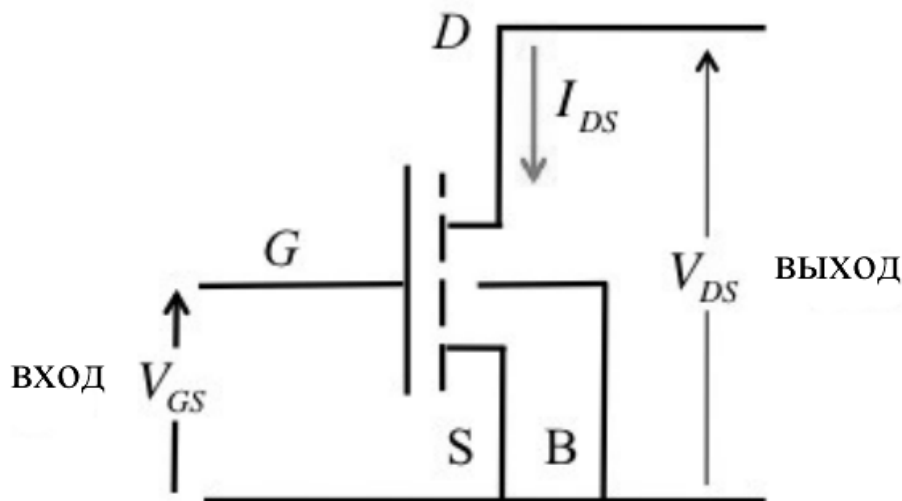


Рис. 5.  $n$ -MOSFET, включенный по схеме с общим истоком.

Входным напряжением является  $V_{GS}$ , а выходным –  $V_{DS}$ .

Выходным током является  $I_{DS}$ , а ток затвора обычно пренебрежимо мал, так что входной постоянный ток предполагается нулевым.

В этом случае постоянным выходным током является ток от стока к истоку  $I_{DS}$ , а постоянным выходным напряжением является напряжение между стоком и истоком  $V_{DS}$ . Постоянным входным напряжением является напряжение между затвором и истоком  $V_{GS}$ . Этот постоянный входной ток обычно очень мал и им, как правило, пренебрегают.

Нас будут в первую очередь интересовать вольт-амперные характеристики транзисторов. Поясним используемую терминологию. Два типа ВАХ особенно интересны: выходные и передаточные характеристики. Для построения выходных ВАХ откладывают выходной ток  $I_{DS}$  против выходного напряжения  $V_{DS}$  при постоянном значении входного напряжения  $V_{GS}$ . В случае передаточных ВАХ откладывают выходной ток  $I_{DS}$  против входного напряжения  $V_{GS}$  при

постоянном выходном напряжении  $V_{DS}$ . Рассмотрим подробнее оба типа характеристик и поясним используемую терминологию.

### 1.5. Вольт-амперные характеристики

На рис. 6 показана вольт-амперная характеристика идеального резистора. В этом случае ток пропорционален приложенному напряжению  $I = V/R$ , где  $R$  есть сопротивление резистора.

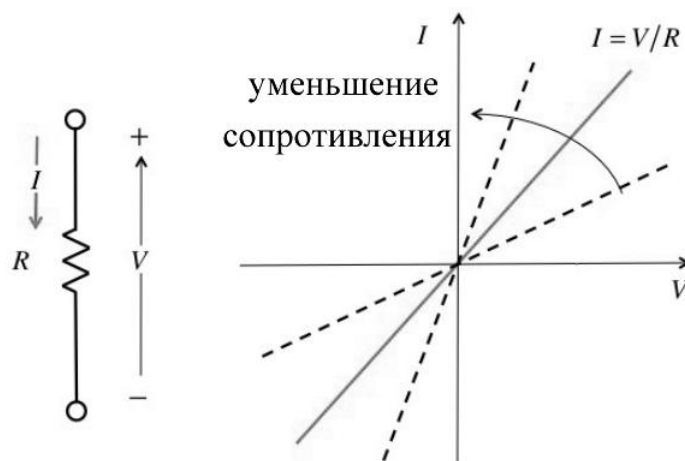


Рис. 6. Вольт-амперная характеристика идеального резистора.

На рис. 7 показана вольт-амперная характеристика источника тока.

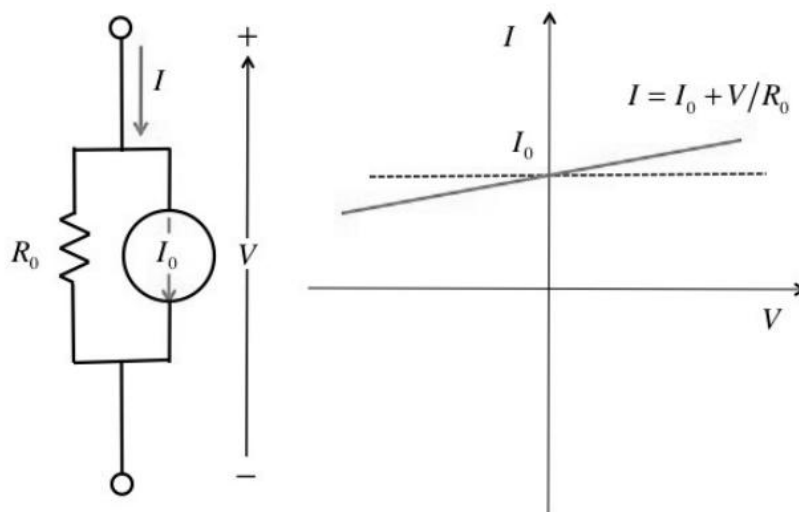


Рис. 7. Вольт-амперная характеристика источника тока. Когда ток не зависит от напряжения, его величина постоянна  $I_0$  (пунктирная линия). В реальных источниках тока наблюдается зависимость его от напряжения, что эквивалентно включению резистора  $R_0$  параллельно источнику тока  $I_0$ .

Выходные характеристики транзистора выглядят как ВАХ для резистора при небольших значениях  $V_{DS}$  и как ВАХ источника тока при больших значениях выходного напряжения  $V_{DS}$ .

Выходные характеристики  $n$ -MOSFET показаны на рис. 2. Каждая характеристика в этом семействе ВАХ соответствует определенному значению  $V_{GS}$ . Для  $V_{DS}$ , меньшего некоторого критического значения  $V_{DSAT}$ , ток пропорционален напряжению. В этой линейной (омической) области транзистор ведет себя как резистор, сопротивление которого определяется входным напряжением  $V_{GS}$ .

В области  $V_{DS} > V_{DSAT}$  (область насыщения) транзистор ведет себя как источник тока, величина которого определяется значением  $V_{GS}$ . Ток немного увеличивается с ростом  $V_{DS}$ , что эквивалентно тому, что этот источник тока характеризуется конечным значением выходного сопротивления  $r_d$ .

Третья область функционирования транзистора это допороговая область, для которой  $V_{GS}$  меньше некоторого критического значения, порогового напряжения  $V_T$ . Для  $V_{GS} < V_T$  ток  $I_{DS}$  очень небольшой и не виден на линейном участке.

На рис. 8 сравниваются выходные и передаточные ВАХ для  $n$ -MOSFET.

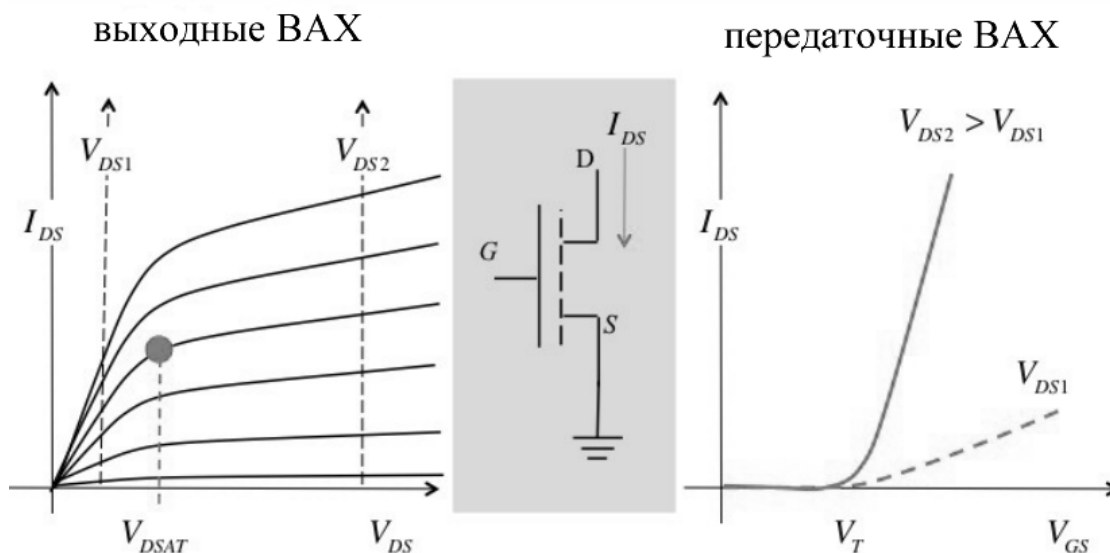


Рис. 8. Сравнение выходных ВАХ  $n$ -MOSFET, включенного по схеме с общим истоком (слева) с передаточными ВАХ этого же транзистора (справа). График, помеченный  $V_{DS1}$  на передаточных ВАХ, соответствует небольшому значению  $V_{DS}$  на выходных ВАХ слева, а помеченный  $V_{DS2}$  – большому значению  $V_{DS}$ .

Выходные характеристики показаны слева. Зафиксируем небольшое значение  $V_{DS}$ , равное  $V_{DS1}$ , и прогоним ток  $I_{DS}$  по соответствующим значениям  $V_{GS}$ . Это нам даст передаточный график  $V_{DS1}$  среди передаточных ВАХ справа. Если возьмем большое значение  $V_{DS}$ , равное  $V_{DS2} > V_{DS1}$ , получим передаточную характеристику, помеченную  $V_{DS2}$ . Передаточные характеристики показывают, что при  $V_{GS} < V_T$  ток  $I_{DS}$  очень небольшой.

Далее мы увидим, что график  $\log_{10} I_{DS} \propto V_{GS}$  (рис. 11) используется для анализа тока  $I_{DS}$  в области  $V_{GS} < V_T$  (рис. 8/справа).

## 1.6. Метрика MOSFET

Функционирование MOSFET можно описать с помощью следующих десяти метрических показателей:

Ток  $I_{ON}$ ,  $\mu A/\mu m$

Линейное сопротивление  $R_{ON}$ ,  $Om \cdot \mu m$

Выходное сопротивление  $r_d$ ,  $Om \cdot \mu m$

Транспроводимость  $g_m$ ,  $\mu Cm/\mu m$

Ток  $I_{OFF}$ ,  $\mu A/\mu m$

Допороговый разброс (Subthreshold Swing)  $SS$ ,  $mB/декада$

Понижение барьера, индуцированное стоком (Drain-Induced Barrier Lowering),  $DIBL$ ,  $mB/B$

Пороговые напряжения  $V_T(\text{lin}), V_T(\text{sat})$ ,  $B$

Напряжение насыщения на стоке  $V_{DSAT}$ ,  $B$

Указанные единицы измерения используются чаще всего. Например, транспроводимость обычно не приводится в  $Cm/m$ , а дается в единицах  $mCm/mкм$  или  $mлCm/mм$ .

Некоторые из метрических показателей транзисторов можно проиллюстрировать на выходных характеристиках  $n$ -MOSFET, включенного по схеме с общим истоком (рис. 9). Ток  $I_{ON}$  – это максимальный ток стока  $I_{DS}|_{V_{GS}=V_{DS}=V_{DD}}$ . Ток  $I_{DS}$  обычно измеряется в  $\mu A/\mu m$  (или  $mA/\mu m$ ), поскольку он меняется линейно с шириной канала проводимости  $W$ . Линейное сопротивление – это минимальное сопротивление канала проводимости  $R_{ON} = (dI_{DS}/dV_{DS})_{V_{GS}=V_{DD}}^{-1}$ . Выходное сопротивление вычисляется аналогично

$r_d = (dI_{DS}/dV_{DS})_{V_{GS}=V_{DD}}^{-1}$ , но в области насыщения. Транспроводимость  $g_m = (dI_{DS}/dV_{GS})_{V_{DS}=\text{Const}}$  обычно приводится при  $V_{DS} = V_{DD}$ . Чтобы получить истинные значения тока стока и транспроводимости, их нужно умножить еще на ширину  $W(\mu\text{м})$ , а для получения истинных линейного и выходного сопротивлений делим их на ширину  $W(\mu\text{м})$ .

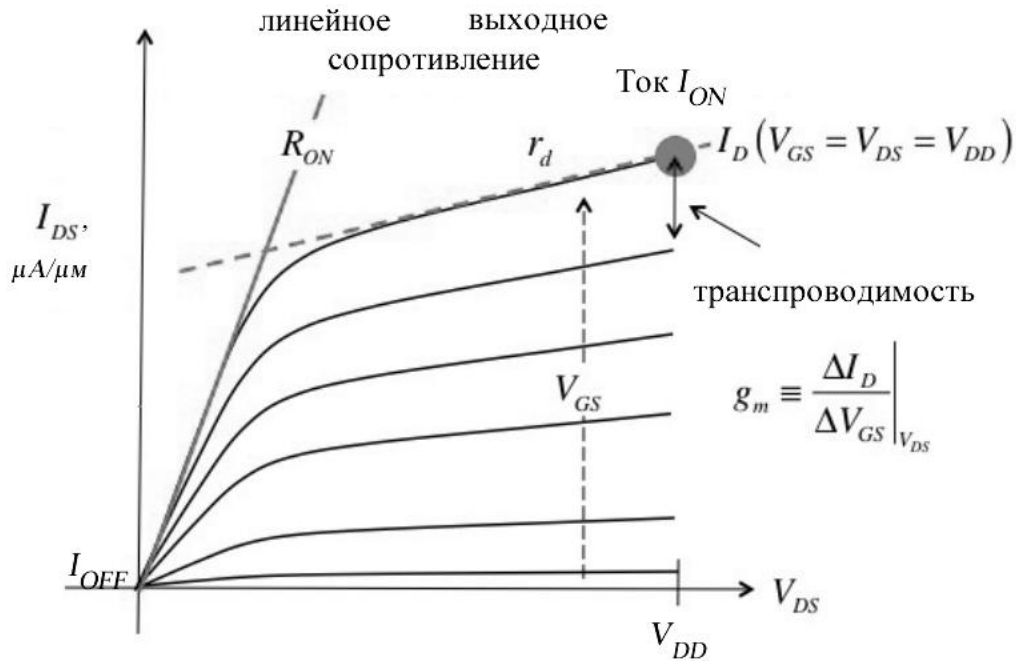


Рис. 9. Выходные характеристики  $n$ -MOSFET, включенного по схеме с общим истоком, с демонстрацией приведенных выше четырех метрик  $I_{ON}$ ,  $R_{ON}$ ,  $r_d$ ,  $g_m$ .

Еще три метрики можно проиллюстрировать на передаточных характеристиках  $n$ -MOSFET, включенного по схеме с общим истоком (рис. 10). Показаны две разные характеристики – для низких напряжений  $V_{DS}$  (линейная область) и для высоких  $V_{DS}$  (область насыщения). Помечен также ток  $I_{ON}$ , показанный на рис. 9. Если определить максимальный наклон обеих характеристик и пересечение соответствующих прямых (они показаны пунктиром) с осью напряжений, то найдем пороговое напряжение  $V_T$ . Обратим внимание, что есть два пороговых напряжения: одно соответствует линейной (омической) области  $V_T(\text{lin})$ , а другое – области насыщения  $V_T(\text{sat})$ , причем  $V_T(\text{sat}) < V_T(\text{lin})$ . Заметим также, что переход  $I_{OFF} \Rightarrow I_{ON}$  происходит постепенно и точка, в которой этот переход завершается, и есть пороговое напряжение  $V_T$ .

Наконец, ток  $I_{OFF} = I_{DS}|_{V_{GS}=0, V_{DS}=V_{DD}}$  фактически также показан на рис. 10, но он слишком незначительный, чтобы просматриваться в этом масштабе.

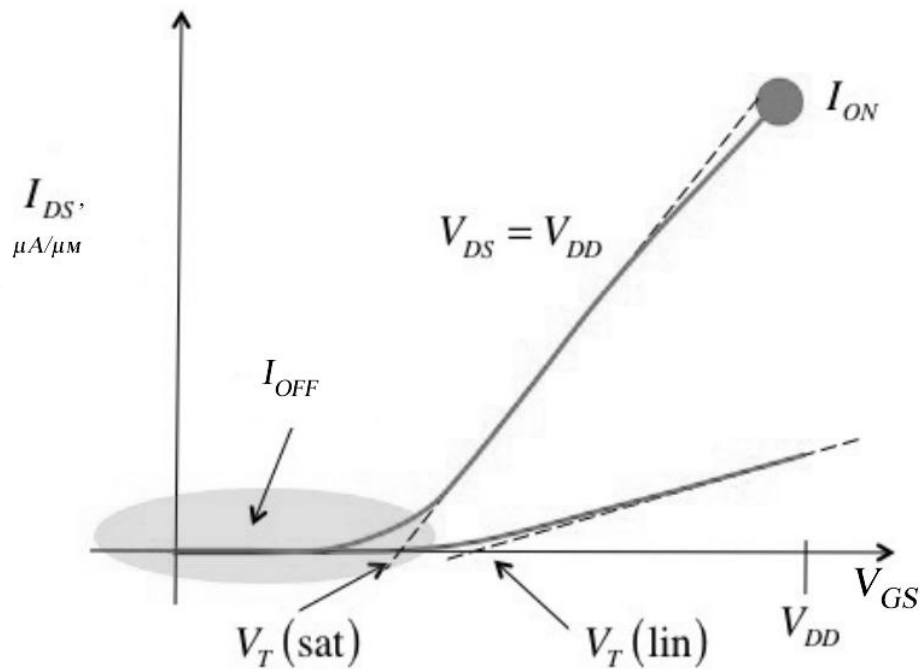


Рис. 10. Передаточные характеристики  $n$ -MOSFET, включенного по схеме с общим истоком, с демонстрацией приведенных выше трех метрик  $I_{OFF}$ ,  $V_T(\text{lin})$ ,  $V_T(\text{sat})$ .

Разобраться с допороговыми характеристиками удобнее, если откладывать выходной ток  $I_{DS}$  в логарифмической шкале, как это показано на рис. 11. Оба тока  $I_{OFF} = I_{DS}|_{V_{GS}=0, V_{DS}=V_{DD}}$  и  $I_{ON} = I_{DS}|_{V_{GS}=V_{DD}, V_{DS}=V_{DD}}$  также показаны на этом рисунке. Допороговый ток MOSFET увеличивается экспоненциально с  $V_{GS}$  (рис. 10). Допороговый разброс  $SS$  дается выражением

$$SS = \left[ d(\log_{10} I_{DS}) / dV_{GS} \right]^{-1} \quad (4)$$

и обычно приводится в  $mB/\text{декада}$ . Допороговый разброс это такое изменение напряжения на затворе (обычно в  $mB$ ), которое приводит к изменению тока  $I_{DS}$  в 10 раз. Чем меньше  $SS$ , тем меньшее напряжение на затворе необходимо для переключения транзистора из режима «OFF» в режим «ON». Позже мы узнаем, что физика подпороговой проводимости предписывает значение  $SS \geq 60 mB/\text{декада}$ . У транзистора хорошего качества допороговый разброс практически одинаков как для передаточных характеристик низкого, так и

высокого напряжения  $V_{DS}$  (рис. 10). Часто наблюдается увеличение  $SS$  с ростом  $V_{DS}$ , что связывают с проявлением  $2D$  электростатики MOS, к детальному рассмотрению которой мы перейдем позже.

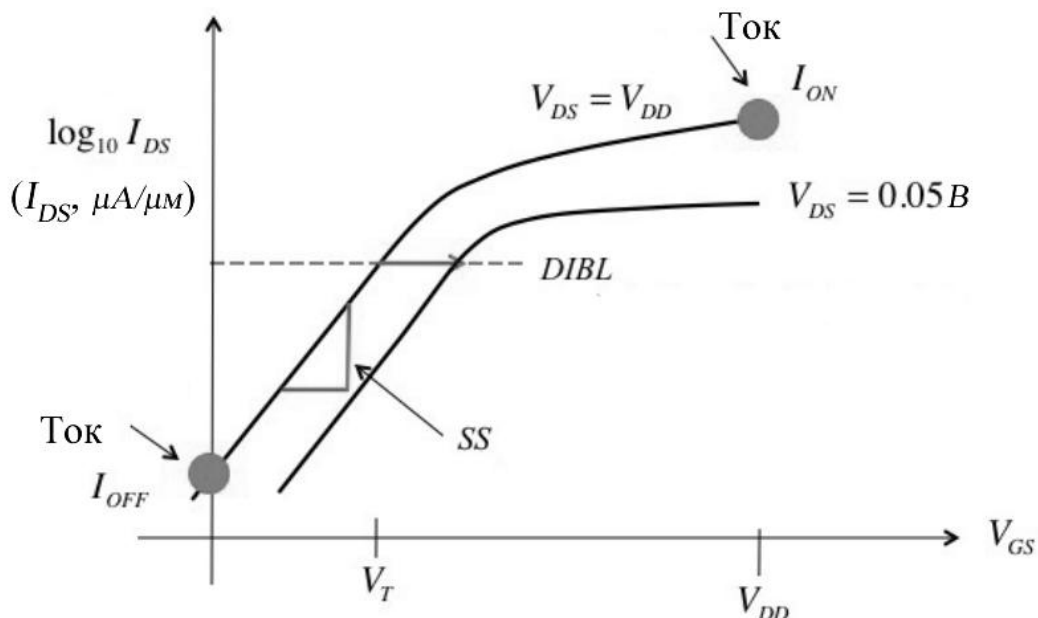


Рис. 11. Передаточные характеристики  $n$ -MOSFET, включенного по схеме с общим истоком, с демонстрацией приведенных выше еще двух метрик  $SS$  и  $DIBL$ . Ток стока  $I_{DS}$  на этом рисунке отложен в логарифмической шкале.

Подпороговые характеристики сдвигаются влево по мере роста стокового напряжения  $V_{DS}$ . Этот сдвиг связывают с эффектом понижения барьера, индуцированного стоком, известным как  $DIBL$ . Он определяется как горизонтальный сдвиг низко- и высоковольтных подпороговых характеристик, деленный на разницу в стоковых напряжениях.  $DIBL$  тесно увязан с двумя пороговыми напряжениями, показанными на рис. 11. У идеального транзистора  $DIBL$  нулевой и пороговое напряжение не зависит от  $V_{DS}$ :  $V_T(\text{lin}) = V_T(\text{sat}) = V_T$ .

Как уже отмечалось, пороговое напряжение не является строго определенным понятием. Приблизительно можно считать его таким напряжением на затворе, при котором уже заметен стоковый ток, и есть разные подходы как определить это напряжение. Например, можно определить его таким образом, как показано на графике  $I_{DS} \propto V_{GS}$  (рис. 10). А можно поступить по-другому, как на рис. 11: выбрать небольшой стоковый ток, например,  $10^{-7} \text{ A}/\mu\text{m}$ , отмеченный на этом рисунке пунктирной линией, и определить  $V_T$  как напряжение на



затворе  $V_{GS}$ , при котором наблюдается этот ток. Когда приводится пороговое напряжение  $V_T$ , необходимо точно указывать, каким именно способом оно было определено.

Наконец, еще об использованных обозначениях. На рисунках 1 и 4 ток, входящий в сток транзистора  $n$ -MOSFET, обозначен как  $I_D$ . В идеальном случае тот же самый ток вытекает из канала проводимости:  $I_D = -I_S \equiv I_{DS}$ . На практике имеет место некоторый ток утечки (в подложку), так что  $I_D > I_S$ . В дальнейшем изложении ток утечки не учитывается и предполагается, что  $I_D = |I_S| \equiv I_{DS}$ .

### 1.7. Управление транзистором

Большинство типов транзисторов управляются путем контролирования высоты энергетического барьера в результате изменения напряжения на затворе. К ним относятся полевые транзисторы MOSFET, HEMT, JFET\*, а также биполярные транзисторы HBT\* и BJT\* [11, 20, 21]. Принципы управления транзисторами легко понять, обратившись к их зонной структуре.

Транзистор MOSFET является существенно двухмерной структурой (рис. 12).

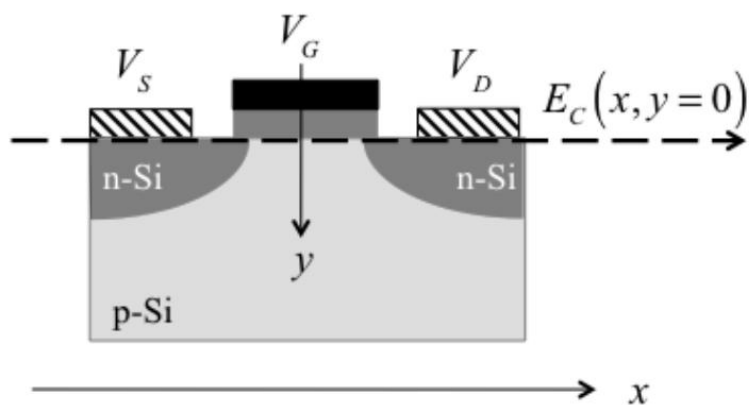


Рис. 12. Сечение  $n$ -MOSFET и система координат, используемая для графического представления энергии в зависимости от координат точки при движении от истока по каналу проводимости до стока. Ось  $z$  направлена за рисунок вдоль ширины канала проводимости  $W$ .

\* JFET / Junction Field-Effect Transistor – Полевой транзистор с управляющим  $p$ - $n$ -переходом,  
HBT / Heterojunction Bipolar Transistor – Биполярный гетеротранзистор,  
BJT / Bipolar Junction Transistor – Биполярный транзистор  $n$ - $p$ - $n$  или  $p$ - $n$ - $p$ .

Для полного понимания функционирования транзистора желательно, конечно, знать его многомерную зонную структуру, однако, бóльшая часть важных деталей работы MOSFET может быть продемонстрирована с использованием только одномерных зонных диаграмм вдоль оси  $x$ .

Контактные области истока и стока  $n$ -MOSFET  $n$ -допированы, а канал проводимости –  $p$ -типа. В однородно допированном массивном полупроводнике энергия зон не зависит от координаты с уровнем Ферми  $E_F$  вблизи дна зоны проводимости  $E_C$  для полупроводника  $n$ -типа и вблизи потолка валентной зоны  $E_V$  для полупроводника  $p$ -типа.

В верхней части рис. 13 показаны пространственно разделенные области истока и стока  $n$ -типа и канала проводимости  $p$ -типа. В разделенном состоянии в истоке и стоке  $E_F \approx E_C$ , а в канале проводимости  $E_F \approx E_V$ . Для получения одномерной зонной диаграммы транзистора в состоянии равновесия все три области нужно объединить. Примем во внимание, что электрохимический потенциал в состоянии равновесия везде должен быть одинаков ( $E_F = \text{Const}$ ).

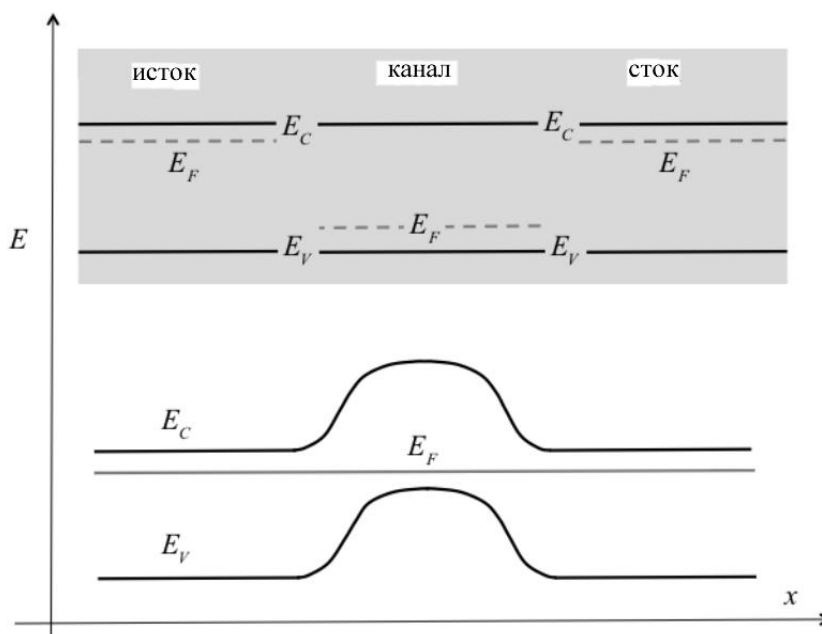


Рис. 13. Одномерные зонные диаграммы в состоянии равновесия. Наверху показаны уровни энергии разделенных истока и стока  $n$ -типа ( $E_C$  – энергия дна зоны проводимости) и канала проводимости  $p$ -типа ( $E_V$  – энергия потолка валентной зоны). Внизу показан одномерный ход энергии в состоянии равновесия когда все три области транзистора соединены и  $V_S = V_G = V_D = 0$ .

Для выравнивания электрохимического потенциала во всех трех областях зонные уровни энергии истока и стока должны понизиться или, что то же самое, уровни энергии валентной зоны канала проводимости должны повыситься. Выравнивание фермиевских уровней энергии происходит в результате того, что электроны перетекают с выше расположенных фермиевских уровней истока и стока на ниже расположенный фермиевский уровень канала, в результате чего зарядовое равновесие нарушается и возникает электростатический потенциал между истоком (стоком) и каналом. Первые приобретают положительный потенциал, так называемый потенциал поля  $p$ - $n$ -перехода, который понижает уровни энергии:

$$\begin{aligned} E_C(x) &= E_{C0} - q\psi(x), \\ E_V(x) &= E_{V0} - q\psi(x), \end{aligned} \quad (5)$$

где индекс «0» показывает величину соответствующей энергии в отсутствие электростатического потенциала  $\psi$ .

Поскольку транзистор находится в равновесии, ток по каналу проводимости не течет. Возник потенциальный барьер, разделяющий электроны в истоке от электронов в стоке. Этот энергетический барьер играет решающую роль в понимании и объяснении работы полевого транзистора. Теперь нужно понять каким образом изменяется зонная структура транзистора в результате подачи напряжения на затвор и на сток.

Рис. 14 показывает, что происходит с профилем зоны проводимости транзистора  $n$ -MOSFET при подаче положительного напряжения на затвор. Поскольку  $V_S = V_D = 0$ , фермиевские уровни энергии во всех трех областях практически одинаковы, транзистор находится в равновесии, ток по каналу проводимости не идет.

Терминал затвора отделен от канала проводимости в кремниевой подложке слоем диэлектрика  $\text{SiO}_2$ , однако, положительный потенциал, подаваемый на затвор, влияет на зонную структуру полупроводника. Он увеличивает электростатический потенциал в канале, что, согласно (5), понижает зону проводимости.

Важно, что подача напряжения на затвор не смещает фермиевский уровень канала проводимости. Подача положительного напряжения на затвор понижает уровень Ферми металлического терминала затвора, однако, он изолирован от подложки. Фермиевский уровень канала проводимости в подложке можно

изменить только лишь подавая напряжение на исток и/или сток, поскольку их фермиевские уровни связаны с уровнем Ферми канала проводимости.

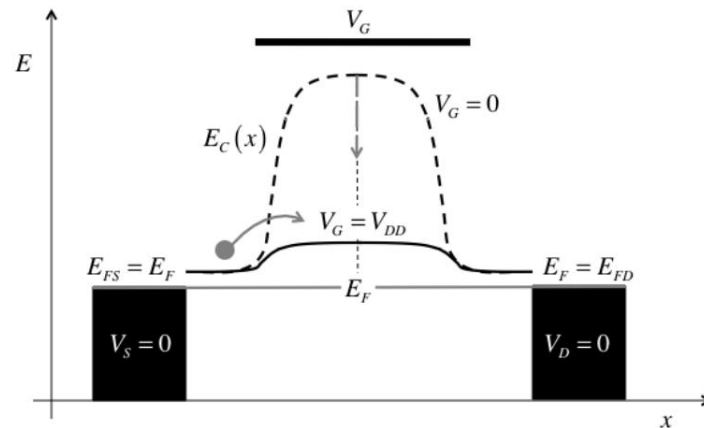


Рис. 14. Зонная структура  $n$ -MOSFET в равновесном состоянии при подаче на затвор близкого к нулю напряжения (пунктирная кривая) и значительного напряжения (сплошная кривая). Подача напряжения на затвор не нарушает равновесия, поскольку терминал затвора изолирован от канала проводимости.

Итак, подача потенциала на затвор просто лишь повышает или понижает потенциальный барьер между истоком и стоком. Транзистор остается в равновесии, ток в канале отсутствует. Тот факт, что равновесие в транзисторе не нарушается при подаче потенциала на затвор, как мы увидим позже, существенно упрощает анализ электростатики MOS.

Что происходит при подаче большого напряжения на сток показано на рис. 15.

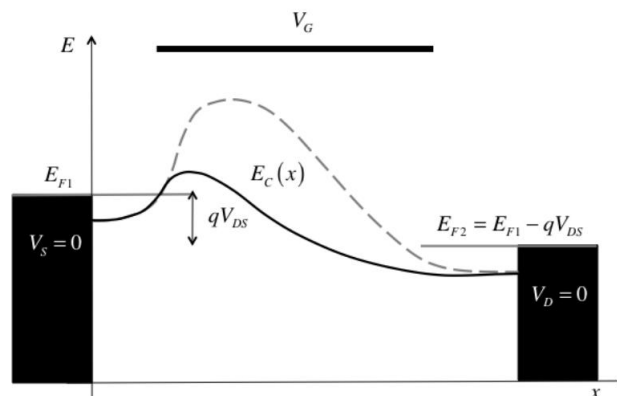


Рис. 15. Изменение зависимости  $E_C(x)$  вдоль канала проводимости  $n$ -MOSFET при подаче напряжения на сток и на затвор: пунктирная линия соответствует большому значению  $V_D$  и малому  $V_G$ , а сплошная линия – большим  $V_D$  и  $V_G$ .

Исток заземлен, так что его фермиевский уровень не изменился по сравнению с состоянием равновесия, а положительное напряжение, поданное на сток, понизило фермиевский уровень стока на величину  $qV_{DS}$  (рис. 15).

Рассмотрим сначала тот случай, когда на затвор подается небольшое напряжение, а на сток – большое. Этот случай показан пунктирной линией на рис. 15. В транзисторе хорошего качества высота энергетического барьера между истоком и каналом проводимости контролируется почти исключительно напряжением, приложенным к затвору. Если напряжение на затворе невелико, барьер достаточно высокий, и лишь считанные электроны обладают достаточной энергией преодолеть барьер и достичь стока. Транзистор находится в состоянии «OFF» с током  $I_{DS} \approx 0$ : имеет место незначительный ток утечки  $I_{OFF}$  (рис. 11).

Когда подается уже значительное напряжение на затвор одновременно с большим напряжением на стоке (сплошная линия на рис. 15), напряжение на затворе увеличивает электростатический потенциал в канале проводимости и понижает высоту барьера. Если барьер оказывается достаточно низким, то значительная часть электронов истока перепрыгивает барьер и движется к стоку. Транзистор находится в состоянии «ON» с максимальным током  $I_{ON}$  при  $V_{GS} = V_{DS} = V_{DD}$  (рис. 9).

Основной принцип работы большинства типов транзисторов – управление током путем изменения высоты барьера в результате подачи напряжения на затвор. Этим фактически исчерпывается физика состояний транзистора  $I_{OFF}$  и  $I_{ON}$  (рис. 9), однако, для понимания всей вольт-амперной характеристики транзистора нужно обратиться к диаграммам зонной структуры.

На рис. 16 показаны результаты моделирования изменения  $E_c(x)$  при росте напряжения на затворе в линейной области ВАХ. Обращает на себя внимание тот факт, что при высоких значениях напряжения на затворе величина  $E_c(x)$  изменяется практически линейно вдоль канала проводимости, что соответствует практически постоянному электрическому полю, как и ожидается для линейной области ВАХ, которая соответствует резистору, контролируемому напряжением на затворе (рис. 7).

Рис. 17 демонстрирует результаты моделирования функции  $E_c(x)$  для нанотранзисторов в области насыщения. По мере того, как растущее напряжение на затворе понижает барьер, электроны из истока преодолевают барьер и далее устремляются вниз по склону барьера к стоку.

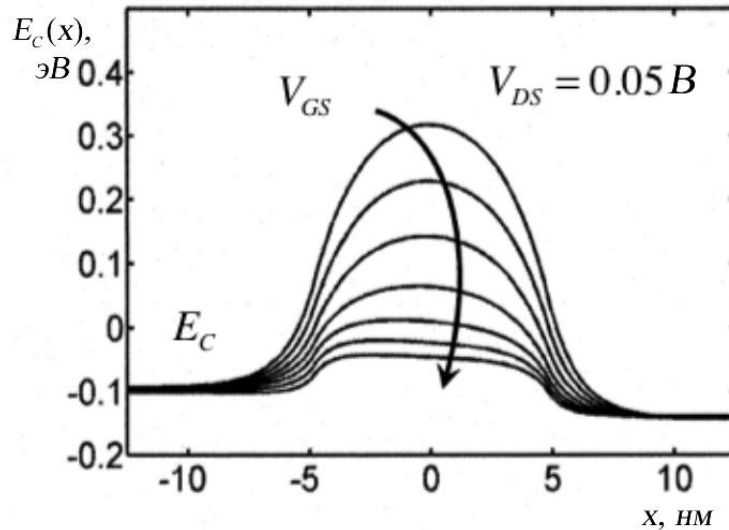


Рис. 16. Моделирование функции  $E_C(x)$  для нанотранзисторов. Приложено небольшое напряжение на стоке: транзистор работает в линейном режиме. Каждый из графиков соответствует определенному значению  $V_{GS}$ , величина которого увеличивается сверху вниз [22].

Рисунок 17 также показывает, почему ток стока насыщается с увеличением напряжения на стоке. Все дело в барьере между истоком и каналом, который и ограничивает ток к стоку.

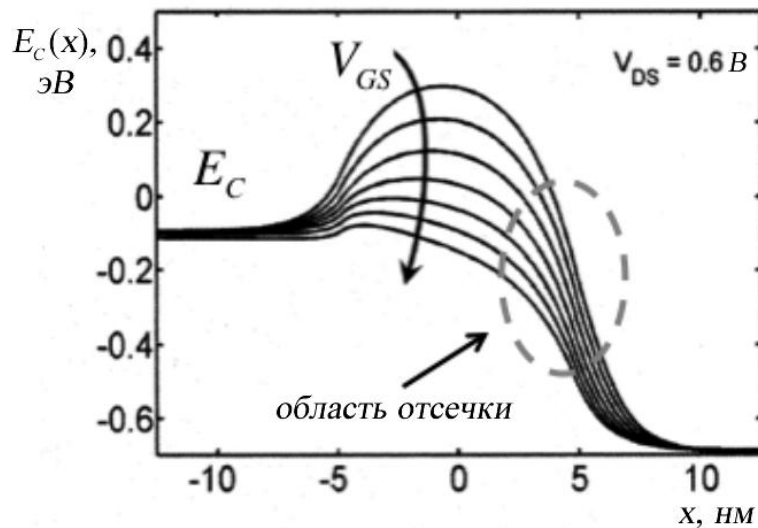


Рис. 17. Моделирование функции  $E_C(x)$  для нанотранзисторов. Приложено большое напряжение на стоке: транзистор работает в области насыщения. Каждый из графиков соответствует определенному значению  $V_{GS}$ , величина которого увеличивается сверху вниз [22]. Область отсечки обсудим позже.

Этот барьер хорошо виден на рис. 17 при самом большом из показанных значений  $V_{GS}$  (нижний график), а также на более подробных рисунках 18: на кривых (б) и (г) для  $V_{DS} = V_{GS} = 0.6B$ . Мы пользуемся классическим описанием транспорта электронов в нанотранзисторах с каналом проводимости длиной не менее 10 нм, что было убедительно продемонстрировано в [22 – 24].

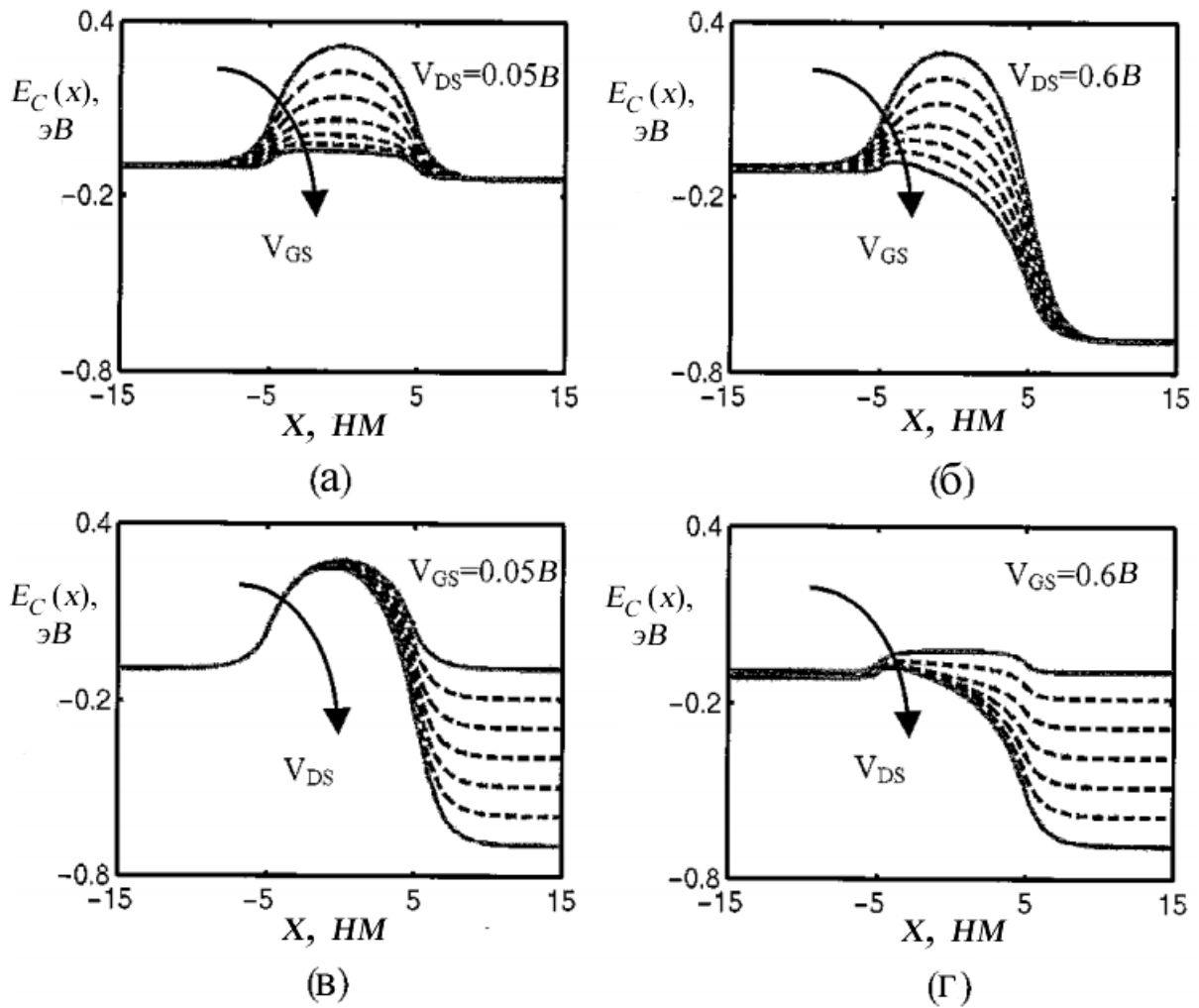


Рис. 18. Моделирование функции  $E_C(x)$  для нанотранзисторов:  
 (а)  $V_{DS} = 0.05B$ ,  $V_{GS} \propto 0.0 \Leftrightarrow 0.6B$  (рис. 16); (б)  $V_{DS} = 0.6B$ ,  $V_{GS} \propto 0.0 \Leftrightarrow 0.6B$  (рис. 17);  
 (в)  $V_{GS} = 0.05B$ ,  $V_{DS} \propto 0.0 \Leftrightarrow 0.6B$ ; (г)  $V_{GS} = 0.6B$ ,  $V_{DS} \propto 0.0 \Leftrightarrow 0.6B$  [22].

Увеличивая напряжение на затворе, барьер между истоком и каналом проводимости не исчезает (рис. 18г), что не позволяет расти току. Даже при весьма значительном напряжении на затворе этот небольшой барьер остается. Без этого барьера и управления им со стороны затвора не было бы и транзистора.

## 1.8. Вольт-амперная характеристика MOSFET с позиций модели термоэлектронной эмиссии

Описать вольт-амперную характеристику транзистора математически можно с привлечением простой модели термоэлектронной эмиссии. Рассмотрим ВАХ  $n$ -MOSFET, включенного по схеме с общим истоком (рис. 9). Суммарный ток стока есть разность двух токов – от истока к стоку (слева – направо) и от стока к истоку (справа – налево) с преодолением барьера в обоих направлениях:

$$I_{DS} = I_{LR} - I_{RL}. \quad (6)$$

Ток  $I_{LR}$  пропорционален вероятности преодоления барьера  $E_{SB}$  на пути электрона от истока к стоку:

$$I_{LR} \propto e^{-E_{SB}/kT}, \quad (7)$$

и аналогично для тока  $I_{RL}$ :

$$I_{RL} \propto e^{-E_{DB}/kT}, \quad (8)$$

где  $E_{DB}$  есть барьер на пути электрона от стока к истоку. Поскольку подача напряжения на сток понижает  $E_C$  на стоке, то  $E_{DB} > E_{SB}$ . При нулевом  $DIBL$

$$E_{DB} = E_{SB} + qV_{DS},$$

так что

$$I_{RL} / I_{LR} = e^{-qV_{DS}/kT}.$$

В итоге

$$I_{DS} = I_{LR} \left( 1 - e^{-qV_{DS}/kT} \right). \quad (9)$$

На вершине барьера есть два потока электронов – один поток движется направо, а другой – налево, с одной и той же кинетической энергией, стало быть с одинаковыми скоростями  $v_T$ . Ток есть произведение заряда на скорость. В случае MOSFET электроны движутся в 2D-канале проводимости, так что нужно учитывать поверхностный заряд ( $K/cm^2$ ). Ток

$$I_{LR} = W \underline{Q^+}_{x=0} v_T,$$

где  $\underline{Q^+}_{x=0}$  есть поверхностный заряд на вершине барьера, obligatory для электронам, движущимся с положительными скоростями. Аналогично

$$I_{RL} = W \underline{Q^-}_{x=0} v_T,$$



так что суммарный заряд от обоих потоков

$$\underline{Q}|_{x=0} = \frac{I_{LR} + I_{RL}}{Wv_T} = \frac{I_{LR}}{Wv_T} \left(1 + e^{-qV_{DS}/kT}\right). \quad (10)$$

Отсюда находим ток  $I_{LR}$ , подставляем его в (9) и окончательно получаем ВАХ  $I_{DS} \propto V_{DS}$  для баллистического MOSFET:

$$I_{DS} = W \left| \underline{Q} \right|_{x=0} |v_T \frac{(1 - e^{-qV_{DS}/kT})}{(1 + e^{-qV_{DS}/kT})}. \quad (11)$$

Позже мы выведем уравнение (11) более строго, обсудим его ограничения и определим скорость  $v_T$ . Вместе с тем, поведение баллистической характеристики вполне объяснимо с позиций простой модели термоэлектронной эмиссии для транзистора с контролируемым барьером.

Проанализируем уравнение (11) в условиях низких и высоких значений напряжения на стоке  $V_{DS}$ . Для низких значений  $V_{DS}$  разложение в ряд Тейлора экспонент дает

$$I_{DS} = W \left| \underline{Q} \right|_{x=0} |v_T \frac{V_{DS}}{2kT/q} = G_{CH} V_{DS} = \frac{V_{DS}}{R_{CH}}, \quad (12)$$

где  $G_{CH}$  и  $R_{CH}$  есть проводимость и сопротивление канала транзистора. Уравнение (12) есть баллистический аналог линейной области на рис. 9.

Рассмотрим уравнение (11) в случае высоких значений  $V_{DS}$ , в области насыщения (рис. 9). В этом случае  $I_{RL} \ll I_{LR}$  и ток стока насыщается при  $I_{DS} = I_{LR}$ . В этом пределе  $V_{DS} \gg kT/q$  и (11) упрощается до

$$I_{DS} = W \left| \underline{Q} \right|_{x=0} |v_T, \quad (13)$$

т. е. при высоких значениях  $V_{DS}$  ток не зависит от напряжения на стоке. Однако, позже мы увидим, что *DIBL* вызывает рост заряда  $\underline{Q}|_{x=0}$  с увеличением напряжения на стоке, так что ток не полностью насыщается при высоких значениях  $V_{DS}$ .

Теперь обратимся к передаточным характеристикам на рис. 11. Они представляют собой зависимости  $I_{DS} \propto V_{GS}$  при фиксированных значениях  $V_{DS}$ . Пусть мы фиксировали напряжение на стоке высоким, так что ток дается

выражением (13). Зададимся вопросом: как изменяется заряд  $Q|_{x=0}$  с напряжением на стоке?

Для высоких значений  $V_{DS}$  ток  $I_{RL} = 0$ , так что из (10) следует, что

$$|Q|_{x=0}| = \frac{I_{LR}}{Wv_T}. \quad (14)$$

Ток  $I_{LR}$  вызывается термоэлектронной эмиссией истока в направлении канального барьера. Приложение напряжения на затворе понижает этот барьер:

$$I_{LR} \propto e^{-E_{SB}/kT} = e^{-(E_{SB}^0 - qV_{GS}/m)/kT}, \quad (15)$$

где  $E_{SB}^0$  есть высота барьера при входе на него со стороны истока при нулевом потенциале на затворе ( $V_{GS} = 0$ ), а  $1/m$  есть доля потенциала на затворе, пробивающая диэлектрическую прокладку затвора. Из (15) и (14) следует, что

$$Q(V_{GS}) = Q|_{V_{GS}=0} e^{qV_{GS}/mkT}. \quad (16)$$

Теперь из (16) и (13) следует, что ток растет экспоненциально с увеличением напряжения на затворе:

$$I_{DS} = W |Q|_{V_{GS}=0}| v_T e^{qV_{GS}/mkT}. \quad (17)$$

Не сложно показать, что для увеличения тока в 10 раз (декада) напряжение на затворе нужно увеличить на  $2.3mkT/q \geq 0.060V$  (при комнатной температуре). Этот допороговый разброс  $SS \geq 60 мВ/декада$  характерен для термоэлектронной эмиссии через барьер.

Итак, согласно (17), ток стока не зависит от напряжения на стоке, однако, в реальных устройствах наблюдается небольшой рост тока стока с увеличением  $V_{DS}$ , поскольку напряжение на стоке как бы помогает затвору уменьшить разницу между истоком и канальным барьером (рис. 19).

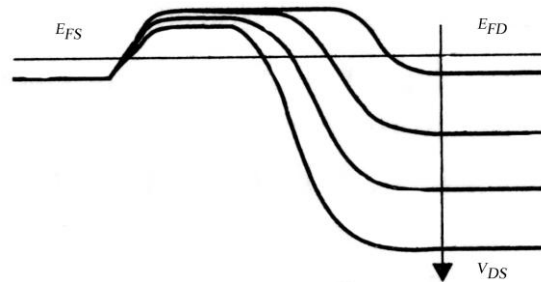


Рис. 19. Схематическая иллюстрация эффекта *DIBL*.

Для объяснения эффекта *DIBL* потребуется учет *2D*-электростатики, чем мы займемся позже.

Уравнение (17) показывает экспоненциальный рост  $I_{DS} \propto V_{GS}$ , наблюдаемый на рис. 11 ниже порога  $V_T$ , однако, выше порога ток не растет экспоненциально, а увеличивается приблизительно линейно с ростом напряжения на затворе. Позже, при обсуждении электростатики MOS будет показано, что заряд (16) выше  $V_T$  растет линейно с  $V_{GS}$  согласно уравнению

$$Q(V_{GS}, V_{DS}) = -C_{ox}(V_{GS} - V_T), \quad V_T = V_{T0} - \delta V_{DS}, \quad C_{ox} = \kappa_{ox} \epsilon_0 / t_{ox}, \quad (18)$$

где  $C_{ox}$  – поверхностная емкость диэлектрической прокладки, обратно пропорциональная её толщине  $t_{ox}$ .

Итак, выходная характеристика баллистического транзистора MOSFET находит в целом достаточно простое объяснение с позиций простой модели термоэлектронной эмиссии с учетом контролируемого затвором барьера.

Вся физика транзистора сводится фактически к электростатике и транспорту электронов. Что же касается инженерного дизайна транзисторов, то его роль критически возрастает по мере их миниатюризации.

Рисунок 20 может служить иллюстрацией ключевых принципов хорошего дизайна нанотранзисторов.

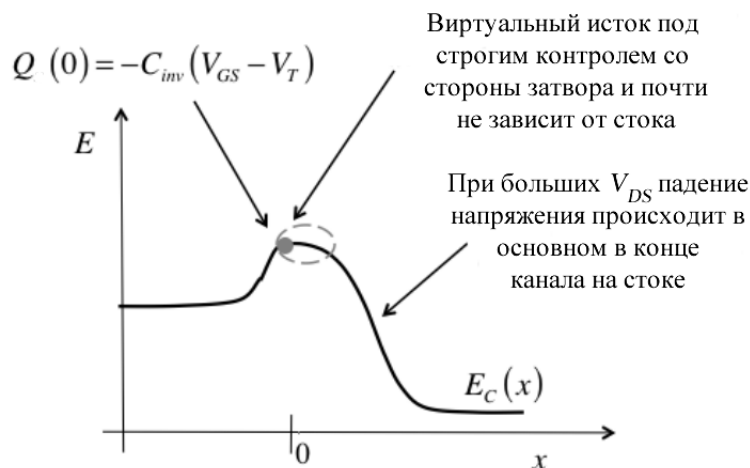


Рис. 20. Ход барьера качественно собранного нанотранзистора при больших напряжениях на затворе и на стоке. В таком нанотранзисторе заряд на вершине барьера весьма близок к значению, которое имеет место в транзисторе с длинным каналом проводимости, когда внешним электрическим полем можно пренебречь. В качественном нанотранзисторе внешнее поле низкое в начале канала проводимости и при высоких  $V_{DS}$  напряжение на стоке лишь незначительно влияет на область в районе вершины барьера.

Вершина барьера является критической точкой, которой соответствует начало канала проводимости и называемой далее виртуальным истоком. У качественного транзистора высота барьера строго контролируется напряжением на затворе и лишь исключительно слабо зависит от напряжения на стоке.

При низких  $V_{DS}$  и высоких  $V_{GS}$  потенциал в канале падает приблизительно линейно, так что электрическое поле приблизительно постоянное. При высоких напряжениях и на стоке и на затворе электрическое поле сильное и изменяется вдоль канала нелинейно. В самом начале канала (у вершины барьера) электрическое поле слабое, однако, вблизи стока поле очень сильное. В области насыщения увеличение напряжения на стоке усиливает падение потенциала на участке канала с сильным полем, однако, оставляет область вблизи вершины барьера почти неизменной (при небольшом значении  $DIBL$ ). Поскольку именно область вблизи вершины барьера контролирует ток  $I_{DS}$ , то он в области насыщения менее чувствителен к напряжению на стоке.

Электроны, преодолевшие барьер (рис. 20), скатываясь вниз к стоку приобретают значительную кинетическую энергию. Часть этой энергии тратится на электрон-фононное рассеяние, однако, в случае нанотранзисторов у электронов недостаточно времени рассеять их энергию в ходе скатывания к стоку. В результате скорость электронов в нанотранзисторах очень велика в концевой части канала. Поскольку ток есть произведение заряда на скорость, то в области насыщения плотность электронов будет очень низкой при их высокой скорости. В классической теории MOS та часть канала, в которой падение потенциала велико, а плотность электронов мала, известна как область отсечки канала. В транзисторах с коротким каналом область отсечки может занимать значительную часть канала, однако, в хорошо электростатически сконструированных наноразмерных MOSFET всегда должна быть небольшая область вблизи истока, находящаяся практически под контролем затвора и характеризующаяся незначительным падением потенциала.

Транзистор с протяженным каналом проводимости при высоких значениях напряжения как на затворе, так и на стоке показан на рис. 21.

Сравнивая его с короткоканальным транзистором на рис. 20, видим, что под контролем затвора находится преобладающая часть канала, но все же есть пусть небольшая область отсечки вблизи стока. Именно эта область отсечки при высоком напряжении на стоке вызывает насыщение тока. В области насыщения (за областью отсечки) ток в основном обусловлен транспортом электронов по низкополевой части канала вблизи истока, а наибольшее падение

потенциала вдоль канала приходится на область с высоким электрическим полем вблизи стока. Как только электроны попали в область отсечки, они очень быстро уносятся в сток.

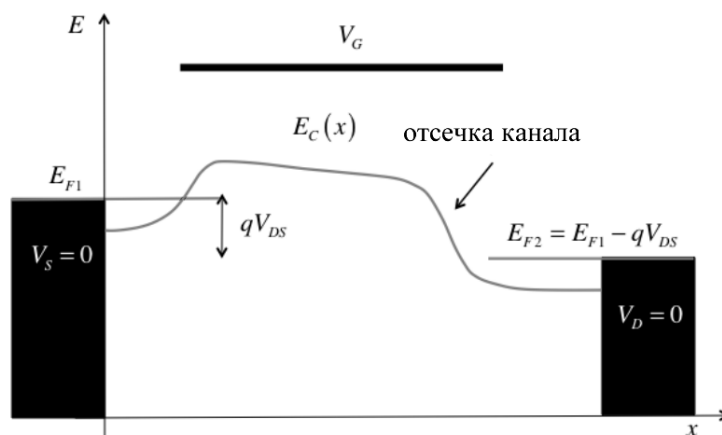


Рис. 21. Схематический ход  $E_c(x)$  для транзистора с протяженным каналом проводимости при высоких значениях напряжения как на затворе, так и на стоке. В этом случае низкое внешнее электрическое поле захватывает значительную часть канала проводимости и область отсечки канала незначительная. Дополнительное увеличение напряжения на стоке несколько увеличит область отсечки, протяженность которой охватывает лишь небольшую часть длинного канала проводимости.

Итак, у качественно собранного нанотранзистора область вблизи вершины барьера находится под строгим контролем затвора и напряжение на затворе практически не должно влиять на эту область. Эта цель является глобальной в проблеме дизайна нанотранзисторов и она становится все более актуальной по мере сближения стока и истока. Как только электроны перепрыгнули барьер (виртуального) истока, они направляются к стоку. Дизайн электростатики MOSFET бросает все более серьезный вызов по мере уменьшения канала проводимости, однако, физические принципы дизайна остаются прежними. Вместе с тем, физическая природа транспорта электронов существенно меняется по мере того как транзисторы становятся все меньше и меньше. Физически корректное описание электронного транспорта в нанотранзисторах становится существенным моментом в инженерном дизайне MOSFET.

Обсуждая ход  $E_c(x)|_{z=0}$ , имелась в виду одномерная зонная структура MOSFET вдоль поверхности кремниевой подложки. Рисунок 22 может служить иллюстрацией 2D зонной поверхности: а) сам  $n$ -MOSFET; б) это равновесное состояние с  $V_s = V_d = 0$  и так подобранным напряжением на затворе, чтобы зоны были плоскими в направлении, перпендикулярном к траектории движения

электронов от истока к стоку; в) в этом случае приложено большое напряжение на затворе при  $V_s = V_d = 0$ , а  $E_c(x, y)$  вдоль поверхности транзистора проецируется в прямую линию на рис. 14; г) теперь большое напряжение приложено и к затвору и к стоку, а  $E_c(x, y)$  вдоль поверхности транзистора проецируется в сплошную линию на рис. 15.

Обратим также внимание на то, что зонные диаграммы MOSFET подобны таковым для биполярных транзисторов [11, 20]. В обоих случаях ток контролируется высотой барьера [26]. Исток MOSFET аналогичен эмиттеру ВТТ, канал MOSFET – базе, а сток – коллектору ВТТ. Эта тесная аналогия окажется полезной для понимания физических принципов функционирования наноразмерных MOSFET.

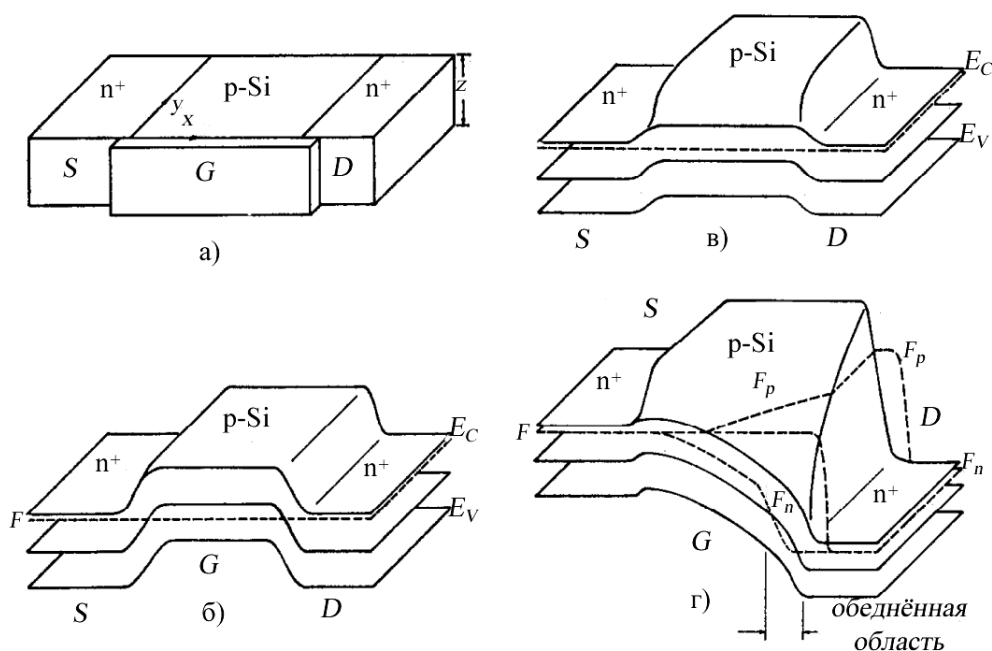


Рис. 22. 2D-зонная структура  $n$ -MOSFET: а)  $n$ -MOSFET; б) равновесная зонная диаграмма; в) зонная диаграмма с большим приложенным напряжением на затворе; г) зонная диаграмма с большим приложенным напряжением как на затворе, так на стоке [25].

Подведем итоги. Основное внимание в первой главе уделено полевым транзисторам с изолированным затвором типа MOSFET. Рассмотрены физическая структура MOSFET, вольт-амперные характеристики обоих типов – выходные и передаточные, подробно описаны 10 метрических показателей MOSFET, достаточных для анализа качества нанотранзисторов, обсуждены принципы управления полевыми транзисторами путем подачи потенциала на затвор, обращено внимание на существование не исчезающего барьера на границе между истоком и каналом проводимости, получившего название

виртуального истока, факт существования которого в конечном итоге привел к разработке наиболее адекватного метода анализа работы нанотранзистров, известного как метод виртуального истока/прохождения, берущего свое начало еще в пионерских работах Р. Ландауэра [14], а позже развитого С. Датта и М. Лундстромом [15 – 18]. Эта глава заканчивается изложением математической модели вольт-амперной характеристики MOSFET с привлечением простой модели термоэлектронной эмиссии, как основы для анализа и построения более совершенных моделей.

Далее будут рассмотрены модель MOSFET в традиционном изложении и начала модели виртуального истока, уравнение Пуассона – Больцмана и приближение истощения, критерии для слабой, умеренной и сильной инверсии, будет показано как связаны между собой поверхностный потенциал и напряжение на затворе, будут рассмотрены подвижный электронный заряд ниже и выше порогового напряжения, его роль в структурах с чрезвычайно тонкой КНИ-подложкой, эффекты учета  $2D$  электростатики и, наконец, сформулируем модель виртуального истока с позиций «сверху – вниз». В заключение будет построена адекватная физическая картина транспортных явлений в нанотранзисторах, исходя из концепции «снизу – вверх» Ландауэра – Датта – Лундстрома, и будет показана связь ее с привычным традиционным подходом «сверху – вниз».

## Литература

1. В. Г. Абліцов, *Галактика «Україна». Українська діаспора: видатні постаті* (Київ: КИТ: 2007).
2. M. Riordan, L. Hoddeson, *Crystal Fire: The Birth of the Information Age* (New York: W.W. Norton & Co.: 1997).
3. Bo Lojek, *History of Semiconductor Engineering* (New York: Springer: 2007).
4. G. E. Moore, *Electronics Magazine*, 4 – 7 (1965).
5. M. Lundstrom, *Science*, **299**: 210 – 211 (2003).
6. S. R. Hofstein, F. P. Heiman, *Proc. IEEE*, 1190 – 1202 (1963).
7. C. T. Sah, *IEEE Trans. Electron Dev.*, **11**: 324 – 345 (1964).
8. H. Shichman, D. A. Hodges, *IEEE J. Solid State Circuits*, **SC-3**: (1968).
9. B. J. Sheu, D. L. Scharfetter, P.-K. Ko, M.-C. Jeng, *IEEE J. Solid State Circuits*, **SC-22**: 558 – 566 (1987).
10. Y. Tsividis, C. McAndrew, *Operation and Modeling of the MOS Transistor* (New York: Oxford Univ. Press: 2011).
11. Y. Taur, T. Ning, *Fundamentals of Modern VLSI Devices* (New York: Oxford Univ. Press: 2013).
12. D. Frank, S. Laux, M. Fischetti, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 553 – 556 (1992).
13. Z. Ren., R. Venugopal, S. Goasguen, S. Datta, M. S. Lundstrom, *IEEE Trans. Electron Dev.*, **50**: 1914 – 1925 (2003).
14. R. Landauer, *IBM J. Res. Dev.*, **1**, № 3: 223 – 231 (1957).
15. S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport* (Singapore: World Scientific: 2012).
16. Ю. А. Кругляк, *Нанoeлектроника «снизу – вверх»* (Одесса: ТЭС: 2015).
17. S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport – Part B: Quantum Transport* (Singapore: World Scientific: 2018).
18. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
19. Kwork K. Ng, *Complete Guide to Semiconductor Devices* (New York: Wiley Interscience: 2002).
20. R. F. Pierret, *Semiconductor Device Fundamentals* (New York: Addison-Wesley: 1996).
21. И. М. Викулин, В. И. Стафеев, *Физика полупроводниковых приборов* (Москва: Радио и связь: 1990).
22. M. Lundstrom, Z. Ren, *IEEE Trans. Electron Dev.*, **49**: 133 – 141 (2002).
23. Y. Naveh, K. K. Likharev, *IEEE Electron Device Lett.*, **21**, 242 – 244 (2000).
24. Z. Ren, R. Venugopal, S. Datta, M. Lundstrom, D. Jovanovic, J. Fossum, *IEDM, Technical Digest*, 715 – 718 (2000).
25. H. C. Pao, C. T. Sah, *Solid-State Electron.*, **9**: 927 – 937 (1966).
26. E. O. Johnson, *RCA Rev.*, **34**: 80 – 94 (1973).





## Глава 2. Теория MOSFET в традиционном изложении, начала модели виртуального истока и приближение истощения

### 2.1. Введение

Теория MOSFET была разработана еще в 60-х годах прошлого века [1 – 5] и хотя она получила существенное развитие, но и ныне основные черты моделирования транзисторов остались такими же, как и 50 лет назад. Эта глава посвящена традиционному изложению теории MOSFET [6, 7]. Мы обсудим лишь наиболее существенные идеи привычного подхода, получившего также название «сверху – вниз». В частности, мы ограничимся моделированием линейной области и области насыщения вольт-амперных характеристик. Более детальное и полное представление о классической теории MOSFET можно найти в учебниках [8, 9]. Позже мы рассмотрим существенно другой подход к теории MOSFET, более адекватный физике нанотранзисторов, и вместе с тем покажем его связь с традиционным подходом.

### 2.2. Ток, заряд и скорость

Ток в  $n$ -MOSFET (рис. 1) дается общим выражением

$$I_{DS} = W |Q(x)| \langle v \rangle, \quad (1)$$

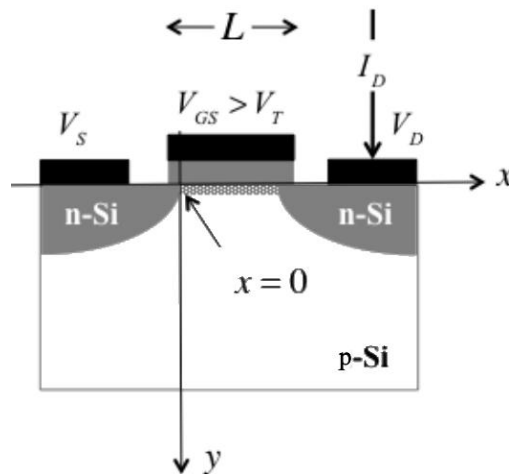


Рис. 1. Упрощенное изображение  $n$ -Si MOSFET в режиме обогащения.

Ось  $y$  направлена вглубь подложки, а ось  $z$  — вдоль ширины канала проводимости с инверсионным зарядом, поскольку  $V_{GS} > V_T$ . Показано равномерное ( $V_S = V_D = 0$ ) распределение электронного заряда в пределах канала проводимости, соединяющего  $n$ -допированные терминалы  $S$  и  $D$ , от начала канала в точке  $x=0$  до  $x=L$ .

в котором  $Q$  есть подвижный поверхностный электронный заряд ( $K/m^2$ ) инверсионного слоя в плоскости  $x-y$  вглубь подложки  $p$ -Si,  $\langle v \rangle$  – средняя скорость электронов в инверсионном слое,  $W$  – ширина слоя. Предполагается, что транзистор однороден вдоль оси  $z$ , направленной за рисунок, а электроны движутся вдоль оси  $x$  от истока к стоку (направление тока  $I_{DS}$  противоположное).

Заряд  $Q$  называют зарядом инверсионного слоя, поскольку речь идет об электронном заряде в полупроводнике  $p$ -типа. Заряд и скорость изменяются при движении вдоль канала, а ток постоянен, если только предполагать, что процессы рекомбинации-генерации электронов в канале не происходят. Поэтому величину тока можно вычислять в любой точке канала, где это удобно.

Пусть  $V_S = V_D = 0$ , а  $V_{GS} > 0$ . Транзистор находится в равновесном состоянии, тока нет. В этом случае заряд  $Q$  распределен равномерно вдоль канала и он небольшой, если  $V_{GS} < V_T$ . При  $V_{GS} > V_T$  отрицательный заряд пропорционален их разности, а именно:

$$Q(V_{GS}) = -C_{ox}(V_{GS} - V_T), \quad (2)$$

$$C_{ox} = \kappa_{ox} \varepsilon_0 / t_{ox}, \quad (3)$$

где  $C_{ox}$  есть поверхностная емкость диэлектрической прокладки ( $\Phi/m^2$ ), обратно пропорциональная её толщине  $t_{ox}$ . При  $V_{GS} \leq V_T$  заряд предполагается пренебрежимо малым.

Когда  $V_D > V_S$  зарядовая плотность в инверсионном слое изменяется по ходу вдоль канала проводимости, как и скорость электронов. Как мы увидим позже при более детальном обсуждении электростатики MOS, в «хорошем» транзисторе зарядовая плотность  $Q$  в самом начале канала ( $x = 0$ ) дается именно выражением (2). С целью вычислить  $I_{DS}|_{x=0}$ , где мы уже знаем заряд  $Q$  согласно (2), потребуется вычислить среднюю скорость  $\langle v \rangle|_{x=0}$  в начале канала.

### *Линейная область*

При небольших напряжениях на стоке  $V_{DS}$  наблюдается линейный участок на выходной характеристике. Выше порога  $V_T$  электрическое поле  $E$  в канале постоянное, и для средней скорости электронов можно написать

$$\langle v \rangle = -\mu E = -\mu V_{DS} / L, \quad (4)$$

где  $\mu$  – подвижность электронов,  $L$  – длина канала проводимости, что с учетом (1) и (2) дает классическое выражение для тока стока MOSFET при небольших напряжениях на стоке, а именно:

$$I_{DS} = \frac{W}{L} \mu C_{ox} (V_{GS} - V_T) V_{DS}. \quad (5)$$

В традиционной теории MOSFET под подвижностью подразумевается некая эффективная подвижность  $\mu_{eff}$ , трактуемая как подвижность, усредненная по глубине инверсионного слоя. Она меньше чем подвижность электронов в массивном образце, что связывают с рассеянием электронов на шероховатостях граничной поверхности раздела SiO<sub>2</sub>/Si.

Уравнение (5) может быть непосредственно использовано для оценки порогового напряжения  $V_T$ . Для напряжений на затворе  $V_{GS}$ , достаточно превышающих значение  $V_T$ , ток стока  $I_{DS}$  растет линейно с  $V_{GS}$ . Экстраполяция этой линейной зависимости на  $I_{DS} = 0$  позволяет определить так называемое экстраполяционное значение порогового напряжения  $V_T$ . Различные методы определения  $V_T$  можно найти в [8].

*Область насыщения: скорость насыщения*

При больших  $V_{DS}$ , в области насыщения выходных характеристик MOSFET ведет себя как источник тока, контролируемый напряжением (рис. 7/гл. 1). Для относительно небольшого напряжения  $V_{DS} \approx 1B$  электрическое поле в канале проводимости нанотранзистора ( $L \approx 20nm$ ) велико – значительно выше  $\approx 10 kB/cm$ , необходимого для насыщения скорости в массивном Si. Если электрическое поле велико вдоль всего канала при  $V_{DS} > V_{DSAT}$ , тогда скорость электронов постоянна, равна  $v_{sat}$ , и среднюю скорость можно записать в виде

$$\langle v(x) \rangle = v_{sat} \approx 10^7 cm/c. \quad (6)$$

Подставляя (6) и (2) в (1), получаем классическое выражение для тока стока MOSFET насыщенного по скорости, а именно:

$$I_{DS} = W C_{ox} v_{sat} (V_{GS} - V_T). \quad (7)$$

В реальных транзисторах ток не насыщается полностью, а медленно увеличивается с ростом напряжения на стоке.

Наконец, нужно обратить внимание на то, что в нанотранзисторах максимальная скорость в канале проводимости не насыщается даже при очень высоком электрическом поле. В учебниках, излагающих традиционный подход к теории MOSFET, тем не менее предполагается, что скорость электронов насыщается при высоком электрическом поле в канале проводимости.

*Область насыщения: классическая область отсечки*

Рассмотрим MOSFET с достаточно протяженным каналом при высоком напряжении на стоке. В этом случае электрическое поле умеренное и насыщение по скорости не ожидается. Тем не менее, ток стока насыщается, и должна быть какая-то другая причина этого. Явление это наблюдалось еще на заре науки о полевых транзисторах, когда длина канала составляла около  $10 \mu\text{м}$  ( $=10000 \text{ нм}$ ) и объяснением насыщения тока стока считалась отсечка вблизи стока.

При высоком напряжении на стоке потенциал в канале существенно меняется от значения  $V_s$  на истоке до значения  $V_D$  на стоке, к более подробному рассмотрению хода которого мы вскоре вернемся. Перепишем (2) в виде

$$Q(V_{GS}, x) = -C_{ox}[V_{GS} - V_T - V(x)], \quad (8)$$

где  $V(x)$  есть потенциал, изменяющийся вдоль канала проводимости длины  $L$ . Согласно (8), если  $V_D = V_{GS} - V_T$  на стоке ( $x = L$ ), заряд на стоке  $Q(V_{GS}, L) = 0$ : говорят об отсечке канала на стоке. Согласно (1), если заряд  $Q = 0$ , то ток стока должен зануляться, на самом же деле наблюдается значительный ток стока. Это связано с тем, что в области отсечки канала электроны в сильном электрическом поле движутся очень быстро, так что плотность заряда конечна, хотя и очень небольшая. Ток насыщается при напряжении на стоке выше  $V_{GS} - V_T$ , поскольку в этой области добавляется еще потенциал отсечки. Рассмотрим ситуацию более подробно.

Итак, мы имеем достаточно протяженный канал MOSFET (рис. 2) при высоком напряжении на затворе и при напряжении на стоке, большем  $V_{GS} - V_T$ . Практически вдоль всего канала имеется инверсионный слой, и  $v(x) = -\mu E(x)$ . Когда электроны заходят в область отсечки, сильное электрическое поле быстро переносит их к стоку (рис. 21/гл. 1).

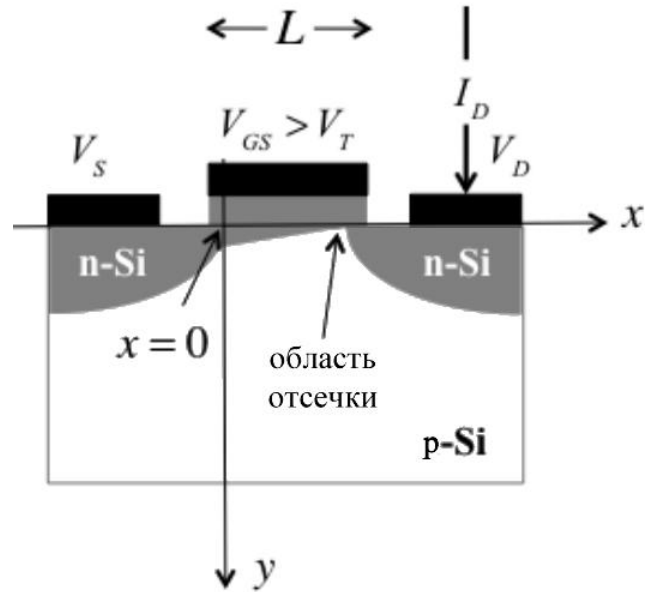


Рис. 2. Длинноканальный MOSFET с указанием на область отсечки. Переменная глубина канала на рисунке иллюстрирует изменение зарядовой плотности вдоль канала – плотность заряда больше в начале истока и уменьшается по мере приближения к стоку. Физическая глубина канала ведет себя противоположным образом: глубина канала растет от истока, где потенциал поля затвора наибольший, к стоку, где потенциал поля затвора наименьший. Длина канала, на протяжении которой плотность заряда существенна, есть  $L' < L$ .

Для той части инверсионного канала, для которой плотность заряда велика, среднюю скорость можно записать в виде

$$\langle v(x) \rangle = -\mu E(x). \quad (9)$$

Напряжение в начале канала  $V(0) = V_S = 0$ , а напряжение в конце канала в области отсечки есть  $V_{GS} - V_T$ . Электрическое поле в начале канала, как будет показано ниже,

$$E(0) = \frac{V_{GS} - V_T}{2L'}, \quad (10)$$

где множитель двойка требуется для корректного учета нелинейности электрического поля в канале, в чем мы убедимся позже, а  $L'$  есть длина канала проводимости, не затронутого отсечкой. Подстановка (10) в (9) для начала канала дает

$$\langle v(0) \rangle = -\mu E(0) = -\mu \frac{V_{GS} - V_T}{2L'}. \quad (11)$$

Подставляя, наконец, (2) и (11) в (1), получаем так называемый квадратичный закон для выходной характеристики достаточно длинного канала MOSFET:

$$I_{DS} = \frac{W}{2L'} \mu C_{ox} (V_{GS} - V_T)^2. \quad (12)$$

В реальных условиях ток не насыщается полностью, а медленно растет с увеличением напряжения на стоке, в то время как область отсечки медленно сдвигается по направлению к истоку, что эффективно уменьшает длину  $L'$  проводящей части канала проводимости.

*Вычисление цельной ВАХ: от линейной части до квадратичной области насыщения*

Уравнения (5) и (12) описывают линейную часть и область насыщения в рамках традиционной теории MOSFET. Покажем как написать уравнение цельной выходной характеристики с позиций этого же подхода. Начнем с общего уравнения (1) с подстановкой в него уравнения (4) через потенциал электрического поля:

$$I_{DS} = W |Q(x)| \langle v(x) \rangle = W |Q(x)| \mu \frac{dV}{dx}. \quad (13)$$

После подстановки выражения (8) для заряда получаем

$$I_{DS} = W \mu C_{ox} [V_{GS} - V_T - V(x)] \frac{dV}{dx}, \quad (14)$$

разделяем переменные и интегрируем вдоль канала проводимости

$$I_{DS} \int_0^{L'} dx = W \mu C_{ox} \int_{V_S}^{V_D} (V_{GS} - V_T - V) dV \quad (15)$$

в предположении постоянства вдоль канала как подвижности, так и тока стока. После интегрирования окончательно получаем искомое уравнение:

$$I_{DS} = \frac{W}{L'} \mu C_{ox} [(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}]. \quad (16)$$

Это уравнение дает ток стока для  $V_{GS} > V_T$  и  $V_{DS} \leq (V_{GS} - V_T)$ . Заряд в (8) зануляется при  $V_{DS} = V_{GS} - V_T$ , обозначая начало области отсечки. Ток за отсечкой находим из

уравнения (16) при  $V_{DS} = V_{GS} - V_T$ , а именно:

$$I_{DS} = \frac{W}{2L'} \mu C_{ox} (V_{GS} - V_T)^2, \quad (17)$$

и изменяется этот ток только с ростом  $V_{DS}$ , поскольку эффективная длина канала уменьшается в результате отсечки ( $L' < L$ ).

Уравнения (16) и (17) исчерпывают суть квадратичного закона выходных характеристик MOSFET, не только в линейной области и в области насыщения, а при любых напряжениях.

### 2.3. Изменение электрического поля вдоль канала

Излагая традиционную модель MOSFET, обычно предполагают, что электрическое поле в достаточно длинном канале при низком напряжении на стоке есть  $V_{DS}/L$ , а при высоком напряжении есть  $(V_{GS} - V_T)/2L'$ . Вычислим ход электрического поля в канале и покажем, что эти предположения действительно справедливы.

Из (14) и (16) имеем

$$\frac{1}{L'} [(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}] = [V_{GS} - V_T - V(x)] \frac{dV}{dx}, \quad (18)$$

разделяем переменные и после интегрирования от истока при  $x=0, V_s=0$  до произвольной точки в канале  $x$  с  $V=V(x)$  получаем квадратное уравнение для потенциала  $V(x)$

$$[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}] \frac{x}{L'} = (V_{GS} - V_T)V(x) - \frac{V^2(x)}{2}, \quad (19)$$

решая которое, получаем

$$V(x) = (V_{GS} - V_T) \left[ 1 - \sqrt{1 - \frac{2(V_{GS} - V_T)V_{DS} - V_{DS}^2/2}{(V_{GS} - V_T)^2} \left( \frac{x}{L'} \right)} \right]. \quad (20)$$

Дифференцируя это уравнение, находим электрическое поле. Рассмотрим два случая. Для небольших значений напряжения на стоке  $V_{DS}$ , на линейном участке входной характеристики имеем

$$V(x) = (V_{GS} - V_T) \left[ 1 - \sqrt{1 - \frac{2V_{DS}}{(V_{GS} - V_T)} \left( \frac{x}{L'} \right)} \right]. \quad (21)$$



Поскольку для малых аргументов  $\sqrt{1-x} \approx 1-x/2$ , уравнение (21) упрощается до

$$V(x) = V_{DS} \frac{x}{L}, \quad (22)$$

имея в виду, что при небольших значениях напряжения на стоке  $L' = L$ . Дифференцируя (22), получаем ожидаемый результат при малых напряжениях на стоке  $V_{DS}$ :

$$-\frac{dV(x)}{dx} = E = -\frac{V_{DS}}{L}. \quad (23)$$

Оценим электрическое поле в условиях отсечки канала,  $V_{DS} = V_{GS} - V_T$ . Из (20) имеем:

$$V(x) = (V_{GS} - V_T)(1 - \sqrt{1 - x/L'}) \quad (24)$$

и электрическое поле

$$E(x) = -\frac{(V_{GS} - V_T)}{2L'} \left[ \frac{1}{\sqrt{1 - x/L'}} \right]. \quad (25)$$

При  $x=0$  из (25) следует принятый ранее без доказательства результат (10), а при  $x=L'$  наступает отсечка и  $E(L') \rightarrow \infty$ . Это ожидаемый результат, поскольку в этой модели заряд обнуляется в области отсечки, так что для поддержки конечного тока требуется бесконечно большое поле.

## 2.4. Насыщение скорости и насыщение тока стока

Мы рассматриваем сейчас традиционный подход к теории MOSFET. Уравнения (5), (7) и (12) описывают ход выходной характеристики в линейной области и в области насыщения. Были изложены два различных подхода к описанию тока в области насыщения. В одном случае насыщение тока происходило за счет насыщения скорости электронов в той части канала, где электрическое поле велико. В другом случае причина насыщения была связана с развитием области отсечки канала вблизи стока. Когда среднее электрическое поле в канале намного больше критического поля для насыщения скорости ( $\approx 10$  кВ/см), тогда лучше работает первая модель:

$$\frac{V_{GS} - V_T}{2L} \gg E_{cr} \approx 10 \text{ кВ/см}. \quad (26)$$

Подставляя сюда типичные значения  $V_{GS} = V_{DD} = 1B$  и  $V_T = 0.2B$ , находим, что модель насыщения скорости приемлема при  $L \sim /< 1\mu m$ . Действительно, подобные модели впервые начали широко применяться в 80-х годах прошлого века когда длины каналов достигли одного микрометра [10].

На рис. 3 показаны выходные характеристики  $n$ -Si MOSFET, включенного по схеме с общим истоком, при  $L \approx 60nm$ . Очевидно, что при больших  $V_{DS}$  ток  $I_{DS} \propto V_{GS} - V_T$ , так что такой транзистор хорошо описывается моделью насыщения скорости (7).

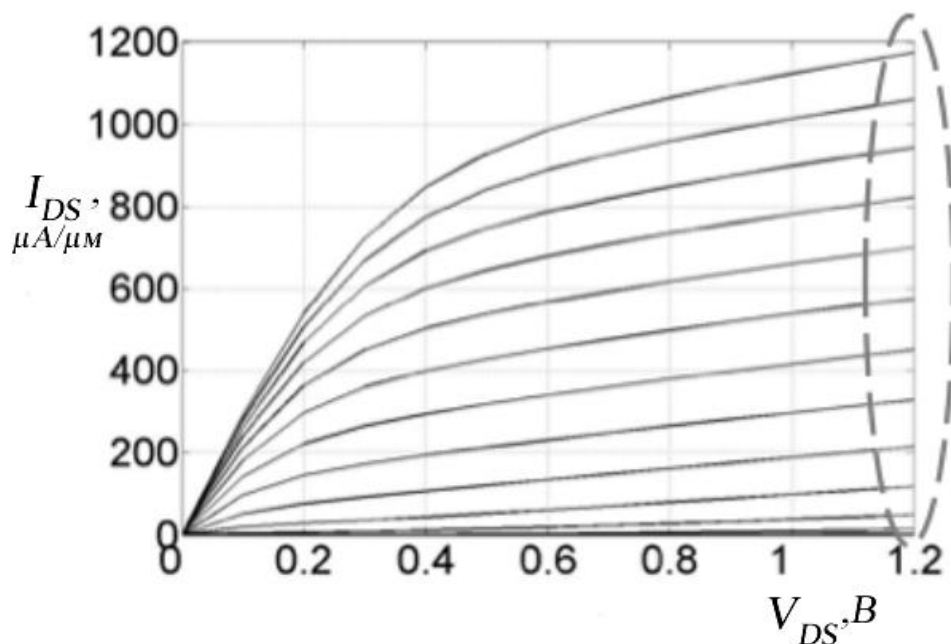


Рис. 3. Выходные характеристики  $n$ -Si MOSFET, включенного по схеме с общим истоком, при  $L \approx 60nm$ . Характеристики показаны с шагом  $0.1 B$ , для верхней кривой  $V_{GS} = 1.2B$ . Обращает на себя внимание что при больших  $V_{DS}$  ток растет линейно с ростом напряжения на затворе. Это признак того, что в канале имеет место насыщение скорости [11].

Для транзистора на этом рисунке  $V_T \approx 0.4B$ . Для верхнего графика с максимальным напряжением на затворе модель отсечки дает  $V_{DSAT} = V_{GS} - V_T = 1.2 - 0.4 \approx 0.8B$ , что слишком много и указывает на то, что ток стока не насыщается за счет классической отсечки. Обсуждение  $V_{DSAT}$  при насыщении скорости можно найти в [8, 9].

Хотя и кажется, что модели насыщения скорости удовлетворительно описывают короткоканальные MOSFET, но есть нечто настораживающее. Тщательное компьютерное моделирование транспорта электронов в

наноразмерных MOSFET неопровержимо показывает, что скорость не насыщается в той части канала нанотранзистора, которая находится под воздействием высокого электрического поля. Просто электронам не хватает времени эффективно рассеяться, чтобы их скорость насытилась. Электроны проскакивают канал и уходят в сток слишком быстро. И все же графики на рис. 3 показывают, что скорость таки насыщается. Это мистическое противоречие найдет свое объяснение когда мы перейдем к последовательному изложению теории нанотранзисторов.

## 2.5. Метрика нанотранзисторов

Уравнения (5) и (7) приемлемы для описания выходных характеристик современных короткоканальных транзисторов и могут быть использованы для оценки их метрических показателей (глава 1), а именно:

$$\begin{aligned}
 I_{ON} &= W C_{ox} v_{sat} (V_{DD} - V_T), \quad V_T = V_{T0} - \delta V_{DS}, \\
 R_{ON} &= \left( \left. \frac{\partial I_{DS}}{\partial V_{DS}} \right|_{V_{GS}=V_{DD}, V_{DS} \approx 0} \right)^{-1} = \left( \frac{W}{L} \mu C_{ox} (V_{GS} - V_T) \right)^{-1}, \\
 g_m^{sat} &= \left. \frac{\partial I_{DS}}{\partial V_{GS}} \right|_{V_{GS}=V_{DS}=V_{DD}} = W C_{ox} v_{sat}, \\
 r_d &= \left( \left. \frac{\partial I_{DS}}{\partial V_{DS}} \right|_{V_{GS}=V_{DD}, V_{DS} > V_{DSAT}} \right)^{-1} = \frac{1}{g_m^{sat} \delta}, \\
 |A_v| &= g_m^{sat} r_d = \delta^{-1}.
 \end{aligned} \tag{27}$$

Метрика  $|A_v|$  – важный показатель качества в аналоговой электронике.

Что же касается зонных диаграмм, то хотя они не использовались нами в явном виде, но они подразумевались в той или иной форме. Начало канала проводимости ( $x=0$ ) это вершина энергетического барьера на рисунках 16 и 17 главы 1 или близко к вершине барьера [13]. Как мы увидим позже, в качественном транзисторе заряд на вершине барьера дается выражением (2). Этот заряд обеспечивается электронами из истока, преодолевшими барьер. Эта локация на вершине барьера, для которой справедливо уравнение (2), известна также как виртуальный исток.

Использование зонных диаграмм особенно полезно для понимания природы отсечки канала. Не сразу понятно как именно электроны покидают канал проводимости и проходят зону отсечки. Проясняют ситуацию зонные диаграммы. Как видно из рис. 17/гл. 1, область отсечки – это область сильного электрического поля в районе стока, и наклон  $E_c(x)$  здесь наиболее крутой. Электроны, попавшие в эту область из канала проводимости, просто скатываются и покидают сток. Ничто не может их остановить, если они уже попали в область отсечки.

## 2.6. Начала модели виртуального истока

У нас уже есть выражения для тока стока в линейной области (5) и в области насыщения (7), а именно:

$$I_{DLIN} = \frac{W}{L} \mu C_{ox} (V_{GS} - V_T) V_{DS},$$

$$I_{DSAT} = W C_{ox} v_{sat} (V_{GS} - V_T),$$
(28)

при этом имеется в виду, что  $V_{GS} > V_T$ , так что уравнения (28) не пригодны для подпороговой области входной характеристики. Эти уравнения приближенно описывают поведение  $I_{DS} \propto V_{DS}$  (рис. 4); рекомендуется учесть *DIBL* как в (18/гл. 1). Уравнивание выражений (28),  $I_{DLIN} = I_{DSAT}$ , позволяет ввести понятие напряжения насыщения на стоке как

$$V_{DSAT} = \frac{v_{sat} L}{\mu}. \quad (29)$$

При  $V_{DS} \ll V_{DSAT}$  ток  $I_{DS} = I_{DLIN}$ , а при  $V_{DS} \gg V_{DSAT}$  ток  $I_{DS} = I_{DSAT}$ .

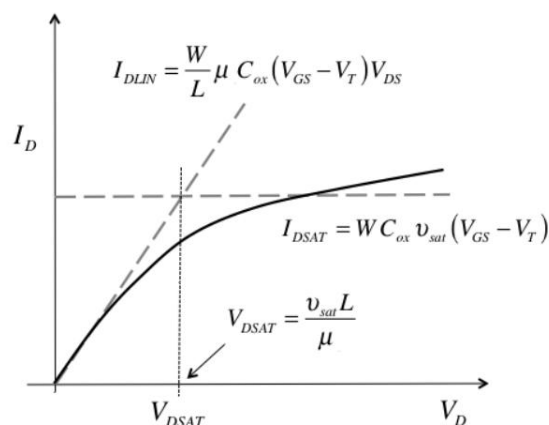


Рис. 4. Выходная характеристика *n*-MOSFET, включенного по схеме с общим истоком, при некотором значении напряжения на затворе. Пунктирные линии дают ток в линейной области и в области насыщения согласно (28).

Традиционный подход к теории MOSFET позволяет получить выражение для зависимости  $I_{DS} \propto V_{DS}$ , плавно переходящей от линейного участка к области насыщения по мере роста  $V_{DS}$  от нуля до  $V_{DD}$  [8, 9]. Наша задача – получить простое полуэмпирическое выражение для всей цельной выходной характеристики  $I_{DS}(V_{DS})$ . Используемый далее подход аналогичен так называемой модели виртуального истока (Virtual Source Model/VS Model) MOSFET, которая нашла успешное применение при моделировании весьма разнообразных нанотранзисторов [13]. Лундстром [14], воспользовавшись изложенным выше традиционным подходом, подошел несколько иначе к построению модели виртуального истока. По мере совершенствования и развития этой модели в конечном итоге будет построена достаточно простая физическая модель, которая обеспечит точное количественное описание современных нанотранзисторов.

#### *Скорость в канале и напряжение на стоке*

Ток стока пропорционален произведению заряда в начале канала и средней скорости электронов там же, а именно:

$$I_{DS}/W = |Q|_{x=0} |v|_{x=0}. \quad (30)$$

Уравнение (28) для линейного тока можно переписать в следующем виде:

$$\begin{aligned} I_{DLIN}/W &= |Q(V_{GS})| v(V_{DS}), \\ Q(V_{GS}) &= -C_{ox}(V_{GS} - V_T), \\ v(V_{DS}) &= \left( \mu \frac{V_{DS}}{L} \right). \end{aligned} \quad (31)$$

Аналогичным образом перепишем (28) для области насыщения:

$$\begin{aligned} I_{DSAT}/W &= |Q(V_{GS})| v(V_{DS}), \\ Q(V_{GS}) &= -C_{ox}(V_{GS} - V_T), \\ v(V_{DS}) &= v_{sat}. \end{aligned} \quad (32)$$

Если найти способ аппроксимировать плавный переход средней скорости от своего значения при низком  $V_{DS}$  до значения  $v_{sat}$  при высоком  $V_{DS}$ , то можно получить модель, которая покрывает весь интервал значений напряжения на стоке.

В оригинальной модели виртуального истока используется эмпирический подход, согласно которому средняя скорость в начале канала записывается следующим образом [13]:

$$\begin{aligned} v(V_{DS}) &= F_{SAT}(V_{DS})v_{sat}, \\ F_{SAT}(V_{DS}) &= \frac{V_{DS}N_{DSAT}}{\left[1 + (V_{DS}N_{DSAT})^\beta\right]^{1/\beta}}, \end{aligned} \quad (33)$$

где  $V_{DSAT}$  определяется уравнением (29), а эмпирический подгоночный параметр  $\beta$  позволяет вписаться в экспериментально измеренную выходную характеристику.

Выбор выражения для функции насыщения тока стока  $F_{SAT}$  был продиктован наблюдением того, что наименьшая из двух скоростей (31) и (32) и есть та скорость, которая ограничивает ток. Это дает основание написать, что

$$\frac{1}{v(V_{DS})} = \frac{1}{\mu V_{DS}/L} + \frac{1}{v_{sat}}, \quad (34)$$

или иначе

$$v(V_{DS}) = \frac{V_{DS}N_{DSAT}}{1 + (V_{DS}N_{DSAT})^\beta} v_{sat}. \quad (35)$$

Уравнение (35) аналогично уравнению (33) за исключением эмпирического параметра  $\beta$ . Типичные значения  $\beta$  для  $n$ - и  $p$ -Si MOSFET лежат в интервале 1.4 – 1.8 [12].

Уравнения (2), (30) и (33) описывают линейный ход и область насыщения входных характеристик MOSFET выше порогового значения напряжения для любого напряжения на стоке.

Итак, простую начальную модель MOSFET, пригодную выше порогового напряжения, можно суммировать в виде следующей сводки формул (36):

$$\begin{aligned} I_{DS}/W &= |Q(0)|v(0), \\ Q(V_{GS}) &= 0, \quad V_{GS} \leq V_T, \\ Q(V_{GS}) &= -C_{ox}(V_{GS} - V_T), \quad V_{GS} > V_T, \quad V_T = V_{T0} - \delta V_{DS}, \\ \langle v(V_{DS}) \rangle &= F_{SAT}(V_{DS})v_{sat}, \\ F_{SAT}(V_{DS}) &= \frac{V_{DS}N_{DSAT}}{\left[1 + (V_{DS}N_{DSAT})^\beta\right]^{1/\beta}}, \quad V_{DSAT} = \frac{v_{sat}L}{\mu}. \end{aligned} \quad (36)$$

В модели виртуального истока нулевого уровня есть только шесть параметров:

$$C_{ox}, V_T, \mu, v_{sat}, L, \beta. \quad (37)$$

Эта модель не пригодна для описания подпороговой области входных характеристик. Физически более корректное обсуждение электростатики MOS позволит нам позже количественно учесть и подпороговую область. Но уже на этом уровне модели виртуального истока можно корректно учесть реальные сопротивления терминалов  $S$  и  $D$  полевого транзистора.

## 2.7. Учет терминальных сопротивлений

На рис. 5 слева показан идеализированный MOSFET. На практике терминалы  $S$  и  $D$  полевого транзистора всегда характеризуются своими сопротивлениями. Они влияют на напряжения, подаваемые на транзистор, и сказываются на его характеристиках. Покажем как можно учесть это влияние.

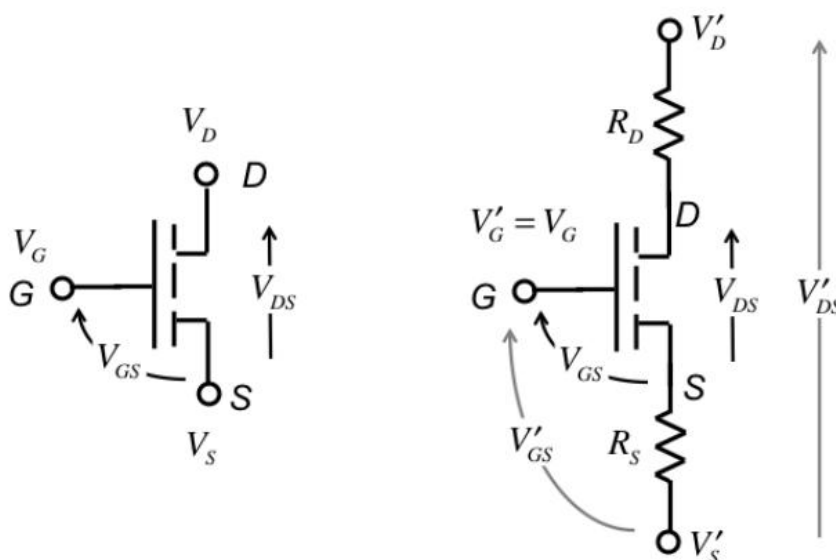


Рис. 5. Учет сопротивлений терминалов  $S$  и  $D$  MOSFET.

Слева: идеализированный MOSFET. Справа: реальный MOSFET, демонстрирующий влияние сопротивлений контактов  $S$  и  $D$  на напряжения в идеализированном полевом транзисторе.

На рис. 5 справа показаны принятые обозначения. Напряжениям  $V_D, V_S, V_G$  идеализированного транзистора соответствуют реальные напряжения  $V'_D, V'_S, V'_G$ . Сопротивление терминала затвора не учитывается, поскольку речь идет о режиме постоянного тока: ток затвора нулевой, так что сопротивление затвора не играет роли. Сопротивление затвора играет важную роль в радиочастотных транзисторах, которые нами не рассматриваются.

Связь между идеализированными, внутренними напряжениями (нештрихованными) и внешними напряжениями (штрихованными) дается простыми выражениями:

$$\begin{aligned} V_D &= V'_D - I_{DS}(V_D, V_S, V_G)R_D, \\ V_S &= V'_S + I_{DS}(V_D, V_S, V_G)R_S, \\ V_G &= V'_G. \end{aligned} \quad (38)$$

Пусть выходная характеристика идеализированного транзистора  $I_{DS}(V_D, V_S, V_G)$  известна. Задавшись измеренными сопротивлениями  $R_S$  и  $R_D$  и экспериментальными значениями напряжений  $V'_D, V'_S, V'_G$ , решение системы уравнений (38) позволяет вычислить внутренние напряжения  $V_D$  и  $V_S$  и определить характеристику  $I_{DS}(V'_D, V'_S, V'_G)$ .

Рисунок 6 иллюстрирует влияние учета сопротивлений контактов  $S$  и  $D$  на входную характеристику. Согласно (31), ток в линейной области без учета сопротивлений  $R_S$  и  $R_D$  дается выражением

$$I_{DLIN} = \frac{W}{L} \mu C_{ox} (V_{GS} - V_T) V_{DS} = V_{DS} / R_{ch}, \quad (39)$$

а с учетом сопротивлений контактов

$$I_{DLIN} = V_{DS} / R_{tot}, \quad (40)$$

где

$$R_{tot} = R_{ch} + R_S + R_D = R_{ch} + R_{DS}. \quad (41)$$

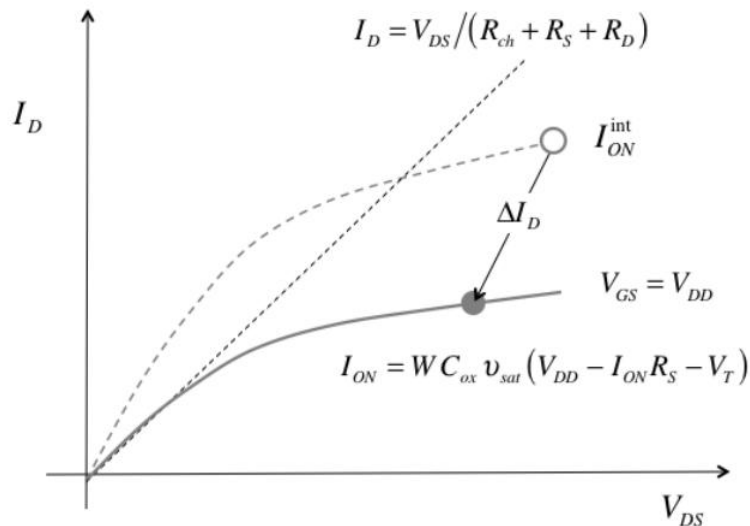


Рис. 6. Влияние учета сопротивлений контактов  $S$  и  $D$  на ВАХ MOSFET.

Пунктиром показана входная характеристика без учета сопротивлений контактов. Учет сопротивлений  $R_S$  и  $R_D$  увеличивает сопротивление канала проводимости и понижает ток  $I_{ON}$ .



Таким образом, влияние учета сопротивлений контактов сводится просто к уменьшению наклона линейного участка входной характеристики (рис. 6).

Как видим, уменьшается также предельный ток насыщения. В идеальном транзисторе сопротивление  $R_D$  не влияет на ту часть области насыщения характеристики, для которой  $V_{DS} > V_{DSAT}$  (рис. 4), а вот сопротивление  $R_S$  уменьшает напряжение на затворе  $V_{GS}$  на величину  $I_{DSAT}R_S$ , так что вместо (28) имеем

$$I_{DSAT} = W C_{ox} v_{sat} (V'_{GS} - I_{DSAT} R_S - V_T). \quad (42)$$

Максимальное значение напряжения  $V_{DS}$  есть напряжение питания транзистора  $V_{DD}$ . Влияние сопротивления контактов незначительно, если  $I_{DSAT} R_S \ll V_{DD}$ . Высокая эффективность транзистора обеспечивается значением

$$R_S \ll \frac{V_{DD}}{I_{DSAT}}. \quad (43)$$

Нынешние Si-MOSFET обеспечивают ток  $I_{ON} = 1 \text{ мА}/\mu\text{м}$  при  $V_{DD} = 1 \text{ В}$ . Соответственно этому, сопротивление  $R_S$  должно быть намного меньше  $1000 \text{ Ом} \cdot \mu\text{м}$ , скажем, равно  $100 \text{ Ом} \cdot \mu\text{м}$ . Хотя нас прежде всего интересует физика идеализированного MOSFET, нужно, однако, при анализе экспериментальных данных помнить о роли сопротивлений контактов. По мере того как длина канала проводимости в транзисторах уменьшается, поддерживать учет сопротивлений контактов на должном уровне становится все сложнее и сложнее.

Воспользуемся уравнением (42) и оценим скорость насыщения  $v_{sat}$  тока  $I_{ON}$  путем анализа входной характеристики MOSFET с  $L \approx 60 \text{ нм}$  на рис. 3. Полученное таким образом значение  $v_{sat}$  как подгоночного эмпирического параметра сравним с известными данными для массивного кремния. Исходные параметры входной характеристики (рис. 3) следующие:

$$\begin{aligned} I_{ON} &= 1180 \text{ } \mu\text{А}/\mu\text{м}, \\ C_{ox} &= 1.55 \times 10^{-6} \text{ Ф}/\text{см}^2, \\ R_{DS} &= 220 \text{ Ом}, \\ V_T &= 0.25 \text{ В}, \\ V_{DD} &= 1.2 \text{ В}, \\ W &= 1 \text{ } \mu\text{м}. \end{aligned}$$

Из (42) для  $v_{sat}$  имеем:

$$v_{sat} = \frac{I_{DSAT}}{W C_{ox}(V_{GS} - V_T)} \equiv v_{inj}, \quad (44)$$

$$V_{GS} = V_{DD} - I_{DSAT} R_{SD} / 2.$$

Для приведенных выше исходных данных получаем  $v_{sat} = 0.92 \times 10^7 \text{ см/с}$ .

Любопытно, что полученное значение скорости насыщения тока в нанотранзисторе практически совпадает с данными для массивного образца кремния ( $1 \times 10^7 \text{ см/с}$ ), однако, физика насыщения тока в нанотранзисторах весьма отличается от таковой в массивном кремнии. По этой причине, а вернемся мы к этому вопросу позже, переименуем скорость насыщения  $v_{sat}$  на скорость впрыска (injection velocity)  $v_{inj}$ .

Казалось бы можно ожидать, что традиционная модель MOSFET применима только для транзисторов с достаточно длинными каналами проводимости, поскольку она основана на таких предположениях как диффузионный транспорт в омической области входной характеристики и насыщение скорости в высоких полях. Удивительно, однако, оказывается, что эта модель достаточно аккуратно описывает входные характеристики MOSFET с длиной канала меньшей 100 нм, как это видно из рис. 7.

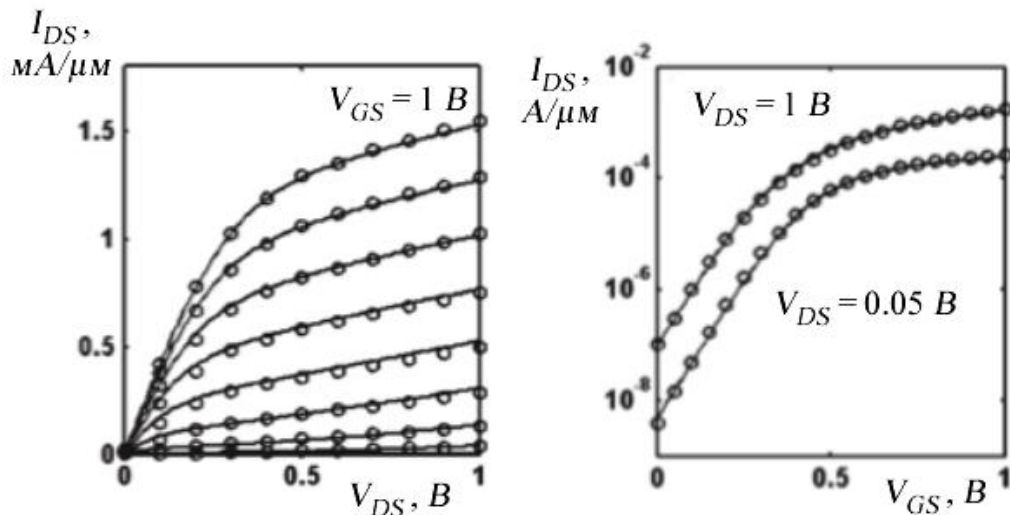


Рис. 7. Измеренные и рассчитанные по модели VS BAX 32-нм *n*-MOSFET. Слева: Выходные характеристики MOSFET, включенного по схеме с общим истоком, с шагом  $\Delta V_{GS} = 0.1 \text{ В}$ . Справа: Передаточные характеристики. Использовалась модель VS более высокого уровня [13], одинаково пригодная как для подпороговых, так и для надпороговых напряжений и подробно излагаемая далее.

Для достижения такого согласия с экспериментальными данными оказалось достаточным двух физических параметров в модели VS, считая их эмпирическими параметрами, всего лишь несколько модифицировать их таким образом, чтобы модель стала пригодной для большинства типов транзисторов. Эти два подгоночных параметра – скорость впрыска  $v_{inj}$  (скорость насыщения  $v_{sat}$  в традиционной модели) и кажущаяся подвижность  $\mu_{app}$  (просто подвижность  $\mu$  в традиционной модели). То обстоятельство, что простая традиционная модель MOSFET вполне удовлетворительно описывает даже современные транзисторы, говорит, скорее всего, о том, что эта модель на самом деле учитывает нечто существенное из физики нанотранзисторов, к более глубокому изучению которой мы перейдем позже.

## 2.8. Приближение истощения и уравнение Пуассона

До сих пор мы обсуждали основные концепции базовой модели MOSFET. Предполагая, что инверсионный заряд в самом начале канала проводимости дается выражениями

$$\begin{aligned} Q(V_{GS}) &= 0, & V_{GS} &\leq V_T, \\ Q(V_{GS}) &= -C_G(V_{GS} - V_T), & V_{GS} &> V_T, \\ V_T &= V_{T0} - \delta V_{DS}, \end{aligned} \quad (45)$$

и используя простые традиционные модели для средней скорости электронов в начале канала, были получены уравнения для моделирования входных характеристик MOSFET. Зададим себе следующие важные вопросы. Почему заряд  $Q$  растет линейно с напряжением на затворе при  $V_{GS} > V_T$ ? Что такое емкость затвора  $C_G$  (она несколько меньше  $C_{ox}$ )? Как небольшой заряд при  $V_{GS} < V_T$  изменяется с напряжением на затворе? Ответы на эти вопросы мы получим рассматривая 1D электростатику MOS. Кроме того возникают вопросы по физике *DIBL*: что именно определяет значение параметра  $\delta$ ? Для ответа на этот вопрос нам нужно будет обратиться уже к 2D электростатике MOS. Понимание 1D и 2D электростатики MOS абсолютно необходимо для построения теории полевых транзисторов, поскольку именно электростатика определяет как именно напряжения на терминалах транзистора влияют на барьер канала проводимости [6, 7].

## Энергия и изгиб зон

Мы хотим понять каким образом терминальные напряжения и собственно даже дизайн MOSFET влияют на электростатический потенциал  $\psi(x, y, z)$ . Направления осей системы координат выберем следующим образом: ось  $x$  направлена от истока к стоку, ось  $y$  – вглубь полупроводниковой подложки, ось  $z$  – вдоль ширины канала проводимости (рис. 8). Потребуется искать решение уравнения Пуассона

$$\begin{aligned}\nabla \cdot \vec{D}(x, y, z) &= \rho(x, y, z), \\ \nabla^2 \psi(x, y, z) &= -\frac{\rho(x, y, z)}{\epsilon_s},\end{aligned}\quad (46)$$

где  $\vec{D}$  – вектор смещения,  $\rho$  – зарядовая плотность,  $\epsilon_s$  – диэлектрическая проницаемость полупроводника, который предполагается пространственно однородным.

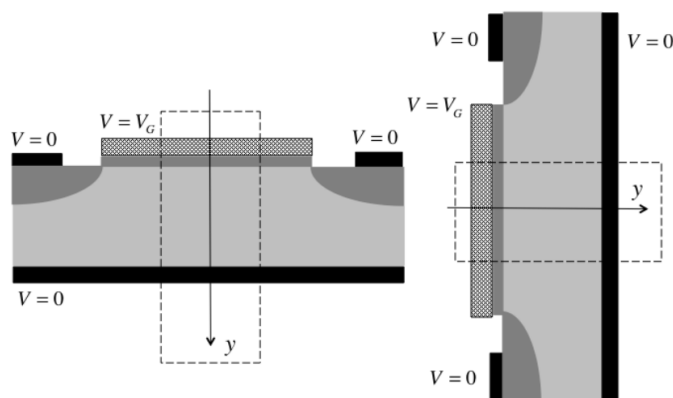


Рис. 8. Выбор направления оси  $y$  вглубь полупроводниковой подложки (слева) и ориентация канала проводимости (снизу – вверх) при изображении зонных диаграмм (справа).

В общем случае нужно искать пространственное решение, однако, обычно предполагается, что канал проводимости однороден по ширине (вдоль ось  $z$ ), так что достаточно 2D решения. Мы начнем обсуждение с 1D электростатики – в направлении, перпендикулярном каналу вглубь полупроводниковой подложки. Итак, имеем достаточно длинный канал и рассмотрим поведение  $\psi(y) \propto y$  в середине канала, где влияние потенциалов со стороны истока и стока минимально, так что 2D эффектами (вдоль оси  $x$ ) можно пренебречь.

Удобным количественным решением уравнения Пуассона служат диаграммы зонной энергии. Рассмотрим влияние напряжения на затворе на зависимость  $E(y)$  вглубь канала проводимости в полупроводниковой подложке. Рисунок 9 иллюстрирует ситуацию когда зоны плоские, т. е. потенциал нулевой

(или постоянный в направлении  $y$ ) и зонная энергия не зависит от координаты (глубины  $y$ ).

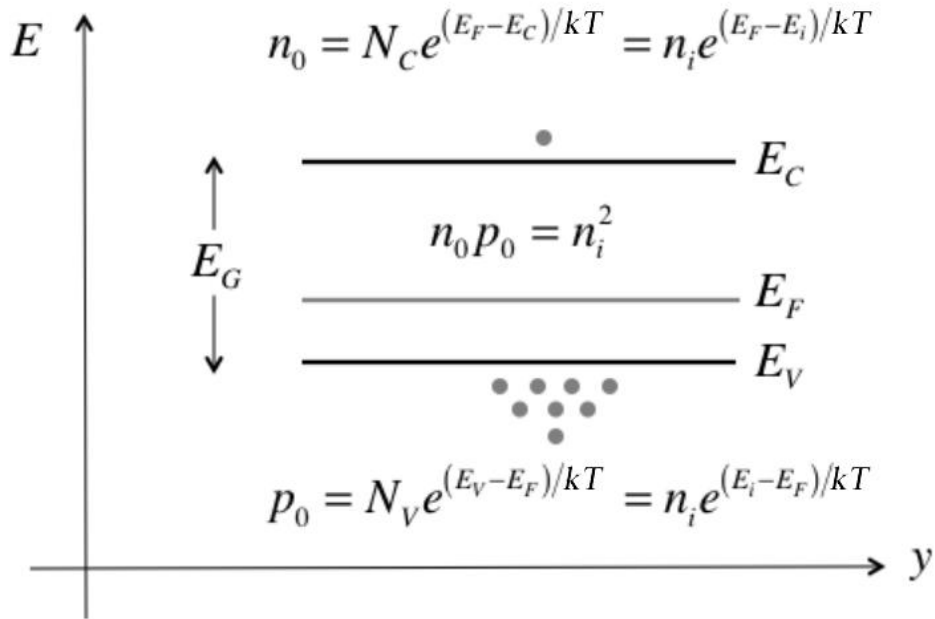


Рис. 9. При постоянном электростатическом потенциале зоны плоские. Показана ширина запрещенной зоны  $E_G$ .

Плотности электронов  $n_0$  и плотности дырок  $p_0$  зависят экспоненциально от разности между границами зон и энергией Ферми, предполагая что носители тока подчиняются статистике Больцмана. Предэкспоненты

$$N_C \equiv 2 \left[ \frac{2\pi m_n kT}{h^2} \right]^{3/2}, \quad N_V \equiv 2 \left[ \frac{2\pi m_p kT}{h^2} \right]^{3/2} \quad (47)$$

носят названия эффективные плотности состояний электронов в зоне проводимости и эффективные плотности состояний дырок в валентной зоне. Уравнения для плотностей электронов  $n_0$  и дырок  $p_0$  легко запомнить, если понять почему  $N_C$  и  $N_V$  называют эффективной плотностью состояний. Это как если бы все уровни энергии, например, в зоне проводимости эффективно совместить с энергией дна зоны проводимости  $E_C$  поместив туда все  $N_C$  электронов. В результате концентрация электронов  $n_0$  есть просто произведение  $N_C$  и вероятности того, что состояние с энергией  $E_C$  занято.

Значения  $N_C$  и  $N_V$  колеблются около  $10^{19} \text{ см}^{-3}$ , отличаясь только за счет различия в эффективных массах электронов  $m_n$  и дырок  $m_p$ . Различны они и для разных полупроводников (табл. 1).

Таблица 1  
Значения  $N_C$  и  $N_V$  для Ge, Si и GaAs при  $300 \text{ K}$ ,  $\text{см}^{-3}$

|       | Ge                    | Si                    | GaAs                 |
|-------|-----------------------|-----------------------|----------------------|
| $N_C$ | $1.04 \times 10^{19}$ | $2.8 \times 10^{19}$  | $4.7 \times 10^{17}$ |
| $N_V$ | $6.0 \times 10^{18}$  | $1.04 \times 10^{19}$ | $7.0 \times 10^{18}$ |

Концентрации мажоритарных носителей тока даются выражениями

$$\begin{aligned} n_0 &= N_D - N_A, \\ p_0 &= N_A - N_D, \end{aligned} \quad (48)$$

где  $N_D$  и  $N_A$  – концентрации доноров и акцепторов, предполагая обычно их полную ионизацию. Концентрации миноритарных носителей тока находятся из

$$n_0 p_0 = n_i^2, \quad (49)$$

где  $n_i$  есть концентрация собственных носителей, составляющая для Si около  $10^{10} \text{ см}^{-3}$  при  $300 \text{ K}$  и зависящая от  $E_G$  и  $T$ .

Рисунок 10 показывает как влияет положительный потенциал на затворе на ход электростатического потенциала. Потенциал  $\psi_s$  на поверхности полупроводника будет положительным:  $0 < \psi_s < V_{GS}$ . Если подложка полупроводника заземлена ( $\psi|_{y \rightarrow \infty} = 0$ ), можно ожидать, что электростатический потенциал экспоненциально ниспадает до нуля.

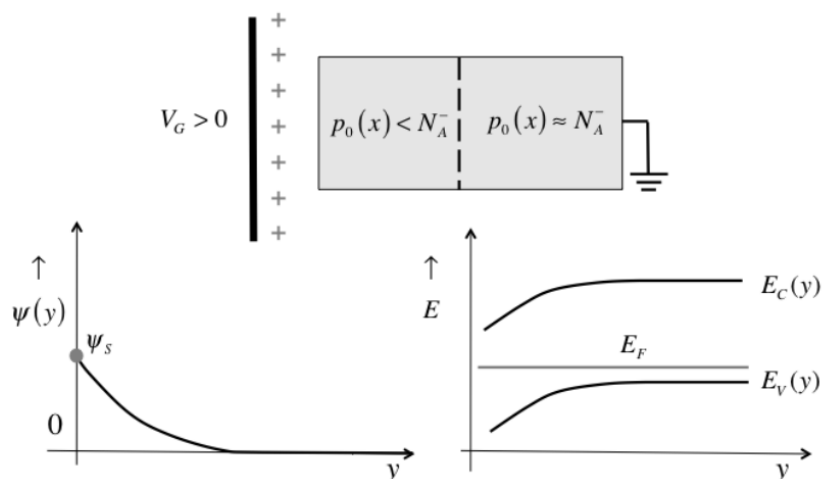


Рис. 10. Влияние положительного напряжения на затворе ( $V_{GS} > 0$ ) на ход электростатического потенциала (слева) и изгиб зон (справа).

Положительный электростатический потенциал понижает потенциальную энергию электронов, так что зоны изгибаются с изменением электростатического потенциала согласно

$$E_c(y) = \text{Const} - q\psi(y). \quad (50)$$

Если электростатический потенциал растет изнутри подложки по направлению к ее поверхности, то энергия зон в том же направлении падает (рис. 10).

Прежде чем переходить к рассмотрению изгиба зон в зависимости от приложенного напряжения на затворе определим некоторые понятия (рис. 11). Для начала рассмотрим гипотетический идеальный «металлический» затвор, фермиевский уровень  $E_{FM}^0$  которого совпадает с таковым  $E_F$  кремниевой подложки (рис. 11/слева). Эту ситуацию называют условием плоских зон (flatband/FB condition) – зоны в кремнии и в его диэлектрическом слое плоские.

Для некоторого гипотетического «металлического» затвора плоские зоны будут при  $V_G' = 0$ . Штрихом подчеркивается то обстоятельство, что затвор сделан из некоторого гипотетического вещества. На практике всегда имеет место разность работ выхода  $\Phi_{MS}$  между металлическим электродом затвора и полупроводниковой подложкой. Условие плоских зон не выполняется при  $V_G = 0$  в отличие от напряжения  $V_G = V_{FB} = \Phi_{MS}/q$  такого, которое нивелирует разность работ выхода.

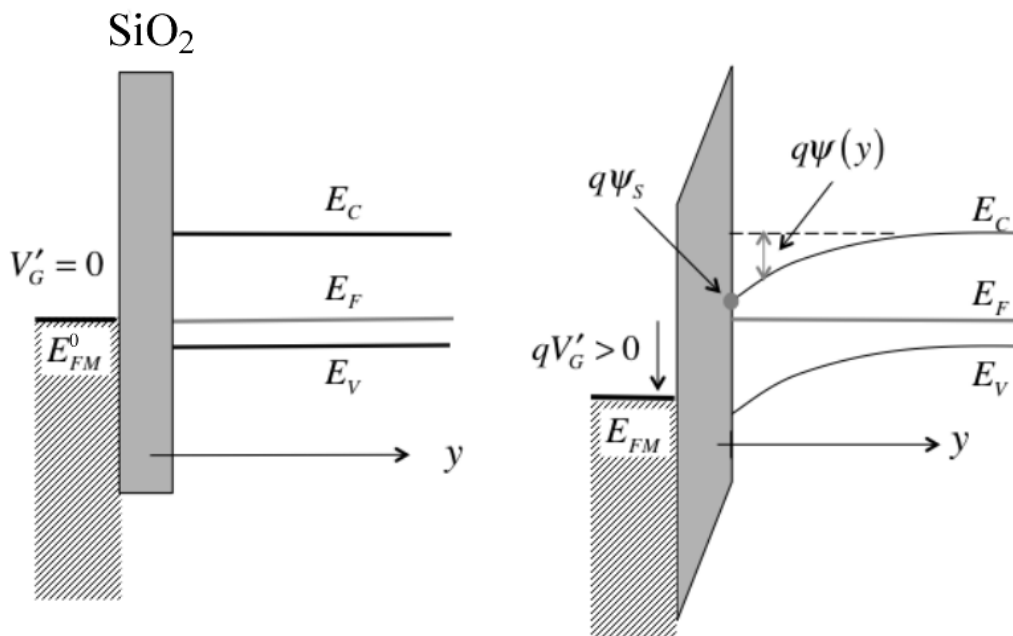


Рис. 11. Слева: Выполнение условия плоских зон в идеализированной структуре MOS. Справа: Изгиб зон в результате подачи положительного напряжения на затвор.

Напомним, что при подаче напряжения на контакт его уровень Ферми понижается. Как показано на рис. 11, фермиевский уровень электрода затвора понижается от значения  $E_{FM}^0$  при  $V'_G = 0$  до значения  $E_{FM} = E_{FM}^0 - qV'_G$ . Положительный потенциал, поданный на затвор, понижает электростатический потенциал в окиси кремния и в полупроводниковой подложке как это следует из решений уравнений Лапласа и Пуассона, к рассмотрению которых мы приступим ниже. Если за ноль отсчета электростатического потенциала принять его значение в массивной подложке,  $\psi|_{y \rightarrow \infty} = 0$ , то электростатический потенциал на глубине  $y$  от поверхности подложки дается выражением

$$\psi(y) = \frac{E_C(\infty) - E_C(y)}{q}. \quad (51)$$

Обращает на себя внимание (рис. 11/справа), что уровень Ферми в полупроводнике плоский даже если приложено напряжение на затворе. Связано это с тем, что диэлектрическая прокладка не дает току затвора пройти в подложку, так что металлический контакт затвора и полупроводниковая подложка порознь находятся в равновесии с различными фермиевскими уровнями (рис. 11/справа).

Теперь мы готовы обсудить изгиб зон при подаче на затвор напряжения разных знаков (рис. 12).

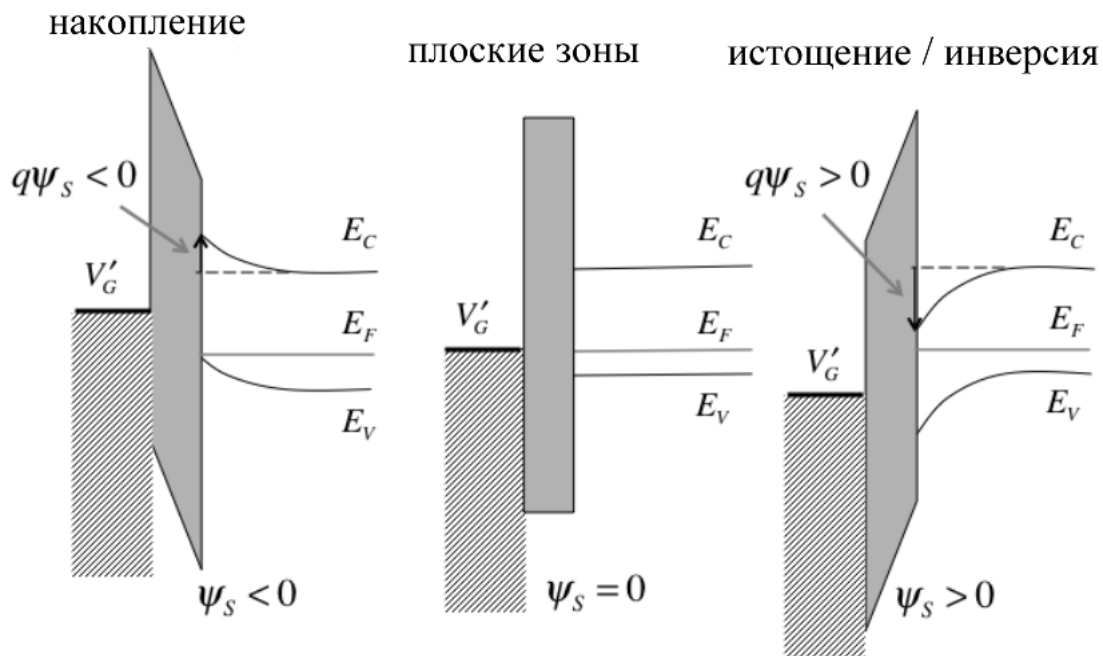


Рис. 12. Характер изгиба зон при трех различных напряжениях на затворе: накопление носителей (слева  $V_{GS} < 0$ ), истощение/инверсия (справа  $V_{GS} > 0$ ), плоские зоны (по центру  $V_{GS} = 0$ ).



Если подается отрицательное напряжение на затвор, в диэлектрическом слое и в полупроводнике индуцируется отрицательный электростатический потенциал и зоны изгибаются вверх, поверхностный потенциал отрицательный:  $\psi|_{y=0} = \psi_s < 0$ . Концентрация дырок увеличивается вблизи границы раздела диэлектрик/полупроводник, поскольку валентная зона изгибается вверх навстречу уровню Ферми. Суммарный заряд вблизи поверхности полупроводника положительный. Этот накопленный заряд (accumulation charge) локализуется очень близко к поверхности подложки и иногда даже аппроксимируется  $\delta$ -функцией.

При подаче положительного напряжения на затвор в диэлектрическом слое и в полупроводнике индуцируется положительный электростатический потенциал и зоны изгибаются вниз, поверхностный потенциал положительный:  $\psi|_{y=0} = \psi_s > 0$ . Поскольку валентная зона изгибается вниз от уровня Ферми (рис. 12/справа), концентрация дырок уменьшается. Легко запомнить в качестве мнемонического правила: положительный потенциал затвора отталкивает положительно заряженные дырки от поверхности раздела диэлектрик/подложка. Нужно только помнить, что «дырки» – это искусственное, а не физическое понятие, вместе с тем чрезвычайно удобное и ток в валентной зоне это поток электронов (как и в зоне проводимости), а не умозрительных «дырок». Результатом подачи положительного напряжения на затвор является появление слоя истощения (depletion layer), слоя толщиной  $W_D$ , в котором концентрация дырок пренебрежимо мала,  $p_0 \ll N_A^-$  (рис. 10). Если изгиб зон не слишком велик, то концентрация электронов также невелика. В противном случае, при значительном изгибе зон электроны начинают накапливаться вблизи поверхности раздела образуя инверсионный слой (inversion layer), связывающий  $n$ -допированные области контактов  $S$  и  $D$  и ответственный за возникновение тока в MOSFET. Явление инверсии подробно обсудим позже.

Наконец, нужно помнить, что зонные диаграммы на рис. 12 относятся к подложке из  $p$ -Si. Эти диаграммы легко перерисовать для подложки  $n$ -типа. Термины «накопление» и «истощение» всегда указывают на накопление и истощение основных носителей тока, а термин «инверсия» всегда относится к миноритарным носителям тока. Нужно только помнить, что отнесение носителей тока к основным или миноритарным является искусственным и условным, однако, весьма удобным, поскольку электрический ток – это всегда поток электронов.

Забегая несколько вперед, приведем сводную таблицу трех режимов – накопления, истощения и инверсии, а также двух переходных локаций – плоских зон и порога.

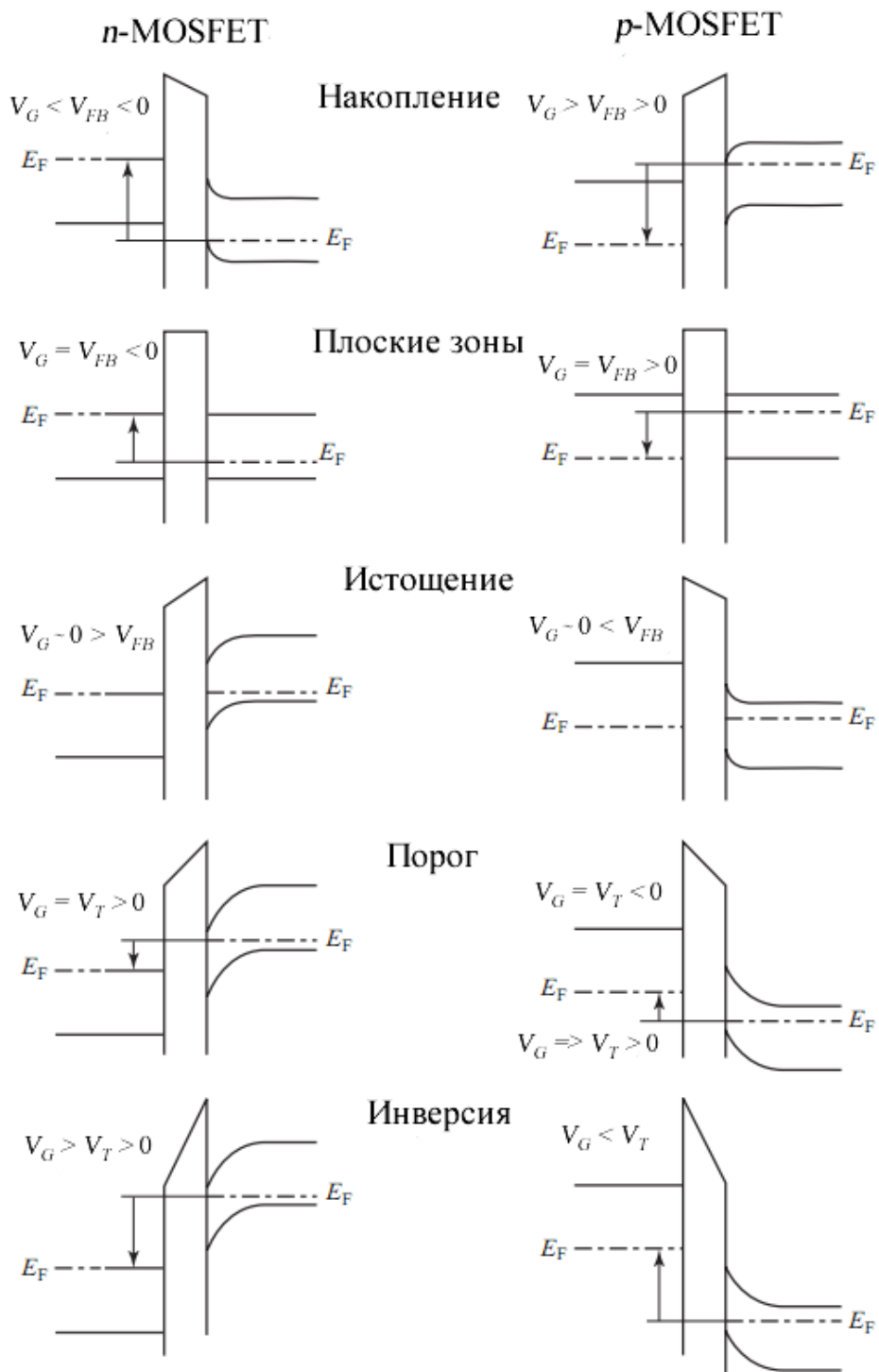


Рис. 13. Характер изгиба зон в режимах накопления, истощения и инверсии, а также в двух переходных локациях – плоских зон и порога. Отрицательному напряжению на затворе соответствуют стрелки «вверх», а положительному – стрелки «вниз» [15].

## Уравнение Пуассона – Больцмана

Наша задача состоит в том, чтобы понять как заряд в полупроводнике

$$Q_S = \int_0^{\infty} \rho(y) dy = q \int_0^{\infty} (p_0(y) - n_0(y) + N_D^+ - N_A^-) dy \quad [K/M^2] \quad (52)$$

зависит от электростатического потенциала  $\psi$ . Нижний индекс «0» означает, что полупроводниковая подложка находится в равновесии. Нас также интересует заряд, обязанный подвижным электронам,

$$Q = -q \int_0^{\infty} n_0(y) dy, \quad [K/M^2] \quad (53)$$

поскольку именно электроны являются носителями тока в MOSFET.

Количественным решением для потенциала и заряда в полупроводнике служат диаграммы зонной структуры. Для нахождения хода потенциала в полупроводнике нужно решать уравнение Пуассона. Запишем это уравнение для 1D полупроводника:

$$\frac{d^2\psi}{dy^2} = \frac{-q}{\varepsilon_S} (p_0(y) - n_0(y) + N_D^+ - N_A^-). \quad (54)$$

Имеется в виду полупроводник  $p$ -типа, для которого  $N_D = 0$ . Подразумевается полная ионизация допантов:  $N_A^- = N_A$ . Подложка ( $B$ ) не заряжена,  $p_B - n_B - N_A = 0$ , так что  $N_A = p_B - n_B$ , и уравнение Пуассона теперь

$$\frac{d^2\psi}{dy^2} = \frac{-q}{\varepsilon_S} (p_0(y) - n_0(y) + n_B - p_B), \quad (55)$$

где

$$p_B \cong N_A, \quad n_B \cong n_i^2/N_A. \quad (56)$$

Индексом « $B$ » подчеркивается, что речь идет о равновесных концентрациях в подложке. Учитывая (56), переписываем (55) в виде

$$\frac{d^2\psi}{dy^2} = \frac{-q}{\varepsilon_S} (p_0(y) - N_A - n_0(y) + n_i^2/N_A). \quad (57)$$

В этом уравнении три неизвестных:  $\psi(y)$ ,  $n_0(y)$  и  $p_0(y)$ . Нужны еще два уравнения.

Вспомним, что MOSFET находится в равновесии при любом напряжении на затворе, поскольку диэлектрическая прокладка в идеале не пропускает ток. А в равновесии плотности носителей тока зависят от положения уровня Ферми (константа в равновесном состоянии) и от границ зон. Поэтому можно записать, что

$$\begin{aligned} n_0(y) &= n_B e^{+q\psi(y)/kT} = \frac{n_i^2}{N_A} e^{+q\psi(y)/kT}, \\ p_0(y) &= p_B e^{-q\psi(y)/kT} = N_A e^{-q\psi(y)/kT}, \end{aligned} \quad (58)$$

и окончательно уравнение (57) принимает вид:

$$\frac{d^2\psi}{dy^2} = \frac{-q}{\epsilon_S} \left[ N_A \left( e^{-q\psi(y)/kT} - 1 \right) - \frac{n_i^2}{N_A} \left( e^{+q\psi(y)/kT} - 1 \right) \right]. \quad (59)$$

Это уравнение известно как уравнение Пуассона – Больцмана. В нашем случае оно описывает 1D полупроводник  $p$ -типа в равновесии, с полностью ионизированными допантами. Осталось записать граничные условия. Имея в виду полубесконечный полупроводник, записываем

$$\begin{aligned} \psi|_{y=0} &= \psi_S, \\ \psi|_{y \rightarrow \infty} &= 0. \end{aligned} \quad (60)$$

На практике значение  $\psi_S$  обычно выбирают в соответствии с потенциалом затвора.

С методами решения нелинейного дифференциального уравнения Пуассона – Больцмана можно подробно ознакомиться в [8, 9, 16]. Можно получить его приближенные решения, если полупроводниковая подложка MOSFET находится в режиме накопления или истощения или сильной инверсии. Начнем мы с режима истощения.

### *Приближение истощения*

Весьма достоверное приближенное решение для электростатического потенциала и электрического поля можно получить, если транзистор находится в режиме истощения. В этом режиме зоны изгибаются вниз и концентрация дырок пренебрежимо мала на глубине  $y \sim \ll W_D$ . В режиме истощения дно зоны проводимости расположено все еще достаточно высоко относительно уровня

Ферми (это не так в режиме инверсии), и концентрация электронов, как и дырок, мала. В итоге, зарядовая плотность

$$\rho(y) = q(p_0(y) - n_0(y) + N_D^+ - N_A^-) \quad (61)$$

существенно упрощается. Пренебрегая незначительной концентрацией носителей тока, предполагая наличие только дорантов  $p$ -типа и их полную ионизацию, уравнение (61) упрощается до

$$\rho(y) = \begin{cases} -qN_A, & (y < W_D) \\ 0, & (y \geq W_D) \end{cases} \quad (62)$$

Это приближение истощения обычно достаточно неплохое и вместе с тем простое и позволяет получить аналитические решения.

Рисунок 14 показывает зонную диаграмму в режиме истощения и соответствующий ход электрического поля в  $n$ -MOSFET, которое получается в результате решения уравнения Пуассона:

$$\begin{aligned} \frac{dD}{dx} &= \frac{d(\varepsilon_s E)}{dx} = \varepsilon_s \frac{dE}{dx} = \rho(y) = -qN_A, \\ \frac{dE}{dx} &= \frac{-qN_A}{\varepsilon_s}. \end{aligned} \quad (63)$$

Если плотность допирования однородная, тогда электрическое поле спадает в глубину подложки линейно с отрицательным наклоном (рис. 14/справа):

$$E(y) = \frac{qN_A}{\varepsilon_s} (W_D - y). \quad (64)$$

Важной величиной является электрическое поле на поверхности полупроводниковой подложки

$$\underline{E}|_{y=0} = E_s = \frac{qN_A}{\varepsilon_s} W_D. \quad (65)$$

Электростатический потенциал определяется через электрическое поле, а именно:

$$\psi(y) = - \int_{\infty}^y E(y') dy'. \quad (66)$$

Падение потенциала на протяжении всей области истощения  $\psi_s$  (рис. 14/слева) есть площадь под кривой  $E(y) \propto y$ , а именно:

$$\psi_s = \frac{1}{2} E_s W_D, \quad (67)$$

откуда, используя (65), получаем важный результат:

$$W_D = \sqrt{\frac{2\epsilon_s \psi_s}{qN_A}}. \quad (68)$$

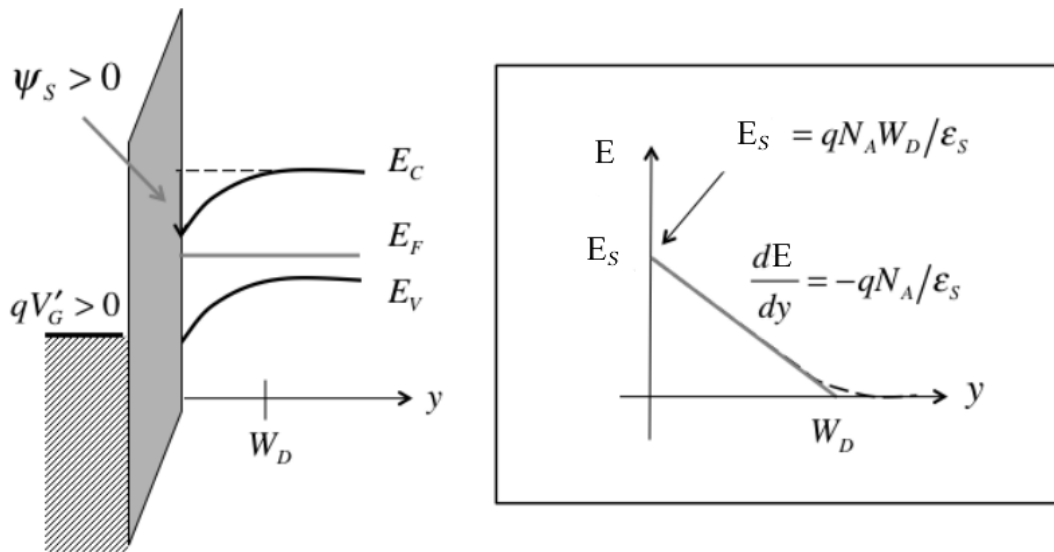


Рис. 14. Иллюстрация эффекта истощения в  $n$ -MOSFET.

Слева: зонная диаграмма. Справа: Ход электрического поля внутри подложки (пунктирная линия), а прямая линия соответствует приближению истощения.

Полный заряд в полупроводниковой подложке

$$Q_s = \int_0^{\infty} \rho(y) dy \approx Q_D = -qN_A W_D = \epsilon_s E_s, \quad [K/m^2] \quad (69)$$

а из (68) и (69) получаем еще один важный результат:

$$Q_D \approx -\sqrt{2qN_A \epsilon_s \psi_s}. \quad (70)$$

Обратим внимание, что в режиме истощения полный заряд в подложке  $Q_s$  практически совпадает с зарядом в слое истощения  $Q_D$ .

Когда полупроводниковая подложка транзистора путем подачи напряжения переведена в режим истощения, подробно рассмотренный выше,

приближение истощения в теории MOSFET обеспечивает достаточно точные решения как для распределения электростатического потенциала, так и электрического поля. Мы рассматривали подложку  $p$ -типа. Соответствующие уравнения легко записать и для подложки  $n$ -типа. Однако, не все так просто с режимами накопления и инверсии.

### Инверсия

Рисунок 15 показывает зонные диаграммы и изменение заряда в условиях инверсии: большой потенциал  $\psi_s$  изгибает зону проводимости у поверхности раздела очень близко к уровню Ферми (рис. 15/слева), так что концентрация электронов у поверхности становится значительной (рис. 15/справа). Связь между концентрациями электронов у поверхности и в глубине подложки дается выражением (58). Зададимся вопросом: каким должен быть потенциал  $\psi_s$ , чтобы область под поверхностью раздела была  $n$ -типа, а в глубине подложки –  $p$ -типа? Из уравнения (58)

$$\underline{n_0}|_{y=0} = \frac{n_i^2}{N_A} e^{+q\psi_s/kT} = N_A \quad (71)$$

получаем ответ:

$$\begin{aligned} \psi_s &= 2\psi_B, \\ \psi_B &= \frac{kT}{q} \ln\left(\frac{N_A}{n_i}\right). \end{aligned} \quad (72)$$

Уравнение (72) подводит итог инверсии: для потенциалов  $\psi_s$ , бóльших  $2\psi_B$ , вблизи поверхности раздела диэлектрика с полупроводником  $p$ -типа реализуется слой  $n$ -типа. Итак, напряжение на затворе, создающее потенциал  $\psi_s$ , превосходящий  $2\psi_B$ , приводит к появлению канала проводимости  $n$ -типа, соединяющего области  $n$ -типа  $S$  и  $D$  и переводящего транзистор в режим «ON». Напряжение на затворе, необходимое для создания такого поверхностного потенциала  $\psi_s$ , и есть пороговое напряжение.

В условиях инверсии область истощения достигает глубины (рис. 15/справа)

$$W_T = W_D(2\psi_B) = \sqrt{\frac{2\varepsilon_s(2\psi_B)}{qN_A}}. \quad (73)$$

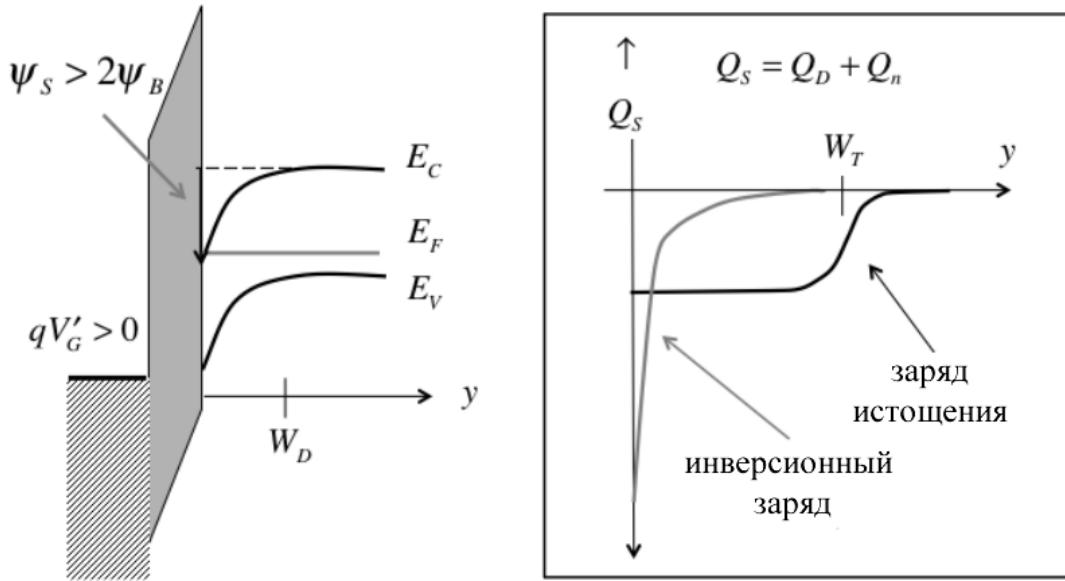


Рис. 15. Режим инверсии в полупроводнике. Слева: зонная диаграмма. Справа: Ход заряда внутрь подложки. Инверсионный заряд резко возрастает вблизи поверхности раздела.

Полный заряд на единицу поверхности в области истощения

$$Q_D = -qN_A W_T. \quad [K/m^2] \quad (74)$$

Есть также значительный заряд, вызванный электронами инверсионного слоя, которые концентрируются вблизи поверхности раздела  $SiO_2/p-Si$ ,

$$Q_n = q \int_0^{\infty} n_0(y) dy. \quad [K/m^2]. \quad (75)$$

Таким образом, полный заряд в полупроводнике в условиях инверсии (рис. 15/справа) есть

$$Q_s = Q_D + Q_n. \quad (76)$$

Ток в MOSFET проводит только инверсионный слой. Позже мы установим связь между зарядом в инверсионном слое и напряжением на затворе.

### *Влияние подложки*

До сих пор мы изучали электростатику MOS в середине достаточно длинного канала проводимости, где побочные электрические поля, создаваемые  $p-n$ -переходами, были малы, так что рассмотрения 1D электростатики казалось



достаточным. Но даже в этом случае  $p$ - $n$ -переходы оказывают существенное влияние, иллюстрацией чего может служить рис. 16.

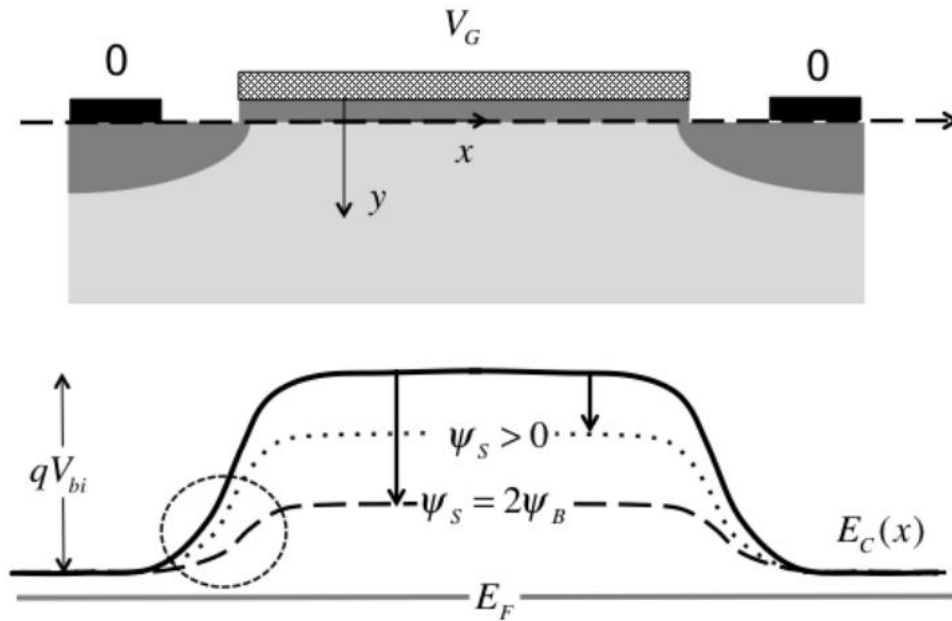


Рис. 16. Зона проводимости вдоль канала на поверхности подложки ( $y = 0$ ) от истока до стока. Сплошная линия соответствует условию плоских зон в подложке. Точечная линия соответствует поверхностному потенциалу  $\psi_s > 0$ , а пунктирная линия соответствует  $\psi_s = 2\psi_B$ .

Высота барьера в условиях нулевых напряжений на терминалах транзистора есть просто  $qV_{bi}$ , где встроенный потенциал (built-in potential)  $p$ - $n$ -переходов в стандартной теории полупроводников [6, 7] дается выражением

$$V_{bi} = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2}. \quad (77)$$

Этот энергетический барьер значительный, так что лишь считанные электроны со стороны как истока, так и стока могут его преодолеть.

Пунктирная линия на рис. 16 соответствует поверхностному потенциалу  $\psi_s = 2\psi_B$ . Барьер понижается до

$$E_b = q(V_{bi} - 2\psi_B) = kT \ln \frac{N_D}{N_A}. \quad (78)$$

Для типичных значений  $N_D = 10^{20} \text{ см}^{-3}$  и  $N_A = 10^{18} \text{ см}^{-3}$  пониженный барьер  $E_b \approx 0.1 \text{ эВ}$ . Электроны из истока могут преодолеть этот весьма небольшой барьер и уйти в инверсионный слой.

Теперь рассмотрим ситуацию на рис. 17, которая показывает зонные диаграммы когда на исток и сток подано положительное напряжение, напряжение обратного смещения (reverse bias)  $V_R$ .

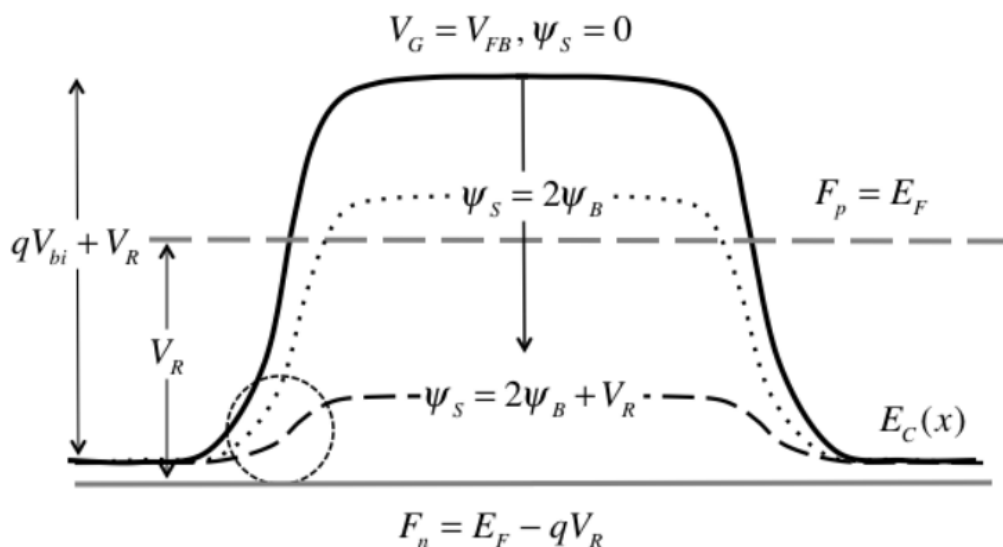


Рис. 17. Зона проводимости вдоль канала на поверхности подложки ( $y = 0$ ) от истока до стока при подаче напряжения обратного смещения  $V_R$  между  $S$  и  $D$  и подложкой  $B$ . Сплошная линия соответствует условию плоских зон в подложке.

Точечная линия соответствует поверхностному потенциалу  $\psi_S = 2\psi_B$ ,

а пунктирная линия соответствует  $\psi_S = 2\psi_B + V_R$ .

В условии плоских зон (сплошная линия) высота барьера увеличивается до  $q(V_{bi} + V_R)$ . Точечная линия соответствует зонной диаграмме для  $\psi_S = 2\psi_B$ , начало инверсии (рис. 16). В этом случае, однако, барьер все еще большой, и электроны не могут войти ни в исток, ни в сток. Получить такой же небольшой барьер, как на рис. 16, можно при поверхностном потенциале  $\psi_S = 2\psi_B + V_R$  (пунктирная линия).

Полезно показать также ход зон вглубь полупроводниковой подложки, в направлении оси  $y$  (рис. 17), а не по ходу канала проводимости (в направлении оси  $x$ ) как показано выше. Результаты для поверхностного потенциала  $\psi_S = 2\psi_B + V_R$  показаны на рис. 18.

Обратите внимание, что дырочный квази-фермиевский уровень  $F_p$  (пунктирная линия на рис. 18) расположен там, где был уровень Ферми при нулевом напряжении на истоке и стоке, а при подаче напряжения обратного смещения квази-фермиевский уровень для электронов понижается на  $qV_R$ .

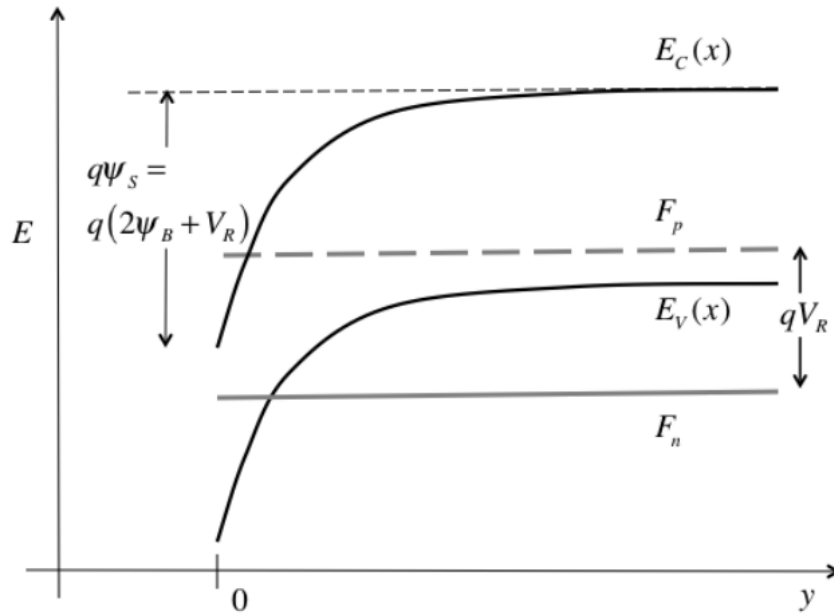


Рис. 18. Ход энергии зон вглубь полупроводниковой подложки для локации посередине канала проводимости при подаче напряжения обратного смещения  $V_R$  между  $S$  и  $D$  и подложкой  $B$ . Графики соответствуют случаю  $\psi_S = 2\psi_B + V_R$  на рис. 17. Обращает на себя внимание расщепление между квази-фермиевскими уровнями при подаче  $V_R$ : электронный уровень понижается на величину  $qV_R$ .

Электронный квази-фермиевский уровень контролирует плотность электронов в полупроводнике. Для достижения той же электронной плотности при наступлении инверсии, как и в случае при  $V_R = 0$ , зоны должны изогнуться вниз дополнительно на величину  $qV_R$ .

Для создания инверсионного слоя на истоке канала проводимости зоны должны изогнуться вниз на  $q(2\psi_B + V_R)$ . Причина роста порогового напряжения связана с увеличением заряда истощения. Из (70) следует

$$Q_D \approx -\sqrt{2qN_A \epsilon_s \psi_S} = -\sqrt{2qN_A \epsilon_s (2\psi_B + V_R)}. \quad (79)$$

Подача напряжения обратного смещения на исток существенно увеличивает заряд в слое истощения при входе в режим инверсии.

Наконец, вы можете задать вопрос, почему мы продолжаем считать полупроводник находящимся в равновесии если  $p$ - $n$ -переходы подавлены. Ответ состоит в том, что в режиме подачи напряжения обратного смещения или даже при подаче малого прямого смещения электрический ток настолько мал что естественно считать полупроводник все еще в состоянии равновесия, не допуская при этом сколь-нибудь серьезных ошибок.

Проиллюстрируем численно некоторые полученные результаты на примере кремния:

$$\begin{aligned} N_A &= 1 \times 10^{18} \text{ см}^{-3} \\ N_V &= 1.81 \times 10^{19} \text{ см}^{-3} \\ n_i &= 1.00 \times 10^{10} \text{ см}^{-3} \\ \varepsilon_s &= 11.7 \\ T &= 300 \text{ K} \end{aligned}$$

Ответим на следующие пять вопросов.

1) Где находится уровень Ферми в кремниевой подложке ( $B$ )?

Перефразируя вопрос, определяем насколько высоко над валентной зоной находится фермиевский уровень. Ответ следует из

$$p_{0B} = N_A = N_V e^{(E_V - E_F)/kT} \Rightarrow E_F - E_V = kT \ln \frac{N_V}{N_A} = 0.075 \text{ эВ}.$$

Положение уровня Ферми иначе можно определить относительно собственного уровня энергии  $E_i$ :

$$p_{0B} = N_A = n_i e^{(E_i - E_F)/kT} \Rightarrow E_i - E_F = kT \ln \frac{N_V}{n_i} = 0.48 \text{ эВ}.$$

2) Каков поверхностный потенциал  $\psi_s$  при входе в режим инверсии? Ответ:

$$\begin{aligned} \psi_s &= 2\psi_B, \\ \psi_B &= \frac{kT}{q} \ln \frac{N_A}{n_i} = 0.48 \text{ В}, \\ \psi_s &= 0.96 \text{ В}. \end{aligned} \tag{80}$$

В массивной подложке уровень Ферми расположен очень близко к валентной зоне. Чтобы сделать область под поверхностью подложки  $n$ -типа в то время как сама подложка  $p$ -типа, нужно зону проводимости изогнуть вниз как можно ближе к уровню Ферми, т. е. изогнуть вниз практически на величину энергии запрещенной зоны ( $\approx 1 \text{ эВ}$ ).

3) Какова глубина слоя истощения при входе в режим инверсии? Ответ:

$$W_T = \sqrt{4\varepsilon_s \psi_B / q N_A} = 36 \text{ нм}.$$

4) Каков полный поверхностный заряд  $Q_D$  в слое истощения? Ответ:

$$Q_D = -qN_A W_T = -\sqrt{2q\epsilon_s N_A (2\psi_B)} = -5.8 \times 10^{-6} \text{ [К/см}^2\text{]} \quad (81)$$

или через плотность заряда

$$|Q_D|/q = 3.6 \times 10^{12} \text{ [см}^{-2}\text{]}$$

5) Каково электрическое поле на поверхности подложки? Ответ:

$$E_S = -Q_D/\epsilon_s = 5.6 \times 10^6 \text{ В/см.}$$

Поле достаточно сильное.

Обсудим заряд в полупроводнике и изгиб зон. Мы выше рассматривали каким образом изгиб зон (поверхностный потенциал  $\psi_s$ ) контролирует заряд в полупроводнике. Рисунок 12 показывает изгиб зон в режимах накопления и истощения/инверсии, а рис. 19 демонстрирует, что в режиме накопления дырочный заряд мажоритарных носителей растет экспоненциально с увеличивающимся отрицательным  $\psi_s$ . В области истощения, как видно из уравнения (70), заряд  $|Q_S| \propto \sqrt{\psi_s}$ . В области инверсии электронный заряд миноритарных носителей растет экспоненциально с увеличивающимся положительным  $\psi_s > 2\psi_B$ .

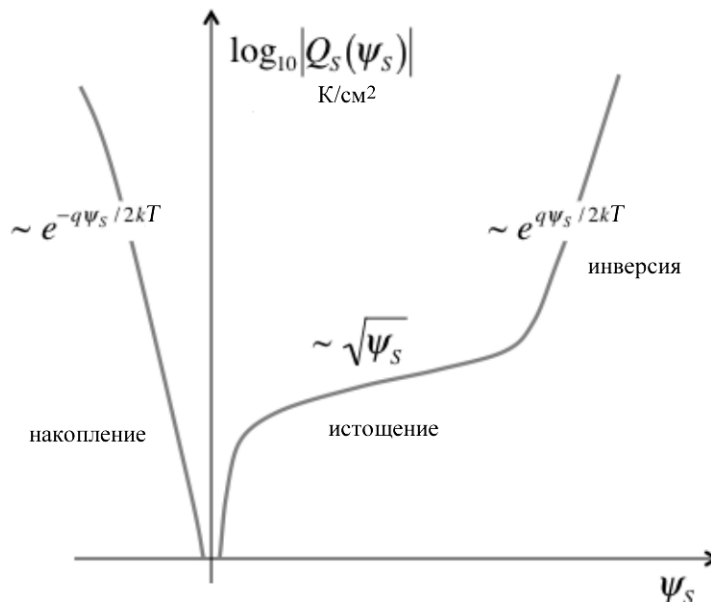


Рис. 19. Ход заряда в полупроводнике  $p$ -типа с изменением поверхностного потенциала.

Позже мы получим приближенные решения уравнения Пуассона – Больцмана для режимов накопления и инверсии, но ход  $Q_s(\psi_s)$  в целом уже не сложно представить.

Приведем критерии для слабой, умеренной и сильной инверсии. Ранее было установлено, что для режима инверсии характерно  $\psi_s > 2\psi_B$ , однако, режим инверсии устанавливается постепенно. Заметим, что при  $\psi_s = \psi_B$  полупроводник на поверхности беспримесный:  $n_0(0) = p_0(0)$ . При  $\psi_s > \psi_B$  на поверхности есть небольшой избыток электронов. Мы увидим позже, что эта небольшая избыточная концентрация электронов является причиной тока поверхностной утечки. Принято считать, что  $\psi_s = \psi_B$  есть критерий начала слабой инверсии.

При  $\psi_s = 2\psi_B$  подповерхностный слой уже  $n$ -типа, тогда как подложка –  $p$ -типа, однако, в этом подповерхностном слое концентрация электронов все еще невелика. Принято считать, что критерий  $\psi_s = 2\psi_B$  означает конец слабой инверсии и начало умеренной инверсии.

Мы позже увидим, что режим  $I_{ON}$  начинается с превышения  $2\psi_B$  всего на несколько  $kT/q$ . Когда поверхностный потенциал  $\psi_s$  становится немного больше  $2\psi_B$ , заряд  $Q_n \gg Q_D$ : заканчивается умеренная инверсия и начинается сильная инверсия. Нам в дальнейшем не потребуются более строгие критерии инверсии, однако, дизайнерам и схемотехникам при моделировании MOSFET таковые исключительно важны [8].

Подведем итоги. Была рассмотрена традиционная теория MOSFET. Возникают, конечно, некоторые вопросы, например, какова роль истощения заряда в уравнении (8), характер тока при произвольном напряжении на стоке и т. п. [6 – 9], однако, существенные черты традиционного подхода вполне приемлемы и могут служить отправной точкой при разработке во многом совершенно иной физической картины процессов в нанотранзисторах.

Мы также переписали традиционную модель MOSFET в форме, близкой к модели виртуального истока нулевого уровня. Применение этой усовершенствованной модели к современным нанотранзисторам оказалось удивительно удачным, что является следствием учета электростатики MOS через такой существенный фактор как контроль высоты барьера между истоком и каналом со стороны затвора. Наиболее слабым местом этой модели является описание транспорта электронов, поскольку оно базируется на использовании таких понятий как подвижность и скорость насыщения. Именно поэтому эти

две характеристики и были выбраны в качестве подгоночных параметров с целью вписаться в экспериментальные данные для входных характеристик.

Далее мы вернемся к более глубокому рассмотрению электростатики MOS и покажем как более корректно можно описать подпороговую и надпороговую области, результатом чего явится улучшенная модель виртуального истока, однако, подвижность и скорость насыщения все же останутся подгоночными параметрами. После этого мы физически более корректно опишем транспорт электронов и окончательно построим модель виртуального истока.

## Литература

1. S. R. Hofstein, F. P. Heiman, *Proc. IEEE*, 1190 – 1202 (1963).
2. C. T. Sah, *IEEE Trans. Electron Dev.*, **11**: 324 – 345 (1964).
3. H. Shichman, D. A. Hodges, *IEEE J. Solid State Circuits*, **SC-3**: (1968).
4. B. J. Sheu, D. L. Scharfetter, P.-K. Ko, M.-C. Jeng, *IEEE J. Solid State Circuits*, **SC-22**: 558 – 566 (1987).
5. И. М. Викулин, В. И. Стафеев, *Физика полупроводниковых приборов* (Москва: Радио и связь: 1990).
6. R. F. Pierret, *Semiconductor Device Fundamentals* (New York: Addison-Wesley: 1996).
7. B. Streetman, S. Banerjee, *Solid State Electronic Devices* (New York: Prentice Hall: 2005).
8. Y. Tsididis, C. McAndrew, *Operation and Modeling of the MOS Transistor* (New York: Oxford Univ. Press: 2011).
9. Y. Taur, T. Ning, *Fundamentals of Modern VLSI Devices* (New York: Oxford Univ. Press: 2013).
10. C. G. Sodini, P.-K. Ko, J. L. Moll, *IEEE Trans. Electron Dev.*, **31**: 1386 – 1393 (1984).
11. C. Geong, D. A. Antoniadis, M. S. Lundstrom, *IEEE Trans. Electron Dev.*, **56**: 2762 – 2769 (2009).
12. Y. Liu, M. Luisier, A. Majumdar, D. Antoniadis, M. S. Lundstrom, *IEEE Trans. Electron Dev.*, **59**: 994 – 1001 (2012).
13. A. Khakifirooz, O. M. Nayfeh, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **56**: 1674 – 1680 (2009).
14. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
15. Chenming Hu, *Modern Semiconductor Devices for Integrated Circuits* (London, UK: Pearson India: 2009).
16. M. Lundstrom, Xingshu Sun, *Notes on the Solution of the Poisson – Boltzmann Equation for MOS Capacitors and MOSFETs* (West Lafayette, Indiana: Purdue University, USA); [www.nanohub.org/resources/5338](http://www.nanohub.org/resources/5338).

# Глава 3. Напряжение на затворе и поверхностный потенциал, подвижный электронный заряд в массивной структуре MOS и в исключительно тонкой SOI

## 3.1. Введение

Рассмотрим как связаны между собой поверхностный потенциал и напряжение на затворе,  $\psi_s \propto V_G$  (рис. 1).

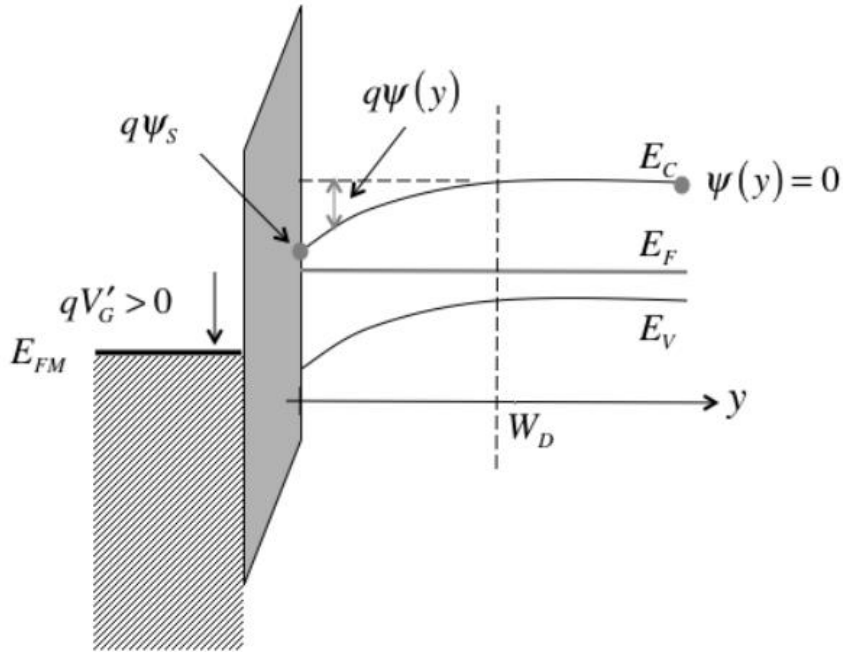


Рис. 1. Зонные диаграммы при положительном напряжении на затворе  $V'_G > 0$ , которое генерирует положительный поверхностный потенциал в полупроводнике и падение напряжения на оксидной пленке.

Сначала покажем как установить эту точную зависимость  $\psi_s(V_G)$ , а также найдем приближенное решение. Напряжение на затворе, необходимое для перевода полупроводника в режим инверсии, известно как пороговое напряжение  $V_T$ . Это то напряжение на затворе, которое необходимо для уравнивания  $\psi_s = 2\psi_B$  и которое необходимо для перевода MOSFET в режим «ON». Также обсудим емкость затвора и ее связь с напряжением на затворе, которая часто используется для характеристики структуры MOS. Затем подробно рассмотрим подвижный электронный заряд в массивной структуре MOS и в структуре SOI (Silicon-On-Insulator) с исключительной тонкой кремниевой подложкой. Более подробное обсуждение этих вопросов можно найти в [1 – 4].



### 3.2. Напряжение на затворе и поверхностный потенциал

Для установления зависимости  $\psi_S \propto V_G$  вспомним, что напряжение на затворе есть сумма падений напряжения на окисной пленке и на полупроводниковой подложке, а именно:

$$V'_G = \Delta V_{ox} + \Delta V_{semi} = \Delta V_{ox} + \psi_S. \quad (1)$$

Падение напряжения на окисной пленке есть произведение электрического поля на толщину окисной пленки,

$$\Delta V_{ox} = E_{ox} t_{ox}, \quad (2)$$

где электрическое поле определяется поверхностным зарядом полупроводника (теорема Гаусса), пренебрегая пока что возможным зарядом на границе раздела  $\text{SiO}_2/\text{Si}$ ,

$$\varepsilon_{ox} E_{ox} = -Q_S(\psi_S). \quad (3)$$

Из (2) и (3) следует, что

$$\Delta V_{ox} = -\frac{Q_S(\psi_S)}{C_{ox}}, \quad (4)$$

где

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \quad [\Phi/\text{M}^2] \quad (5)$$

есть поверхностная электрическая емкость конденсатора, каковым по существу является MOSFET (рис. 2).

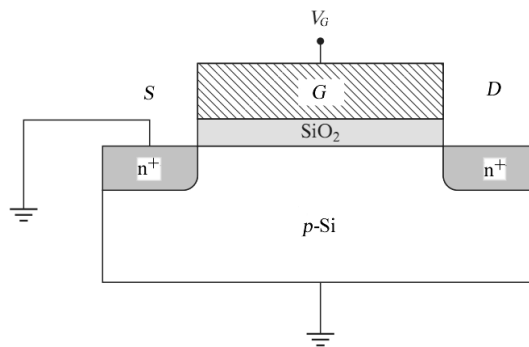


Рис. 2. MOSFET – это конденсатор с  $p\text{-n}$ -переходами.

Наконец, из (1) и (4) находим искомую связь напряжения на затворе с поверхностным потенциалом:

$$V'_G = -\frac{Q_S(\psi_S)}{C_{ox}} + \psi_S. \quad (6)$$

Уравнение (6) справедливо для идеального электрода затвора в отсутствии заряда на границе раздела  $\text{SiO}_2/\text{Si}$ , так что при  $V'_G = 0$  зоны плоские и  $\psi_s = \psi_\infty = 0$ .

Рассмотрим ситуацию, когда работа выхода электрода затвора  $\Phi_M$  меньше работы выхода полупроводника  $\Phi_S$  (рис. 3). Диаграмма равновесной зонной структуры показывает, что имеется встроенный потенциал  $\Delta V_{ox}$ . При нулевом напряжении на затворе электростатический потенциал затвора есть  $-(\Phi_M - \Phi_S)/q$ . Очевидно, что если подать на затвор напряжение, эквивалентное разности работ выхода металл/полупроводник  $\Phi_M - \Phi_S = \Phi_{MS}$ , то влияние этой разности будет нивелировано и зоны будут плоские. Соответственно этому напряжение режима плоских зон не будет при  $V_G = 0$ , а будет при  $V_G = V_{FB}$ , где

$$qV_{FB} = \Phi_M - \Phi_S = \Phi_{MS}. \quad (7)$$

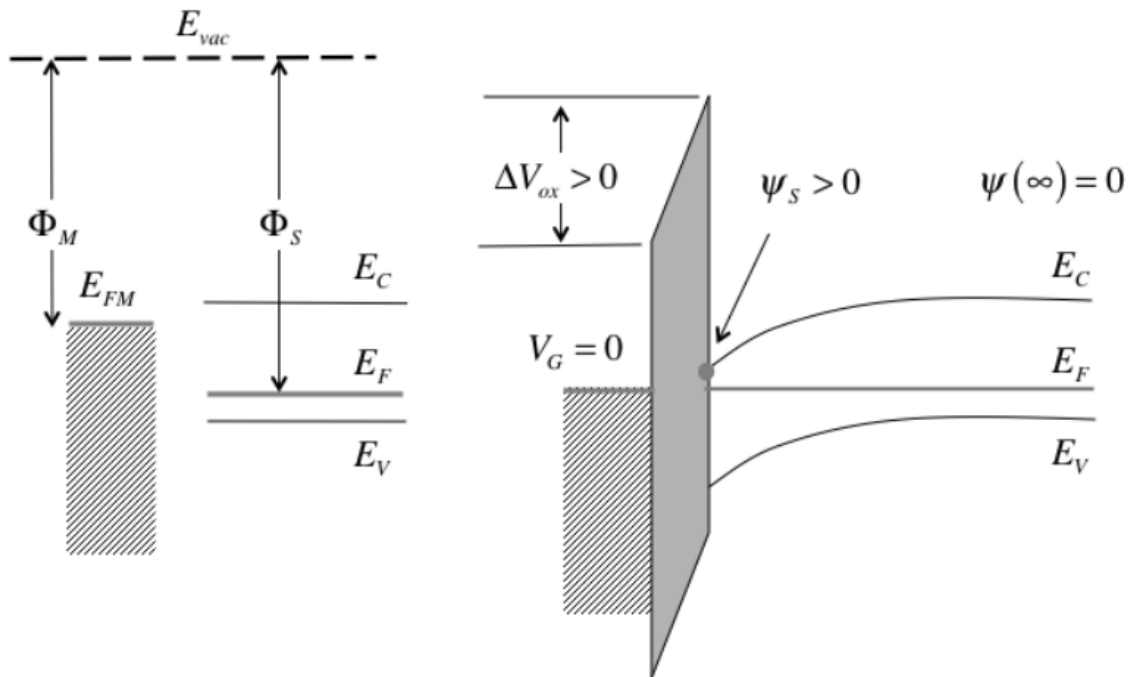


Рис. 3. Влияние разности работ выхода полупроводника  $\Phi_S$  и металла электрода затвора  $\Phi_M$  на 1D структуру MOS. Слева: изолированные друг от друга полупроводник со своим фермиевским уровнем и электрод затвора со своим уровнем Ферми  $E_{FM}$ . Справа: результирующая равновесная зонная структура при нулевом напряжении на затворе  $V_G = 0$ . Встроенный в эту структуру потенциал  $\Delta V_{ox}$  аналогичен встроенному потенциалу  $p$ - $n$ -переходов (77/гл. 2) и как и в случае  $p$ - $n$ -переходов не может быть непосредственно измерен.

Рассмотрим ситуацию подробнее и проиллюстрируем её количественно (рис. 4).

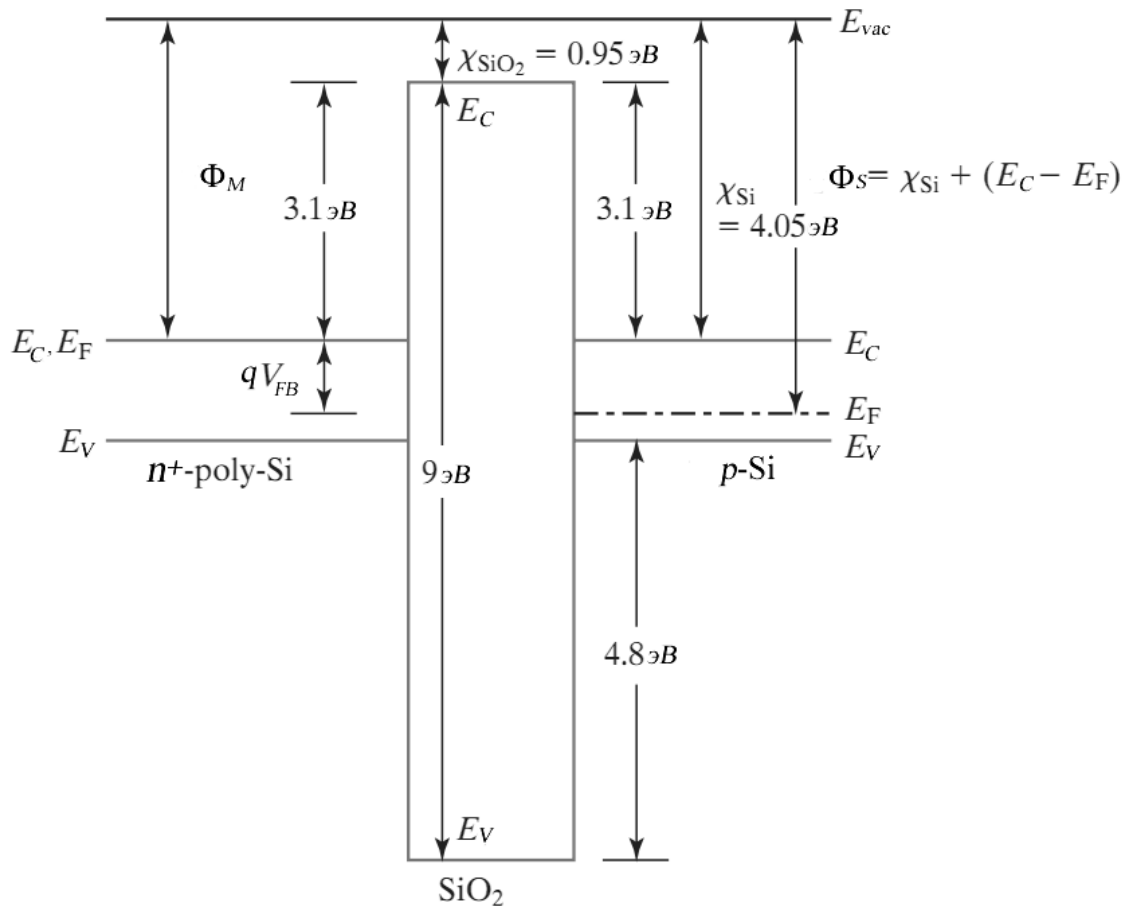


Рис. 4. Диаграмма зонной структуры MOS при выполнении условия плоских зон. Для достижения этого условия к затвору прикладывается напряжение  $V_{FB}$  между затвором  $n^+$ -poly-Si и подложкой  $p$ -Si. Показаны работы выхода затвора  $\Phi_M$  и полупроводника  $\Phi_S > \Phi_M$  [3].

При выполнении условия плоских зон на границе раздела Si/SiO<sub>2</sub> зоны плоские. Когда зоны плоские, электрическое поле на поверхности полупроводника нулевое. Поэтому электрическое поле в SiO<sub>2</sub> тоже нулевое (в отсутствии пограничного заряда):  $E_C$  и  $E_V$  в SiO<sub>2</sub> также плоские. Энергии  $E_C$  и  $E_V$  в SiO<sub>2</sub> отстоят друг от друга на  $9 \text{ eV}$ . Это ширина запрещенной зоны  $E_g$ . Разница между  $E_{vac}$  и  $E_C$  для SiO<sub>2</sub> составляет  $0.95 \text{ eV}$ . Это сродство к электрону  $\chi_{\text{SiO}_2}$ . У кремния  $\chi_{\text{Si}} = 4.05 \text{ eV}$ . Разница в  $3.1 \text{ eV}$  есть высота энергетического барьера на границе раздела Si/SiO<sub>2</sub> для электронов. Барьер для дырок составляет  $4.8 \text{ eV}$ . Из-за больших величин этих барьеров электроны и дырки не могут проникнуть через слой диэлектрика. Уровень Ферми  $E_F$  в  $n^+$ -poly-Si

обычно предполагают совпадающим с  $E_c$ . В  $\text{SiO}_2$  точное положение фермиевского уровня не существенно: если расположить его где-то посередине запрещенной зоны, то величина  $n = N_c e^{(E_c - E_F)/kT}$  будет бессмысленно малым числом, например, таким как  $10^{-60} \text{ см}^{-3}$ . Поэтому понятие положения  $E_F$  в  $\text{SiO}_2$  бессодержательное.

Напряжение плоских зон определяется разностью работ выхода затвора и подложки:

$$V_{FB} = (\Phi_M - \Phi_S)/q \approx -0.7 \text{ В}.$$

Теперь рассмотрим случай, когда между работами выхода нет разницы, а на границе раздела  $\text{Si}/\text{SiO}_2$  есть поверхностный заряд  $Q_F, \text{ К/м}^2$ . В этом случае вместо (3) имеем:

$$\varepsilon_{ox} E_{ox} = -Q_S(\psi_S) - Q_F, \quad (8)$$

а уравнение (4) будет таким:

$$\Delta V_{ox} = -\frac{Q_S(\psi_S)}{C_{ox}} - \frac{Q_F}{C_{ox}}. \quad (9)$$

Когда  $\psi_S = 0$ , заряд  $Q_S = 0$  и зоны в полупроводнике плоские. Согласно (1), это условие плоских зон реализуется при  $V_G = V_{FB} = -\frac{Q_F}{C_{ox}}$ .

В общем случае, когда имеет место разница работ выхода и одновременно есть поверхностный заряд, условие плоских зон реализуется при напряжении на затворе

$$V_{FB} = \frac{\Phi_{MS}}{q} - \frac{Q_F}{C_{ox}}, \quad (10)$$

и в качестве общего соотношения между напряжением на затворе и поверхностным потенциалом имеем

$$V'_G = V_G - V_{FB} = -\frac{Q_S(\psi_S)}{C_{ox}} + \psi_S. \quad (11)$$

Не исключено, что заряд на границе раздела  $\text{Si}/\text{SiO}_2$  не постоянный, а зависит от поверхностного потенциала и характеризуется некоторой функцией распределения по глубине оксидной пленки [1].

Уравнение (11) это наше искомое соотношение между напряжением на затворе и поверхностным потенциалом в полупроводнике. Решить его в общем

случае аналитически и найти  $\psi_s \propto V_G$  невозможно. На практике можно задаться потенциалом  $\psi_s$ , а затем вычислить напряжение  $V_G$ , которое создает такой потенциал. Мы уже видели как можно рассчитать  $Q_s(\psi_s)$  в режиме истощения. Позже мы обсудим как рассчитать эту зависимость в других режимах и рассмотрим ее подробнее.

Вычисление зависимости  $\psi_s \propto V_G$  весьма трудоемкое дело [5 – 7], однако, качественный ход этой зависимости, показанный на рис. 5, понять несложно. Обратимся к зависимости  $Q_s \propto \psi_s$  на рис. 19/гл. 2. По мере того как  $\psi_s$  растет от нуля и приобретает положительное значение заряд в слое истощения медленно возрастает как  $\sqrt{\psi_s}$ : заряд в полупроводнике умеренный, так что из (11) следует, что значительная доля напряжения на затворе падает на полупроводнике. Как только поверхностный потенциал превысил  $2\psi_B$ , становится важным инверсионный заряд; он растет экспоненциально и падение напряжения на диэлектрической прокладке становится ощутимо большим. Значительная часть напряжения на затворе падает на диэлектрической пленке, так что очень трудно поверхностный потенциал увеличить сверх  $2\psi_B$ . Когда напряжение на затворе отрицательное, быстро создается слой накопления. В этом режиме значительная доля напряжения на затворе падает на диэлектрической прослойке и намного меньшая доля – на полупроводниковой основе.

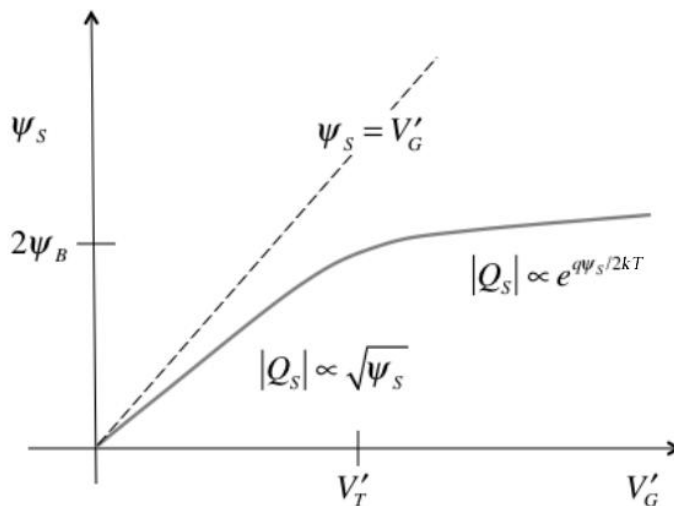


Рис. 5. Ожидаемый ход  $Q_s \propto \psi_s$ . Ниже порогового напряжения поверхностный потенциал изменяется линейно с  $V_{GS}$  согласно  $\psi_s = V_{GS}/m$ , где  $m \approx 1$ , однако, выше порогового напряжения  $\psi_s \approx 2\psi_B$  и поверхностный потенциал растет с увеличением  $V_{GS}$  медленно, поскольку  $m \gg 1$ .

## Пороговое напряжение

Одна из возможностей фиксировать пороговое напряжение  $V_T$  состоит в определении такого напряжения на затворе  $V_{GS}$ , которое необходимо для изгиба зон так, чтобы  $\psi_S$  стало равным  $2\psi_B$ , что и есть границей, после которой начинает образовываться значительный инверсионный заряд [5].

Из уравнения (11)

$$V_T = V_{FB} - \frac{Q_S(2\psi_B)}{C_{ox}} + 2\psi_B. \quad (12)$$

С началом инверсии поверхностный заряд  $Q_S = Q_D + Q$  формируется в основном из заряда истощения  $Q_D$ : заряд инверсионного слоя  $Q$  все еще небольшой. Полагая  $Q_S(2\psi_B) \approx Q_D(2\psi_B)$ , вместо (12) имеем

$$V_T = V_{FB} - \frac{Q_D(2\psi_B)}{C_{ox}} + 2\psi_B,$$

а после подстановки (80/гл. 2) окончательно имеем

$$V_T = V_{FB} + \frac{\sqrt{2q\epsilon_s N_A(2\psi_B)}}{C_{ox}} + 2\psi_B. \quad (13)$$

Это уравнение является ключевым для вычисления порогового напряжения, если мы знаем уровень допирования канала и толщину диэлектрического слоя. С увеличением плотности допирования канала пороговое напряжение растет, а с утончением диэлектрика – падает. Мы предполагаем однородное допирование канала, однако, изучаются и неоднородные профили допирования [8].

Как ранее обсуждалось в связи с уравнением (79/гл. 2), подача напряжения обратного смещения на исток понижает квази-фермиевский уровень электронов и увеличивает поверхностный потенциал при входе в режим инверсии от  $2\psi_B$  до  $2\psi_B + V_{SB}$ , где  $V_{SB}$  это напряжение обратного смещения между истоком  $S$  и подложкой  $B$ . Соответственно этому напряжение между затвором  $G$  и подложкой, необходимое для изгиба зон при входе в режим инверсии увеличивается до

$$V_{GB} = V_{FB} - \frac{Q_D(2\psi_B + V_{SB})}{C_{ox}} + 2\psi_B + V_{SB} = V_{FB} + \frac{\sqrt{2q\epsilon_s N_A(2\psi_B + V_{SB})}}{C_{ox}} + 2\psi_B + V_{SB}. \quad (14)$$

Напряжение между истоком и подложкой есть  $V_{SB}$ , так что напряжение между затвором и истоком  $V_{GS}$  при входе в режим инверсии будет  $V_{GS} = V_T$ , где

$$V_T = V_{GB} - V_{SB} = V_{FB} - \frac{Q_D(2\psi_B + V_{SB})}{C_{ox}} + 2\psi_B = V_{FB} + \frac{\sqrt{2q\epsilon_s N_A(2\psi_B + V_{SB})}}{C_{ox}} + 2\psi_B. \quad (15)$$

Мы видим, что сильно допированный канал не только увеличивает  $V_T$ , но и делает пороговое напряжение более чувствительным к напряжению  $V_{SB}$  обратного смещения между истоком  $S$  и подложкой  $B$ . Зависимость порогового напряжения от напряжения между истоком и подложкой известна как эффект подложки.

Пороговое напряжение обычно связывают с наступлением сильной инверсии. Как уже обсуждалось, при сильной инверсии  $\psi_S > 2\psi_B$ , так что  $2\psi_B$  в (15) нужно было бы заменить на потенциал, который на несколько  $kT/q$  больше  $2\psi_B$ . Тем не менее, в сложившейся практике в уравнении для  $V_T$  используют  $\psi_S = 2\psi_B$ , исключая разве что прецизионное моделирование MOSFET [1].

### Ёмкость затвора

Стандартная практика характеризовать структуру MOS состоит в измерении небольшого сигнала – переменной ёмкости между электродом затвора и дном подложки в зависимости от напряжения на затворе. Ситуация показана на рис. 6.

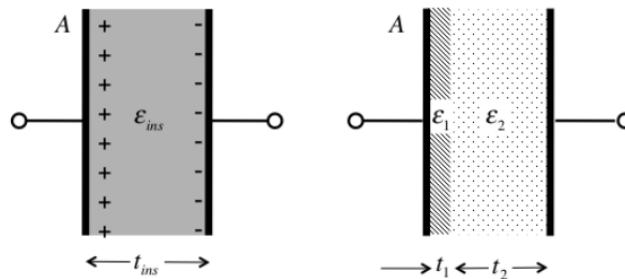


Рис. 6. Слева: обычный конденсатор с одним диэлектриком между двух параллельных обкладок с площадью поперечного сечения  $A$ .

Справа: Конденсатор с двумя различными диэлектриками.

Для простого конденсатора с двумя параллельными обкладками (рис. 6/слева) ёмкость на единицу поверхности

$$\frac{C}{A} = \frac{\epsilon_{ins}}{t_{ins}}. \quad [\Phi/M^2] \quad (16)$$

При наличии в конденсаторе двух диэлектриков с различной диэлектрической проницаемостью и разной толщины (рис. 6/справа) ёмкость  $C$  такого конденсатора связана с ёмкостями его слоев соотношением

$$\frac{1}{C/A} = \frac{1}{C_1/A} + \frac{1}{C_2/A} = \frac{1}{\epsilon_1/t_1} + \frac{1}{\epsilon_2/t_2}. \quad (17)$$

Рассмотрим ёмкость MOS при трех различных режимах: истощения, инверсии и накопления (рис. 7). В первом случае, в режиме истощения, электрод затвора есть первая «металлическая» обкладка конденсатора, диэлектрический слой затвора – первый диэлектрик, истощенный слой полупроводника – второй диэлектрик, неистощенный  $p$ -полупроводник – вторая «металлическая» обкладка конденсатора. Соответственно этому

$$\frac{1}{C_G^{dep}} = \frac{1}{C_{ox}} + \frac{1}{C_D}, \quad (18)$$

где  $C_G^{dep}$  – ёмкость затвора в режиме истощения,

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} \quad (19)$$

есть ёмкость диэлектрического оксидного слоя и

$$C_D = \frac{\epsilon_s}{W_D(\psi_s)} \quad (20)$$

есть ёмкость слоя истощения в полупроводниковой подложке. Все ёмкости имеют размерность  $[\Phi/M^2]$ .

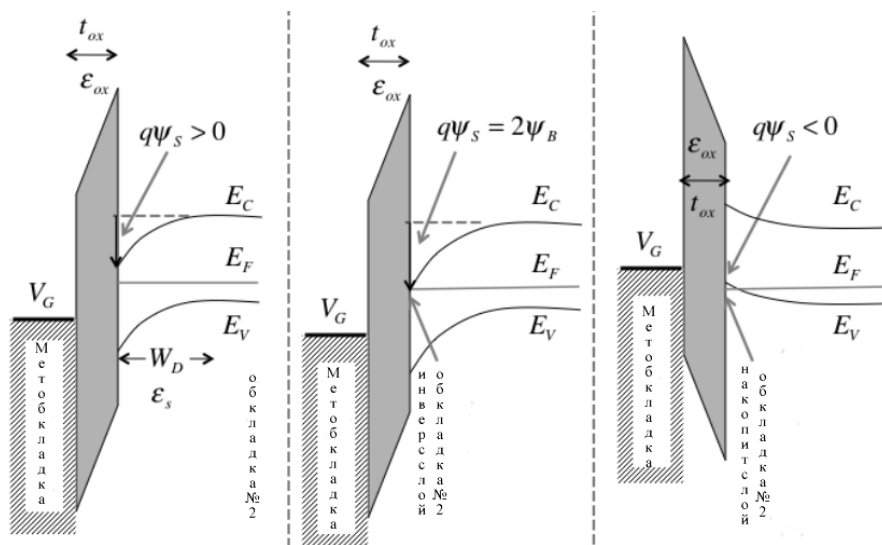


Рис. 7. Диаграммы зонной структуры при трех различных напряжениях. Режимы: истощения (слева), инверсии (по центру), накопления (справа).



Рассмотрим инверсионную ёмкость (рис. 7/по центру). В этом случае первым диэлектриком, как и в случае режима истощения, является диэлектрический слой затвора, однако, второй «металлической» обкладкой служит высокопроводящий инверсионный слой электронов на границе раздела Si/SiO<sub>2</sub>. Поэтому ёмкость затвора в режиме инверсии

$$C_G^{inv} \approx C_{ox}. \quad (21)$$

Рассмотрим, наконец, накопительную ёмкость (рис. 7/справа). В этом случае первым диэлектриком, как и прежде, является диэлектрический слой затвора, а второй «металлической» обкладкой служит высокопроводящий накопительный слой дырок на границе раздела Si/SiO<sub>2</sub>. Соответственно этому ёмкость затвора в режиме накопления

$$C_G^{acc} \approx C_{ox}. \quad (22)$$

Итак, ёмкость затвора есть последовательное соединение двух конденсаторов, а именно:

$$\frac{1}{C_G} = \frac{1}{C_{ox}} + \frac{1}{C_S(\psi_S)}, \quad (23)$$

где  $C_S(\psi_S)$  есть ёмкость полупроводника, которая сильно зависит от поверхностного потенциала.

Этим качественным рассуждениям о зависимости ёмкости затвора от напряжения на затворе можно придать количественный вид. Ёмкость затвора определяется как

$$C_G \equiv \frac{dQ_G}{dV_G}, \quad (24)$$

где  $Q_G = -Q_S$  есть поверхностный заряд электрода затвора, равный поверхностному заряду полупроводника, как того требует баланс зарядов. Дифференцируя (11), находим

$$\frac{dQ_G}{d(-Q_S)} = \frac{d\psi_S}{d(-Q_S)} + \frac{1}{C_{ox}}, \quad (25)$$

что можно переписать в виде

$$\frac{1}{C_G} = \frac{1}{C_{ox}} + \frac{1}{C_S}, \quad (26)$$

где

$$C_s \equiv \frac{d(-Q_s)}{d\psi_s} \quad (27)$$

есть емкость полупроводника. Обратите внимание, что увеличение поверхностного потенциала увеличивает отрицательный заряд в слоях истощения и инверсии, так что емкость полупроводника положительная величина.

Эквивалентная схема ёмкости затвора показана на рис. 8. Для вычисления  $C_G \propto V_G$  нужно знать как  $C_s = d(-Q_s)/d\psi_s$  зависит от  $V_{GS}$ .

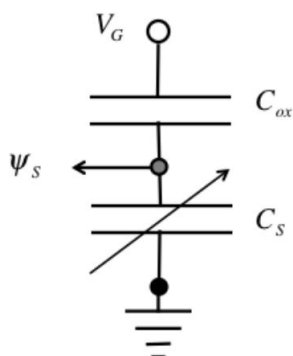


Рис. 8. Эквивалентная схема показывает, что ёмкость затвора есть последовательное соединение ёмкости оксидного слоя  $C_{ox}$  и ёмкости полупроводниковой подложки  $C_s$ .

Качественный ответ на этот вопрос представлен на рис. 19/гл. 2. В режимах накопления и инверсии ёмкость полупроводника очень высокая, так что суммарная ёмкость близка к ёмкости оксидного слоя. В режиме же истощения ёмкость полупроводника умеренная, так что суммарная ёмкость уменьшается (рис. 9).

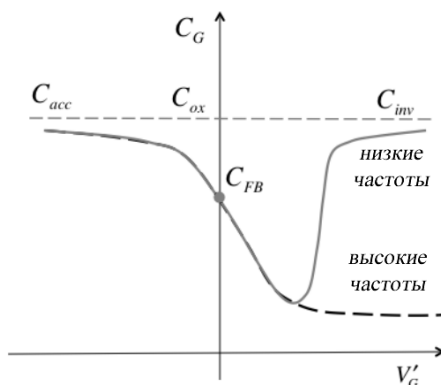


Рис. 9. Качественный ход ёмкости затвора с напряжением на затворе. Сплошная линия соответствует низким частотам, а пунктирная – высоким.

*Приближенная зависимость напряжения на затворе от поверхностного потенциала*

Уравнение (11) связывает напряжение на затворе с поверхностным потенциалом полупроводниковой подложки. В общем случае решать уравнение приходится численно, а в режиме истощения ( $Q_S \approx Q_D$ ) его можно решить аналитически. В предположении равенства заряда полупроводника заряду слоя истощения ( $Q_S = Q_D$ ) и учитывая (80/гл. 2) при  $2\psi_B = \psi_S$  вместо (11) имеем квадратичное по  $\sqrt{\psi_S}$  уравнение

$$V_G = V_{FB} + \frac{\sqrt{2q\epsilon_s N_A (\psi_S)}}{C_{ox}} + \psi_S, \quad (28)$$

решение которого можно найти в [1]. Для многих приложений желательно иметь более простое соотношение  $V_G \propto \psi_S$ , которое подсказывает нам эквивалентная схема на рис. 8.

Ёмкость полупроводника зависит от поверхностного потенциала, однако, в режиме истощения ёмкость полупроводника есть ёмкость истощения, которая с учетом (68/гл. 2) меняется весьма медленно с поверхностным потенциалом  $\psi_S$ , а именно:

$$C_S \approx C_D = \frac{\epsilon_s}{W_D(\psi_S)} = \frac{\epsilon_s}{\sqrt{2\epsilon_s \psi_S / (qN_A)}}. \quad (29)$$

Если аппроксимировать ёмкость истощения ее средним значением в режиме истощения, тогда вместо рис. 8 имеем просто два постоянных конденсатора, соединенных последовательно, что дает

$$\psi_S = V_G \left( \frac{C_{ox}}{C_{ox} + C_D} \right) = \frac{V_G}{m}, \quad (30)$$

где

$$m = 1 + \frac{C_D}{C_{ox}} \quad (31)$$

известен как коэффициент эффекта подложки в режиме истощения.

Этот коэффициент  $m$  говорит нам, какая доля приложенного напряжения на затворе падает на полупроводниковой подложке. Для очень тонких диэлектрических пленок  $C_{ox} \gg C_D$  и коэффициент  $m \rightarrow 1$ : все приложенное напряжение на затворе падает на полупроводниковой подложке. Так

происходит, поскольку лишь незначительное падение напряжения возможно на тонкой оксидной пленке. Для слабодопированного полупроводника  $C_D \ll C_{ox}$  и коэффициент  $m \rightarrow 1$ : опять все приложенное напряжение на затворе падает на полупроводниковой подложке. Так происходит, поскольку слабое допирование порождает лишь небольшой заряд в полупроводнике, отсюда слабое электрическое поле в диэлектрической пленке и соответственно незначительное падение напряжения на ней. Для  $m$  типичные значения 1.1 – 1.3, так что наклон графика  $\psi_s \propto V_G$  в режиме истощения меньше единицы (рис. 5).

Проиллюстрируем численно некоторые полученные результаты на уже ранее рассмотренном примере кремния (гл. 2). Дополнительно примем  $t_{ox} = 1.8 \text{ нм}$ ,  $\epsilon_{ox} = 4.0$ , затвор  $n^+$ -poly-Si и отсутствие заряда на границе Si/SiO<sub>2</sub>.

Ответим на следующие три вопроса.

1) Какова разница работ выхода металла затвора и полупроводника и каково напряжение плоских зон?

Ранее мы уже определили положение уровня Ферми в полупроводнике выше на 0.075 эВ потолка валентной зоны (гл.2). В случае сильнодопированного затвора  $n^+$ -poly-Si разумно предположить, что  $E_F = E_C$  (рис. 4). Разница между уровнями Ферми в металлоподобном затворе  $n^+$ -poly-Si и в полупроводнике  $p$ -типа на 0.075 эВ меньше величины запрещенной зоны Si:

$$\Phi_{MS} = -(1.1 - 0.075) = -1.03 \text{ эВ},$$

так что напряжение плоских зон  $V_{FB} = \Phi_{MS}/q = -1.03 \text{ В}$ .

2) Каково пороговое напряжение?

Ранее в численном примере с кремнием (гл. 2) нашли, что

$$\psi_B = \frac{kT}{q} \ln \frac{N_A}{n_i} = 0.48 \text{ В},$$

так что при входе в режим инверсии  $\psi_s = 2\psi_B = 0.96 \text{ В}$ . В этом режиме заряд полупроводника есть в основном заряд слоя истощения  $Q_D$ , который выражается через  $2\psi_B$ . Емкость оксидной пленки по (19)

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = 2.0 \times 10^{-6} \text{ Ф/см}^2. \quad (32)$$

Окончательно из (13) находим пороговое напряжение  $V_T = 0.19 \text{ В}$ .

3) Чему равен коэффициент эффекта подложки в режиме истощения?

Сначала определим ёмкость слоя истощения. Оценим ее по (29) при  $\psi_s = \psi_B = 0.48 V$ :

$$C_D = \frac{\epsilon_s}{W_D} = \frac{\epsilon_s}{\sqrt{2\epsilon_s\psi_s/(qN_A)}} = 8.3 \times 10^{-7}. \quad [\Phi/cm^2]$$

Наконец, по (31) находим  $m = 1.4$ : напряжение на затворе в режиме истощения падает на кремниевой подложке на 70%.

На рис. 10 показан типичный стек затвора. Роль электрода затвора играет не металл, а сильно допированный поликристаллический кремний  $n^+$ -poly-Si. Если достаточно сильно допирован,  $n^+$ -poly-Si ведет себя как металл.

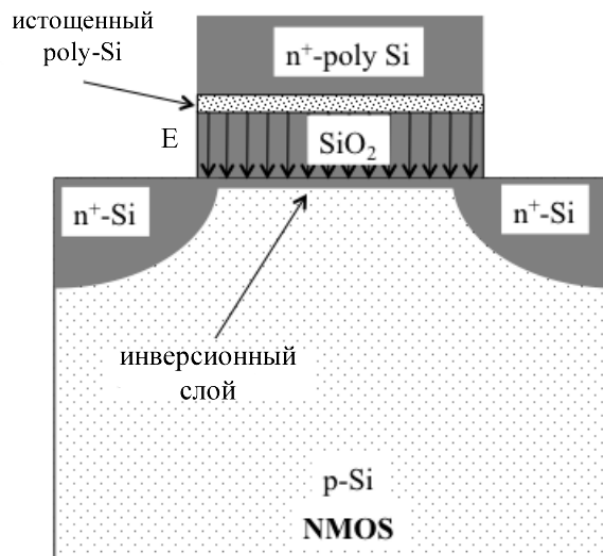


Рис. 10. Типичный стек затвора с участием сильно допированного  $n^+$ -poly-Si. NMOS  $\equiv n$ -канальный MOSFET.

Как показано на рис. 8, ёмкость затвора определяется последовательным соединением ёмкости оксидного слоя  $C_{ox}$  и ёмкости полупроводниковой подложки  $C_s$ , так что суммарная ёмкость затвора меньше чем  $C_{ox}$ . В режиме истощения суммарная ёмкость существенно меньше чем  $C_{ox}$ , однако, в режиме инверсии ёмкость полупроводника становится очень большой. В идеале мы стремимся к существенному превышению  $C_s$  над  $C_{ox}$  в режиме инверсии с тем, чтобы  $C_G \approx C_{ox}$ . По мере того как оксидная пленка все утончалась и утончалась в течение последних нескольких десятилетий, понижение ёмкости затвора в режиме инверсии за счет ёмкости полупроводника стало играть важную роль. Мы позже еще обсудим эту ситуацию.

В случае использования в затворе  $n^+$ -poly-Si появляется еще один фактор, который понижает суммарную ёмкость затвора: это так называемое poly-истощение. Как показано на рис. 10, в режиме инверсии возникает сильное электрическое поле в направлении  $+y$  вглубь подложки от положительно заряженного затвора в сторону отрицательного заряда в полупроводнике. Это электрическое поле истощается, а затем инвертирует в полупроводниковой подложке. Но это электрическое поле также немного истощает сильно допированный затвор  $n^+$ -poly-Si. Суммарная ёмкость теперь определяется тремя конденсаторами, соединенными последовательно, – оксидная ёмкость, ёмкость подложки и ёмкость истощения  $n^+$ -poly-Si:

$$\frac{1}{C_G} = \frac{1}{C_{poly}} + \frac{1}{C_{ox}} + \frac{1}{C_S}.$$

Инженеры-дизайнеры MOSFET описывают эти эффекты в терминах толщины эквивалентной ёмкости (capacitance equivalent thickness/CET), которая определяется как толщина  $\text{SiO}_2$ , которая обеспечивает измеряемую ёмкость затвора в условиях сильной инверсии, включая эффекты от всех трех факторов. Величина CET определяется как

$$C_G \equiv \frac{\epsilon_{ox}}{CET}. \quad (33)$$

В приведенном выше примере толщину  $t_{ox} = 1.8 \text{ нм}$  было бы правильнее обозначить как  $CET = 1.8 \text{ нм}$ .

### 3.3. Подвижный электронный заряд: массивная структура MOS

Выше обсуждалось поведение заряда в полупроводнике с изгибом зон, определяемым поверхностным потенциалом. Задача состояла в том, чтобы качественно понять зависимость  $Q_s(\psi_s)$ . Было также показано как решить уравнение Пуассона в режиме истощения и получить зависимость  $Q_D(\psi_s)$ . Далее мы сосредоточимся на изучении инверсионного заряда, заряда подвижных электронов  $Q_n(\psi_s)$  в  $n$ -MOSFET. В  $p$ -MOSFET соответствующая величина есть  $Q_p(\psi_s)$ , заряд подвижных дырок в дырочном инверсионном слое подложки  $n$ -типа. Нужно только не забывать, что носителями тока в этой ситуации также являются реальные электроны валентной зоны, а не умозрительные, но удобные при обсуждении «дырки». Как и ранее, подвижный заряд электронов будем обозначать просто  $Q$ , опуская нижний индекс  $n$  за

ненадобностью. Вычисление инверсионного заряда в зависимости от поверхностного потенциала сводится к решению уравнения Пуассона – Больцмана (59/гл. 2) [5 – 7].

О решении интересующего нас уравнения Пуассона – Больцмана часто говорят как о «полном и точном» решении проблемы MOS, но это не совсем так. Например, это уравнение предполагает справедливой статистику Максвелла – Больцмана, тогда как в режимах сильной инверсии и накопления правильнее было бы использовать статистику Ферми – Дирака. Уравнение Пуассона – Больцмана не учитывает квантовые ограничения, обусловленные потенциальной ямой, образующейся на границе раздела Si/SiO<sub>2</sub> в структуре MOS. Учет квантовых ограничений является особенно существенным в современных наноразмерных структурах MOS. И тем не менее, решения уравнений Пуассона – Больцмана не только широко распространены, но и служат вполне приемлемым приближенным подходом при моделировании MOSFET. Далее будут изложены приближенные аналитические решения уравнений Пуассона – Больцмана для массивной структуры MOS в режимах слабой и сильной инверсии. Под термином «массивная структура MOS» подразумевается, что полупроводниковая подложка начинается с  $y=0$  и условно заканчивается на бесконечности. На практике кремниевую пластину для наших целей можно рассматривать как бесконечно толстую. Мы также найдем приближенные решения для зависимостей  $Q(V_G)$ .

### *Подвижный электронный заряд*

По определению, подвижный электронный заряд

$$Q = -q \int_0^{\infty} n_0(y) dy = -qn_s. \quad [K/m^2]. \quad (34)$$

Поскольку электронная плотность зависит экспоненциально от разности между дном зоны проводимости  $E_c$  и уровнем Ферми (рис. 9/гл. 2), она увеличивается вблизи поверхности, где увеличивается и электростатический потенциал, а зона проводимости  $E_c$  изгибается вниз (рис. 10/гл. 2). Рассматривая структуру, аналогичную представленной на рис. 8/гл. 2 с  $V_s = V_d = 0$ , имеем

$$n_0(y) = \left( \frac{n_i^2}{N_A} \right) e^{q\psi(y)/kT}. \quad (35)$$

Подставив (35) в (34) и выполнив замену переменных, имеем

$$Q = -q \left( \frac{n_i^2}{N_A} \right) \int_0^{\infty} e^{q\psi(y)/kT} dy = -q \left( \frac{n_i^2}{N_A} \right) \int_{\psi_S}^0 e^{q\psi(y)/kT} \frac{dy}{d\psi} d\psi. \quad (36)$$

В общем случае требуется численное моделирование зависимости  $\psi(y)$  с последующим взятием интеграла в уравнении (36), однако, поскольку электроны концентрируются очень близко к поверхности раздела Si/SiO<sub>2</sub>, правдоподобно предположить, что электрическое поле  $E = -d\psi/dy$  приблизительно постоянно в пределах существенной части области интегрирования вблизи поверхности раздела. Пусть среднее значение электрического поля в этом электронном слое есть  $E_{ave}$ . Тогда (36) можно переписать в виде

$$Q \approx -q \left( \frac{n_i^2}{N_A} \right) \frac{1}{E_{ave}} \int_{\psi_S}^0 e^{q\psi(y)/kT} d\psi, \quad (37)$$

а после взятия интеграла получим

$$Q(\psi_S) = -q \left[ \left( \frac{n_i^2}{N_A} \right) e^{q\psi_S/kT} \right] \left( \frac{kT/q}{E_{ave}} \right). \quad (38)$$

Выражение в квадратных скобках, согласно (35), есть электронная плотность на поверхности раздела Si/SiO<sub>2</sub> ( $y = 0$ ), а после введения понятия толщины электронного слоя  $t_{inv}$ , перепишем (38) в окончательном виде

$$\begin{aligned} Q(\psi_S) &= -qn(0)t_{inv}, \\ n(0) &= \frac{n_i^2}{N_A} e^{q\psi_S/kT}, \\ t_{inv} &= \left( \frac{kT/q}{E_{ave}} \right). \end{aligned} \quad (39)$$

Согласно (39), поверхностный электронный заряд  $Q$  есть концентрация электронов на поверхности раздела, помноженная на заряд электрона и эффективную толщину электронного слоя. Это уравнение применимо и ниже и выше порога напряжения. Начнем с рассмотрения подпорогового случая.



### Подвижный заряд ниже порогового напряжения

Уравнение (39) для  $Q(\psi_S)$  ниже порогового напряжения работает, когда для определения  $E_{ave}$  можно воспользоваться приближением истощения. Поскольку электронный слой тонок сравнительно с толщиной слоя истощения, можно предположить, что  $E_{ave} \approx E_S$ . Уравнение (39) электронный заряд выражает через поверхностный потенциал  $\psi_S$ , однако, в рассматриваемом случае заряд  $Q$  удобнее выразить через напряжение на затворе  $V_G$  и коэффициент эффекта подложки  $m$ , а для этого уравнение (38) для  $Q(\psi_S)$  нужно переписать в уравнение для  $Q(V_G)$ .

Имеем

$$E_{ave} \approx E_S = \frac{qN_A W_D}{\epsilon_s} = \frac{qN_A}{C_D}, \quad (40)$$

где  $W_D$  есть толщина слоя истощения, а  $C_D = \epsilon_s/W_D$  есть ёмкость этого слоя. Далее, согласно (31), ёмкость слоя истощения связана с коэффициентом эффекта подложки  $m = 1 + C_D/C_{ox}$ , так что  $C_D = (m-1)C_{ox}$  и вместо (40) имеем

$$E_S = \frac{qN_A}{(m-1)C_{ox}}. \quad (41)$$

Теперь (38) можно переписать иначе, а именно:

$$Q(\psi_S) = -(m-1)C_{ox} \frac{kT}{q} \left( \frac{n_i}{N_A} \right)^2 e^{q\psi_S/kT}. \quad (42)$$

Согласно (72/гл. 2),  $n_i/N_A$  связано с  $\psi_B$  следующим образом:

$$\left( \frac{n_i}{N_A} \right)^2 = e^{-q2\psi_B/kT}, \quad (43)$$

так что вместо (42) имеем:

$$Q(\psi_S) = -(m-1)C_{ox} \frac{kT}{q} e^{q(\psi_S - 2\psi_B)/kT}. \quad (44)$$

Заменить поверхностный потенциал в (44) на напряжение на затворе позволяет уравнение (30). Окончательно имеем:

$$Q(V_G) = -(m-1)C_{ox} \frac{kT}{q} e^{q(V_G - V_T)/mkT}. \quad (45)$$

Это уравнение дает зависимость небольшого подпорогового подвижного заряда от напряжения на затворе. Ниже порогового напряжения этот небольшой заряд должен расти экспоненциально с увеличением  $V_G$ , что не видно на иллюстративном рис. 11 из-за линейности вертикальной шкалы для заряда. Этот рост связан с тем, что по мере того как зоны изгибаются вниз с увеличением напряжения на затворе концентрация электронов растет экспоненциально. Этот экспоненциальный рост заряда с ростом  $V_G$  ниже порогового напряжения является причиной экспоненциально растущего подпорогового тока.

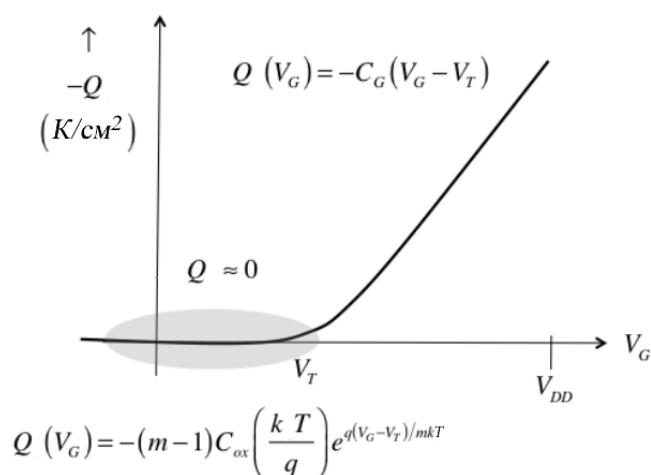


Рис. 11. Ход электронного заряда  $Q$  с ростом напряжения на затворе  $V_G$  для  $n$ -MOSFET. Линейность вертикальной шкалы для заряда не позволяет отразить экспоненциальный рост  $Q \propto V_G$  ниже  $V_G = V_T$ , однако, линейная связь  $Q \propto V_T$  видна при  $V_G > V_T$ .

Нужно отметить, что пороговое напряжение в (45) на самом деле зависит от напряжения на затворе [5]. При входе в режим инверсии пороговое напряжение  $V_T$  дается уравнением (12), однако, зоны не полностью прекращают изгибаться когда  $\psi_s = 2\psi_B$ , они продолжают изгибаться медленно с ростом  $V_G$ , что заставляет  $V_T$  увеличиться при сильной инверсии [5].

Множитель  $(m-1)$  в уравнении (45) также заслуживает внимания. В качественных транзисторах  $m \rightarrow 1$  и множитель  $(m-1)$  означал бы стремление заряда к нулю, что явно нефизично и обусловлено упрощениями при получении формулы (30). Эта формула используется часто, но она справедлива, если  $m \sim 1.1$  [5].

### Подвижный заряд выше порогового напряжения

Уравнение (39) применимо и ниже и выше порога напряжения. Ниже порога напряжения мы использовали для  $E_s$  приближение истощения. При сильной инверсии  $Q_s \approx Q \gg Q_D$ . Вместо уравнения (40) закон Гаусса теперь дает

$$E_s = -\frac{Q}{\varepsilon_s}. \quad (46)$$

Электрическое поле меняется быстро в инверсионном слое от значения  $E_s$  на поверхности до почти нуля на дне инверсионного слоя. Соответственно этому, предположим, что  $E_{ave} \approx E_s/2$ . Тогда (39) с учетом (46) переписывается в виде

$$Q = -\sqrt{2\varepsilon_s kT n(0)} \quad (47)$$

или после подстановки выражения для  $n(0)$

$$Q(\psi_s) = -\sqrt{2\varepsilon_s kT (n_i^2/N_A)} e^{q\psi_s/2kT}. \quad (48)$$

Это уравнение показывает, что в режиме сильной инверсии электронный заряд ведет себя как  $Q \propto e^{q\psi_s/2kT}$ , что ранее уже было отмечено на рис. 19/гл. 2. Аналогичные рассуждения приводят к тому, что в режиме накопления дырочный заряд  $Q_p \propto e^{-q\psi_s/2kT}$ .

Уравнение (48) показывает зависимость электронного заряда от поверхностного потенциала  $\psi_s$ , а мы хотим получить зависимость заряда от напряжения на затворе  $V_G$ . Зависимость  $Q \propto V_G$  можно получить путем решения уравнения (11) численно с использованием (48), что приводит к приблизительно линейной зависимости для  $V_G > V_T$ , как это показано на рис. 11:  $Q \propto (V_G - V_T)$  для  $V_G > V_T$ .

Чтобы удостовериться, что линейная зависимость  $Q \propto V_G$  действительно выполняется выше порогового напряжения, рассмотрим уравнение (11). При входе в режим инверсии большая доля заряда полупроводниковой подложки есть заряд в слое истощения и  $\psi_s = 2\psi_B$ . Вследствие этого из (11) находим, что

$$V_T = V_{FB} - \frac{Q_D(2\psi_B)}{C_{ox}} + 2\psi_B, \quad (49)$$

обозначив напряжение на затворе при входе в режим инверсии как пороговое напряжение  $V_T$ . Для напряжения на затворе, значительно большего  $V_T$ , изгиб зон и заряд истощения изменяются незначительно, однако, формируется большой инверсионный заряд. Из (11) следует, что в этой ситуации

$$V_G \approx V_{FB} - \frac{Q_D(2\psi_B) + Q}{C_{ox}} + 2\psi_B. \quad (50)$$

Вычитая (49) из (50), видим, что

$$Q \approx -C_{ox}(V_G - V_T). \quad (51)$$

В практических расчетах оказывается, что  $d(-Q)/dV_G$  немного меньше  $C_{ox}$ , поскольку  $\psi_s$  не точно уравнивается с  $2\psi_B$ , как это предполагается в уравнении (50). Наклон можно найти из

$$\frac{d(-Q)}{dV_G} \approx \frac{d(-Q_s)}{dV_G} = \frac{d(Q_M)}{dV_G} = C_G, \quad (52)$$

так что выше порогового напряжения инверсионный заряд записываем как

$$Q(V_G) \approx -C_G(V_G - V_T), \quad (53)$$

где  $C_G (< C_{ox}) \approx \text{Const}$ . Ранее, уравнение (23), мы видели, что

$$\frac{1}{C_G} = \frac{1}{C_{ox}} + \frac{1}{C_S}, \quad (54)$$

где  $C_S$  есть ёмкость полупроводника: ёмкость в режиме истощения или ёмкость инверсионного слоя в режиме инверсии,

$$C_S^{inv} = \frac{d(-Q)}{d\psi_s} = \frac{-Q}{2kT/q}, \quad (55)$$

где последнее значение следует из (48). По другому, ёмкость полупроводника в режиме инверсии

$$C_S^{inv} = \frac{\epsilon_s}{t_{inv}}, \quad (56)$$

где толщина инверсионного слоя

$$t_{inv} = \frac{2kT\epsilon_s}{-qQ}. \quad (57)$$

Итак, в режиме сильной инверсии, когда напряжение на затворе сильно превышает пороговое напряжение, электронный заряд инверсионного слоя дается выражениями

$$\begin{aligned}
 Q(V_G) &= -C_G(V_G - V_T), \quad (V_G > V_T) \\
 \frac{1}{C_G} &= \frac{1}{C_{ox}} + \frac{1}{C_S^{inv}}, \\
 C_S^{inv} &= \frac{\varepsilon_s}{t_{inv}}, \\
 t_{inv} &= \frac{2kT\varepsilon_s}{-qQ}.
 \end{aligned}
 \tag{58}$$

Эти результаты свидетельствуют о том, что когда  $C_S \gg C_{ox}$ , то  $C_G \approx C_{ox}$ . Этот вывод доминировал в технологии MOS в течение длительного времени. Однако, по мере все большего утончения диэлектрического слоя в затворе, допущения, приведшие к этому выводу, все больше нарушались. Кроме того, эффекты, связанные со статистикой Ферми – Дирака и с квантовыми ограничениями, до сих пор нами пренебрегаемые, как выяснилось, также понижают ёмкость полупроводниковой подложки  $C_S$ . В результате оказалось, что  $C_S$  существенно понизило ёмкость затвора  $C_G$  ниже  $C_{ox}$  в современных MOSFET.

Приведем численные примеры по ёмкости инверсионного слоя и его толщине на все том же примере кремния с уже приведенными выше исходными данными.

Ответим на следующие четыре вопроса.

1) Чему равна ёмкость полупроводниковой подложки при  $n_s = 1 \times 10^{13} \text{ см}^{-2}$  ?

Такая плотность слоя электронов  $n_s$  типична для современных MOSFET в режиме «ON». Из (55) находим

$$C_S^{inv} = \frac{-Q}{2kT/q} = \frac{qn_s}{2kT/q} = 30.8 \times 10^{-6} \text{ Ф/см}^2.
 \tag{59}$$

Это очень большая ёмкость по сравнению с ёмкостью оксидного слоя  $C_{ox} = 2.0 \times 10^{-6} \text{ Ф/см}^2$ , найденной ранее (32), просто нереально большая. Как будет показано позже, переход к статистике Ферми – Дирака и учет квантовых ограничений существенно понизят величину  $C_S$ .

2) Чему равна ёмкость затвора?

Из уравнения (54) находим

$$C_G = \frac{C_{ox}C_S}{C_{ox} + C_S} = \frac{C_{ox}}{1 + C_{ox}/C_S} = \frac{C_{ox}}{1 + 2.0/30.8} = 0.94C_{ox} = 1.9 \times 10^{-6} \text{ Ф/см}^2. \quad (60)$$

Как и ожидалось, получили  $C_G < C_{ox}$ . Когда перейдем к статистике Ферми – Дирака и учтем квантовые ограничения, отношение  $C_G/C_{ox}$  станет еще меньше.

3) Чему равна толщина эквивалентной ёмкости  $CET$  ?

Как следует из (33),

$$CET = \frac{\varepsilon_{ox}}{C_G} = \frac{4.0 \times 8.854 \times 10^{-14}}{1.9 \times 10^{-6}} = 1.86 \text{ нм}, \quad (61)$$

где учтена электрическая постоянная  $\varepsilon_0 = 8.854 \times 10^{-12} \text{ Ф/м}$ . Толщина эквивалентной ёмкости  $CET$  несколько больше действительной толщины оксидной пленки, заданной равной  $t_{ox} = 1.8 \text{ нм}$ . Увеличение  $CET$  окажется еще бóльшим после перехода к статистике Ферми – Дирака и учета квантовых ограничений. Использование затвора  $n^+$ -poly-Si с poly-истощением (рис. 10) также увеличивает  $CET$ .

4) Каков поверхностный потенциал кремниевой подложки при приведенной выше плотности слоя электронов  $n_s$  ?

Из формулы (48) с заменой  $Q = -qn_s$  имеем

$$\psi_s = 2 \frac{kT}{q} \ln \left( \frac{qn_s}{\sqrt{\varepsilon_s kT (n_i^2/N_A)}} \right) = 1.12 \text{ В}. \quad (62)$$

Поверхностный потенциал подложки получился немного больше чем  $\psi_s = 2\psi_B = 0.96 \text{ В}$  (80/гл. 2) в режиме сильной инверсии, больше приблизительно на  $6kT/q$ . Этот эффект будет еще больше при переходе к статистике Ферми – Дирака и учете квантовых ограничений.

### *Еще о поверхностном потенциале и напряжении на затворе*

Часто говорят, что изгиб зон в структуре MOS ограничен условием  $\psi_s \approx 2\psi_B$ . Мы только что видели на конкретном примере, что поверхностный потенциал в режиме сильной инверсии на несколько  $kT/q$  больше чем  $2\psi_B$ . И все же изгиб зон трудно сильно увеличить при превышении  $2\psi_B$ . Приведем пример того, почему так происходит.

Согласно (48), электронный заряд зависит экспоненциально от поверхностного потенциала в режиме сильной инверсии. Предположим, что напряжение на затворе ведет к изгибу зон с  $n_s = 5 \times 10^{12} \text{ см}^{-2}$ . Зададимся вопросом: насколько нужно увеличить изгиб зон, чтобы удвоить плотность инверсионного слоя до  $n_s = 1 \times 10^{13} \text{ см}^{-2}$ ? Из (48) / (62) ответ будет таким:

$$\Delta\psi_s = 2 \frac{kT}{q} \ln 2 = 0.036 \text{ В},$$

т. е. очень небольшое увеличение поверхностного потенциала удваивает электронную плотность инверсионного слоя, или иначе, инверсионный заряд. А насколько увеличится падение напряжения на оксидной пленке? Ответ таков:

$$\Delta V_{ox} = -\frac{\Delta Q}{C_{ox}} = \frac{1.6 \times 10^{-19} \times (5 \times 10^{12})}{2 \times 10^{-6}} = 0.4 \text{ В},$$

где мы воспользовались значением ёмкости оксидной пленки из последнего равенства (60). Мы видим, что увеличение падения напряжения на оксидной пленке более чем в 10 раз превышает увеличение поверхностного потенциала.

Этот пример показывает, что поскольку небольшое изменение поверхностного потенциала ведет к большому изменению (скачку) электронного заряда в режиме сильной инверсии, в итоге имеет место большой скачок в падении напряжения на оксидной пленке. В этом примере напряжение на затворе должно возрасти на 0.44 В для того, чтобы поверхностный потенциал увеличился на 0.04 В. Таким образом, выше порогового напряжения существенная доля напряжения на затворе падает на оксидной пленке и лишь его незначительная часть падает на полупроводниковой подложке. Это объясняет почему поверхностный потенциал  $\psi_s$  меняется медленно с ростом  $V_G$  при  $V_G > V_T$ , как это показано на рис. 5.

Уравнение (30) позволяет иначе посмотреть на эту ситуацию. Мы ранее нашли, что

$$\psi_s = \frac{V_G}{m},$$

где из (31)

$$m = 1 + \frac{C_s}{C_{ox}},$$

причем ниже порогового напряжения  $C_s < C_{ox}$  ( $C_s = C_D$  ниже  $V_T$ ) и коэффициент эффекта подложки в режиме истощения  $m$  близок к единице, однако, выше  $V_T$  ёмкость полупроводниковой подложки становится очень большой и  $C_s \gg C_{ox}$  с  $m \gg 1$ . Используя результаты подсчетов в приведенном выше примере, для коэффициента  $m$  получаем оценку в  $m \approx 16$ , так что двухемкостной делитель напряжения на рис. 10 экранирует поверхностный потенциал  $\psi_s$  от влияния напряжения на затворе  $V_G$ .

Итак, показано, что электронный заряд  $Q(\psi_s)$  зависит экспоненциально от  $\psi_s$  как ниже, так и выше порогового напряжения  $V_T$ . Зависимость ниже  $V_T$  (44) идет как  $\exp(\psi_s/kT)$ , тогда как выше  $V_T$  (48) – как  $\exp(\psi_s/2kT)$ , однако, экспоненциальная зависимость присутствует в обоих режимах.

Ниже порога  $Q(V_{GS})$  изменяется экспоненциально с напряжением на затворе  $V_{GS}$  (45), поскольку имеет место зависимость  $\psi_s(V_{GS})$  (30). Однако, выше порога ситуация другая: выше порога  $Q(V_{GS})$  меняется линейно с  $V_{GS}$  (58), поскольку  $\psi_s \propto \ln(V_{GS})$ .

Суммируя результаты, имеем следующую картину:

$$\begin{aligned} Q(V_G) &= -(m-1)C_{ox} \frac{kT}{q} e^{q(V_G - V_T)/mkT}, \quad (V_G \ll V_T) \\ Q(V_G) &= -C_G(V_G - V_T). \quad (V_G \gg V_T) \end{aligned} \quad (63)$$

По этим уравнениям с привлечением уравнения (30/гл. 2)

$$I_{DS}/W = |Q(x=0)| \langle v(x=0) \rangle \quad (64)$$

можно вычислить ток стока ниже и выше порогового напряжения. Полезно было бы, конечно, иметь одно уравнение, одинаково хорошо работающее и ниже и выше порога. Соотношение  $Q \propto V_G$  можно просчитать численно, вместе с тем, как будет показано позже, доступно эмпирическое выражение, обеспечивающее корректные результаты как ниже, так и выше порогового напряжения.



### 3.4. Подвижный электронный заряд: исключительно тонкий SOI

До сих пор мы обсуждали свойства и особенности массивной структуры MOS. Сейчас перейдем к рассмотрению структуры с исключительно тонкой кремниевой подложкой. Эта структура MOS характерна для нынешней тенденции миниатюризации транзисторов и носит название чрезвычайно тонкой КНИ-подложки (Extremely Thin Silicon-On-Insulator/ETSOI).

Электроны в тонком слое кремния ведут себя как квантовые частицы в потенциальной яме. Из-за ограничения в одном измерении, по толщине подложки вдоль оси  $y$ , электроны нужно рассматривать как 2D квазичастицы и для моделирования плотности электронов нужно вводить 2D плотность состояний. Отметим также, что даже и в массивной структуре MOSFET, рассматриваемой до сих пор, на поверхности раздела Si/SiO<sub>2</sub>, строго говоря, также нужно учитывать квантовые ограничения. Если бы мы учли эти квантовые ограничения для массивной структуры MOSFET, результаты численного моделирования были бы несколько другими, но качественно картина не изменилась бы.

Далее нас будет интересовать та же задача, что и ранее, – зависимость электронного заряда  $Q$  как от поверхностного потенциала  $\psi_s$ , так и от напряжения на затворе  $V_G$ , но теперь применительно к ETSOI. Будем с самого начала рассматривать электроны как 2D квазичастицы. В конечном итоге мы убедимся, что результаты нашего рассмотрения качественно будут такими же, как и при классическом анализе массивных MOSFET.

*Подвижный электронный заряд*

Электронный заряд в ETSOI определяется как

$$Q = -q \int_0^{t_{Si}} n(y) dy = -qn_s, \quad [K/m^2] \quad (65)$$

где  $t_{Si}$  есть толщина кремниевого слоя. Рассмотрим квантовую яму с несколькими подзонами (рис. 12). Если бы мы рассматривали электроны как классические частицы, плотность электронов была бы однородной [9]

$$\begin{aligned} n_0 &= N_{3D}^c \mathfrak{F}_{1/2}(\eta_F), \quad [M^{-3}] \\ \eta_F &= (E_F - E_C)/kT, \\ n_s &= n_0 t_{Si}, \quad [M^{-2}] \end{aligned} \quad (66)$$

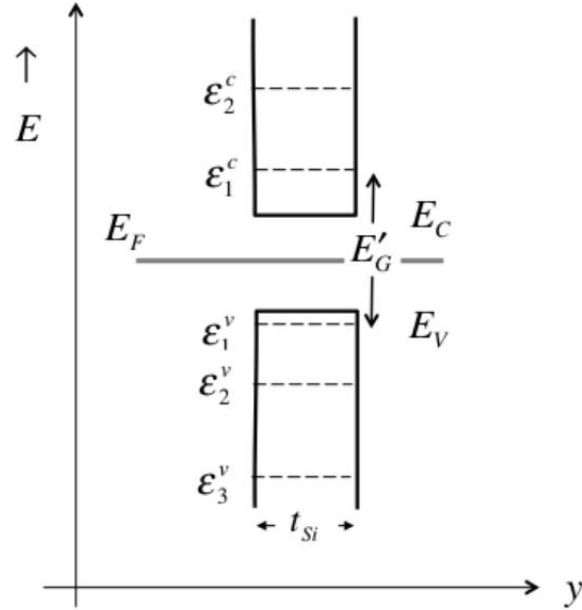


Рис. 12. Зонная диаграмма ETSOI.

где  $N_{3D}^c$  – эффективная плотность состояний,  $n_0$  – объемная и  $n_s$  – поверхностная плотности электронов,  $\mathfrak{F}_j(\eta_F)$  – интеграл Ферми – Дирака порядка  $j$  [9].

Квантовые ограничения порождают зоны проводимости, дно каждой из которых характеризуется энергией  $E_C + \varepsilon_m^c$  (рис. 12). Аналогичное квантование имеет место в валентной зоне. Изменяются также волновые функции: для бесконечно глубокой потенциальной ямы  $\psi(y) \propto \sin(m\pi y/t_{Si})$ . Плотность электронов в каждой валентной подзоне  $n(y)$  пропорциональна  $\psi^*(y)\psi(y) = \sin^2(m\pi y/t_{Si})$  и определяет пространственное распределение электронов внутри ямы. Полную интегральную электронную плотность (65) находим интегрированием 2D плотности состояний, помноженной на функцию Ферми. Для подзоны  $m$  имеем

$$n_S^m = \int_0^{t_{Si}} n(y) dy = \int_0^{\infty} D_{2D}(E) f_0(E) dE = N_{2D}^{m,c} \ln\left(1 + e^{(E_F - E_C - \varepsilon_m^c)/kT}\right), \quad (67)$$

где

$$N_{2D}^{m,c} \equiv g_v^m \frac{m_{D,m}^* kT}{\pi \hbar^2}. \quad (68)$$

Полную поверхностную электронную плотность находят суммированием вкладов от всех подзон.

Далее мы предположим, что только нижняя подзона заселена. Это допущение представляется разумным для очень узкой потенциальной ямы с далеко отстоящими энергиями подзон. Соответственно этому, электронный заряд

$$Q(\psi_S) = -qn_S = -qN_{2D}^{1c} \ln\left(1 + e^{(E_F - E_C - \varepsilon_1^c)/kT}\right). \quad (69)$$

В структуре ETSOI затвор используется для управления поверхностным потенциалом  $\psi_S$  в квантовой яме. Структура MOS ETSOI показана на рис. 13.

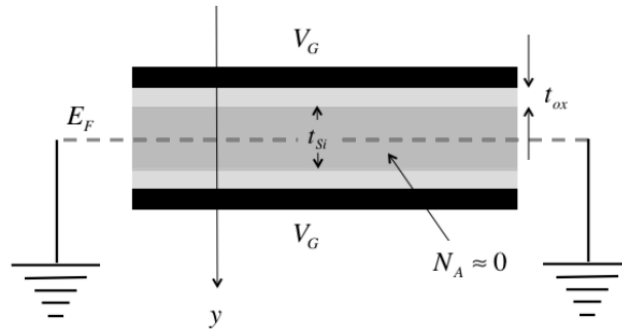


Рис. 13. Структура MOS ETSOI с двойным затвором (Double Gate/DG). Верхний и нижний оксидные слои одинаковы, одно и то же напряжение прилагается к обоим затворам. Тонкая кремниевая подложка заземлена, так что  $E_F$  есть равновесный уровень Ферми. Электрическое поле в направлении  $y$  симметрично относительно пунктирной линии.

Предполагаем также, что кремниевая подложка достаточно тонкая и электронная плотность невелика, так что дно потенциальной ямы практически плоское: электростатический потенциал в яме  $\psi_S$  не зависит от ширины ямы  $u$ . Это позволяет переписать (69) в виде

$$Q(\psi_S) = -qn_S = -qN_{2D}^{1c} \ln\left(1 + e^{(E_F - E_{C0} + q\psi_S - \varepsilon_1^c)/kT}\right), \quad (70)$$

где  $E_C = E_{C0} - q\psi_S$ ,  $\psi_S$  контролируется потенциалом от двух затворов, а  $E_{C0}$  есть значение  $E_C$  при  $\psi_S = 0$ , что определяется работой выхода затворов. Наконец, примем невырожденную статистику носителей тока, что позволит результаты по ETSOI сравнить непосредственно с результатами по массивной подложке MOS. Итак, окончательное выражение для электронного заряда

$$Q(\psi_S) = -qn_S = -qN_{2D}^{1c} e^{(E_F - E_{C0} + q\psi_S - \varepsilon_1^c)/kT}. \quad (71)$$

Перепишем это уравнение для электронов следующим образом

$$Q(\psi_S) = -qn_{S0}e^{q\psi_S/kT}, \quad (72)$$

где

$$n_{S0} = N_{2D}^{1c} e^{(E_F - E_{C0} - \varepsilon_1^c)/kT}. \quad (73)$$

Аналогично запишем уравнение для дырочного заряда

$$Q_p(\psi_S) = qp_{S0}e^{-q\psi_S/kT}, \quad (74)$$

где

$$p_{S0} = N_{2D}^{1v} e^{(E_{V0} - \varepsilon_1^v - E_F)/kT}. \quad (75)$$

Первая и основная задача – понять зависимость  $Q_S(\psi_S)$  для ETSOI и сравнить результаты с таковыми же для массивной структуры MOS, суммированными на рис. 19/гл. 2.

Если структура ETSOI не допирована, то заряд в кремниевой подложке определяется только электронами и дырками, а именно:

$$Q_S(\psi_S) = q(p_S - n_S) = q(p_{S0}e^{-q\psi_S/kT} - n_{S0}e^{q\psi_S/kT}). \quad (76)$$

Примем за нуль отсчета потенциала  $\psi_S = 0|_{n_{S0}=p_{S0}=n_{Si}}$ , тогда

$$\underline{Q_S}|_{\psi_S=0} = 0, \quad (77)$$

и (76) переписывается в таком виде:

$$Q_S(\psi_S) = qn_{Si} \left( e^{-q\psi_S/kT} - e^{q\psi_S/kT} \right). \quad (78)$$

Если мы рассматриваем  $n$ -MOSFET, то поверхностная плотность подвижных электронов

$$n_S(\psi_S) = n_{Si} e^{q\psi_S/kT}. \quad [M^{-2}]. \quad (79)$$

Рисунок 14 фактически демонстрирует каким образом напряжение на затворе влияет на зонную структуру ETSOI.

При положительном напряжении на затворе потенциал в полупроводнике растет (рис. 10/гл. 2), зона проводимости понижает свою энергию и опускается ближе к уровню Ферми, концентрация электронов растет экспоненциально.

При отрицательном напряжении на затворе валентная зона движется вверх и концентрация дырок растет экспоненциально.

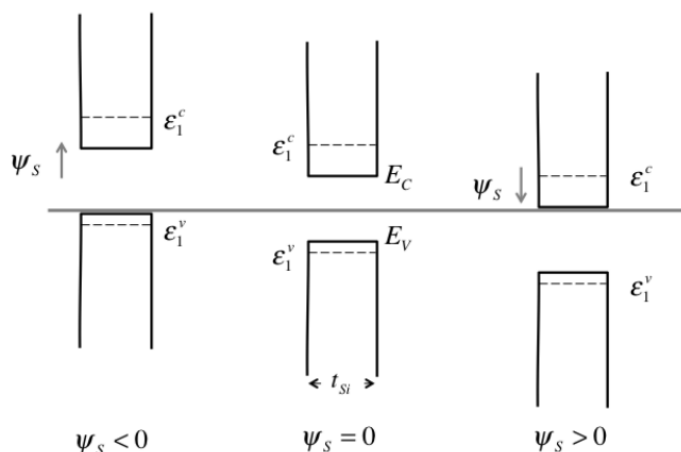


Рис. 14. Иллюстрация того, как отрицательный, нулевой и положительный потенциал  $\psi_s$  влияет на зонную диаграмму ETSOI.

Рис. 15 иллюстрирует итоговую зависимость  $Q_s(\psi_s)$  для ETSOI, которую теперь можно сравнить с таковой для массивной структуры MOSFET на рис. 19/гл. 2.

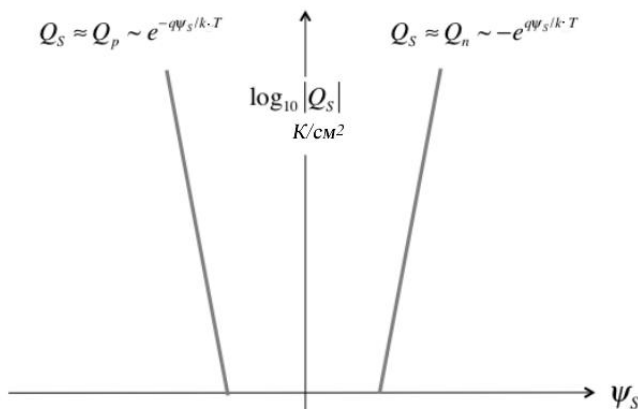


Рис. 15. Иллюстрация зависимости заряда в полупроводнике с изменением потенциала для структуры ETSOI в сравнении с аналогичным графиком для массивной структуры MOSFET на рис. 19/гл. 2.

В случае структуры ETSOI кремниевая подложка не была допирована, так что не должно быть области с зависимостью от  $\sqrt{\psi_s}$ , характерной для режима истощения. Как только потенциал  $\psi_s$  принимает достаточно большое положительное или отрицательное значение, создается большая плотность соответственно электронов или дырок. В режимах сильной инверсии или накопления заряд в массивной структуре MOS растет экспоненциально с поверхностным потенциалом (рис. 19/гл. 2.). Тот же эффект наблюдается и для

структуры ETSOI с тем отличием, что в режимах инверсии или накопления заряд в массивной структуре MOS изменяется как  $\exp(q\psi_s/2kT)$ , а для структуры ETSOI – как  $\exp(q\psi_s/kT)$ . Это различие можно увязать с тем обстоятельством, что потенциальная яма в массивной структуре MOS связана, как уже упоминалось, с электрическим полем на границе раздела Si/SiO<sub>2</sub>, тогда как тонкий кремниевый слой в структуре ETSOI сам по себе является потенциальной ямой.

Уравнение (78) аналогично уравнению (48) для массивной структуры MOS, а чтобы им пользоваться нужно оценить плотность  $n_{Si}$ . В общем случае  $n_{s0}$  (73) есть поверхностная электронная плотность при  $\psi_s = 0$ . Она зависит от того, где находится уровень Ферми, а на это влияет работа выхода электрода затвора.

Рассмотрим численный пример собственного полупроводника с  $\psi_s = 0$ , так что  $n_{s0} = p_{s0} = n_{Si}$ . Для оценки величины  $n_{Si}$  возьмем следующие типичные данные для структуры ETSOI:

$$\begin{aligned}
 t_{Si} &= 5 \text{ нм}, \\
 m_l^* &= 0.92m_0, \quad m_t^* = 0.19m_0, \quad m_{hh}^* = 0.54m_0, \quad m_0 = 9.109 \times 10^{-31} \text{ кг}, \\
 E_G &= 1.125 \text{ эВ}, \\
 T &= 300 \text{ К}.
 \end{aligned}$$

В связи с приведенными исходными данными обратимся к зонной структуре кремния (рис. 16).

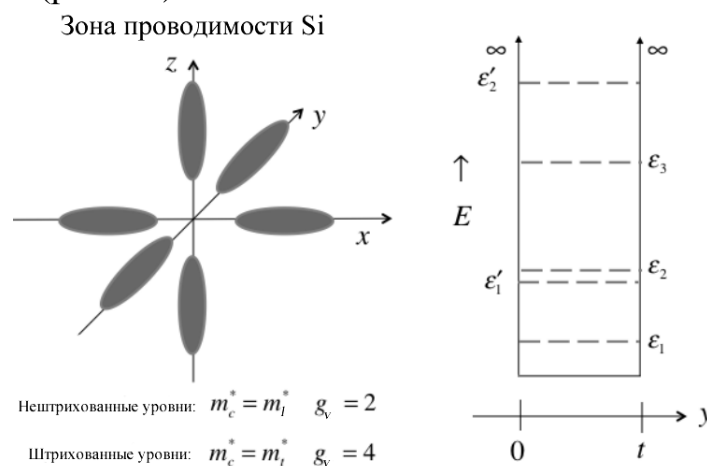


Рис. 16. Эллипсоиды вращения изоэнергетических поверхностей в кремнии направлены в обратном пространстве таким образом, что большая полуось совпадает с направлениями (100). Справа показаны уровни энергии подзон для Si (100). Эффективные массы соответствуют выделенному направлению  $y$ .

На этом рисунке показаны эллипсоиды вращения изоэнергетических поверхностей в зоне проводимости кремния

$$E = \frac{\hbar^2 k_x^2}{2m_{xx}^*} + \frac{\hbar^2 k_y^2}{2m_{yy}^*} + \frac{\hbar^2 k_z^2}{2m_{zz}^*}. \quad (80)$$

Наименьшие энергии соответствуют шести различным ориентациям в зоне Бриллюэна вдоль трех координатных осей (долинное вырождение  $g_v = 6$ ). Вводят две различных эффективных массы: тяжелая продольная (longitudinal) эффективная масса  $m_l^*$  и легкая поперечная (transverse) эффективная масса  $m_t^*$ . Для Si  $m_l^* = 0.92m_0$  и  $m_t^* = 0.19m_0$ . Например, для эллипсоидов, ориентированных вдоль оси  $x$ ,  $m_{xx}^* = m_l^*$ ,  $m_{yy}^* = m_{zz}^* = m_t^*$ .

В соответствии со спектром энергий в прямоугольной потенциальной яме

$$\varepsilon_n = \frac{\hbar^2 n^2 \pi^2}{2m^* t^2}, \quad (81)$$

энергии подзон определяются значением эффективной массы – той, которая соответствует выделенному направлению, вдоль которого наложено ограничение, т. е. вдоль направления  $y$  в нашем случае. Как видим из рис. 16, для Si (100) с ограничением вдоль оси  $y$  двум эллипсоидам из шести соответствуют тяжелые массы  $m_l^*$  и четырем – легкие  $m_t^*$ . В результате имеем две различные серии подзон – нештрихованные уровни энергии, определяемые массой  $m_l^*$  и вырождением  $g_v = 2$ , и штрихованные, определяемые массой  $m_t^*$  и вырождением  $g_v = 4$ . Наини́зшая подзона с  $n = 1$  нештрихованная. В простом примере, рассматриваемом далее, предполагается, что занята только нижняя нештрихованная подзона с массой  $m_l^*$  вдоль оси  $y$  и массой  $m_t^*$  в плоскости  $xz$ .

Сначала вычислим некоторые нужные нам величины. Для подзоны проводимости с  $n = 1$  эффективная масса  $m^* = m_l^*$ . Для дырок в первой валентной подзоне  $m^* = m_{hh}^*$ . Из (81) находим

$$\varepsilon_1^c = 0.016 \text{ эВ}, \quad \varepsilon_1^v = 0.028 \text{ эВ}.$$

Квантовые ограничения увеличивают эффективную запрещенную зону, поскольку дну зоны проводимости теперь соответствует энергия  $E_c + \varepsilon_1^c$ , а потолку валентной зоны – энергия  $E_v - \varepsilon_1^v$ . Эффективная запрещенная зона для структуры ETSOI будет

$$E'_G = E_G + \varepsilon_1^c + \varepsilon_1^v = 1.169 \text{ эВ}. \quad (82)$$

Эффективная 2D плотность состояний дается формулой (68). Для первой подзоны зоны проводимости долинное вырождение равно двум, а эффективная масса в плоскости  $xz$  есть  $m_i^*$ , так что эффективная масса  $m_{D,1}^* = 2m_i^*$  в (68). Для валентной зоны  $g_v = 1$  и  $m_{D,1}^* = m_{hh}^*$ . Окончательно из (68) имеем:

$$N_{2D}^{1,c} = 4.11 \times 10^{12} \text{ см}^{-2}, \quad N_{2D}^{1,v} = 5.84 \times 10^{12} \text{ см}^{-2}. \quad (83)$$

Из (73) и (75) находим поверхностную плотность носителей тока:

$$n_S = N_{2D}^{1,c} e^{(E_F - E_C - \varepsilon_1^c)/kT}, \quad p_S = N_{2D}^{1,v} e^{(E_V - E_F - \varepsilon_1^v)/kT}. \quad (84)$$

Перемножая эти плотности, получаем квадрат собственной плотности

$$n_S p_S = N_{2D}^{1,c} N_{2D}^{1,v} e^{-E_G'/kT} = n_{Si}^2,$$

который не зависит от положения уровня Ферми и

$$n_{Si} = \sqrt{n_S p_S} = \sqrt{N_{2D}^{1,c} N_{2D}^{1,v}} e^{-E_G'/2kT}. \quad [M^{-2}].$$

Обращает на себя внимание схожесть полученного выражения со стандартным выражением для внутренней плотности носителей тока в массивном полупроводнике [1]. Подставляя теперь все известные нам численные значения  $N_{2D}^{1,c}$ ,  $N_{2D}^{1,v}$  и  $E_G'$ , получаем очень маленькую плотность собственных носителей

$$n_{Si} = 8.5 \times 10^2 \text{ см}^{-2}, \quad (85)$$

которая может быть подавлена зарядами на границе раздела Si/SiO<sub>2</sub> или даже случайными допантами в тонкой пленке кремния.

### *Подвижный электронный заряд ниже порогового напряжения*

Выражение для электронного заряда в зависимости от поверхностного потенциала  $Q(\psi_S)$ , одинаково пригодное как ниже, так и выше порогового напряжения, было получено для массивной структуры MOS в виде уравнения (39), однако, в этих двух случаях  $V_G < V_T$  и  $V_G > V_T$  использовались разные выражения для среднего электрического поля. Рассмотрим ситуацию для структуры ETSOI, для которой, согласно (72),

$$Q(\psi_S) = -q n_{Si} e^{q\psi_S/kT}, \quad (86)$$



где предполагается, что фермиевский уровень располагается таким образом, что  $n_{s0} = n_{si}$  при  $\psi_s = 0$ . Выражение (86) справедливо как ниже, так и выше порогового напряжения.

Далее нам нужно увязать  $Q(\psi_s)$  ниже порогового напряжения с напряжением на затворе. Структура DG ETSOI на рис. 13 симметрична относительно пунктирной линии, проходящей через середину тонкой кремниевой подложки. Половина электронного заряда в SOI локализуется под верхним затвором, а другая половина – над нижним затвором. Учитывая симметрию структуры DG, достаточно установить связь между напряжением на верхнем затворе с зарядом в верхней половине SOI. Начнем с общего уравнения (1), а именно:

$$V'_G = \Delta V_{ox} + \psi_s. \quad (87)$$

По закону Гаусса, электрическое поле на верхней границе раздела Si/SiO<sub>2</sub>

$$E_{ox} = -\frac{Q_s(\psi_s)}{2\epsilon_s} = -\frac{Q(\psi_s)}{2\epsilon_s}. \quad (88)$$

Падение потенциала на верхней оксидной пленке

$$\Delta V_{ox} = E_{ox} t_{ox}. \quad (89)$$

Из этих уравнений и из (5) следует, что

$$V'_G = -\frac{Q(\psi_s)}{2C_{ox}} + \psi_s. \quad (90)$$

Ниже порогового напряжения заряд подложки исключительно мал, так что падение напряжения на оксидной пленке  $\Delta V_{ox}$ , можно считать, пренебрежимо мало, и можно было сразу вместо (87) записать, что

$$V'_G = \psi_s. \quad (91)$$

В случае массивной структуры MOS  $\psi_s = V'_G / m$  (30), где  $m > 1$  (31), однако, в случае DG ETSOI  $m \equiv 1$ , что означает полный контроль затвора над поверхностным потенциалом и является преимуществом структуры DG ETSOI. Теперь, благодаря (91), общее выражение (86) в режиме  $V_G < V_T$  таково:

$$Q(\psi_s) = -qn_{si}e^{qV_G/kT}. \quad (92)$$

Уравнение (92) определяет электронный заряд структуры DG ETSOI в подпороговом режиме. В случае массивной структуры MOS окончание режима слабой инверсии связывалось с уравниванием  $\psi_s$  и  $2\psi_B$ . Мы не можем пользоваться подобным критерием в случае ETSOI, поскольку эта структура не допирована. В этом случае можно утверждать, что при понижении первой подзоны зоны проводимости таким образом, что  $E_F = E_C + \varepsilon_1^c$ , концентрация электронов становится существенной. Уравнение (84) показывает, что при выполнении этого условия  $n_s = N_{2D}^{1c}$ . В результате потенциал полупроводниковой подложки в структуре ETSOI на входе в режим инверсии можно найти из

$$n_s(\psi_s) = n_{Si} e^{q\psi_s/kT} = N_{2D}^c, \quad (93)$$

откуда для потенциала при входе в режим инверсии имеем

$$\psi_s^{inv} = \frac{kT}{q} \ln \frac{N_{2D}^c}{n_{Si}}. \quad (94)$$

Из уравнения (91) и учитывая рассуждения, которые привели к этому уравнению, при входе в режим инверсии

$$V_T' = \psi_s^{inv} = \frac{kT}{q} \ln \frac{N_{2D}^c}{n_{Si}}. \quad (95)$$

Из (95) и (92) окончательно имеем

$$Q(V_G) = -C_Q \frac{kT}{q} e^{q(V_G - V_T)/kT}, \quad (96)$$

где

$$C_Q = q^2 D_{2D} \quad (97)$$

называют квантовой ёмкостью, к обсуждению и использованию которой в качестве еще одного параметра ETSOI мы вернемся позже, а  $D_{2D}$  есть 2D плотность состояний применительно к структуре ETSOI со сверхтонкой подложкой.

Мы получили важный результат о зависимости электронного заряда ETSOI от напряжения на затворе (96), который полезно сравнить с аналогичной зависимостью (45) для массивной структуры MOS. В обоих случаях заряд в подпороговом режиме увеличивается экспоненциально с  $V_G$ . Различие между этими двумя структурами в том, что для DG ETSOI характерен идеальный

подпороговый наклон  $m = 1$ , тогда как для массивных структур MOS  $m \approx 1.1 - 1.3$ . Это означает, что при одном и том же увеличении напряжения на затворе электронный заряд в структуре DG ETSOI растет быстрее, чем в массивной структуре MOS.

Для массивных MOSFET при входе в режим инверсии поверхностный потенциал  $\psi_s = 2\psi_B$ . Для структуры ETSOI этот критерий определяется уравнением (94). Насколько сравнимы  $\psi_s^{inv}$  для структуры ETSOI с  $\psi_s^{inv} = 2\psi_B$  для массивной структуры MOS? Подстановка расчетных данных (83) и (85) из предыдущего примера в уравнение (94) приводит к следующему результату:

$$\psi_s^{inv} = \frac{kT}{q} \ln \frac{N_{2D}^c}{n_{Si}} = 0.026 \cdot \ln \frac{4.11 \times 10^{12}}{8.5 \times 10^2} = 0.58 \text{ В},$$

согласно которому можно ожидать, что  $q\psi_s^{inv}$  составляет примерно половину от ширины эффективной запрещенной зоны DG ETSOI (82). Поскольку уровень Ферми при  $\psi_s = 0$  был почти посередине запрещенной зоны (рис. 14), то  $n_{S0} = p_{S0} = n_{Si}$ . Поверхностный потенциал при входе в режим инверсии такой, что он понижает  $E_{C0} + \varepsilon_1^c$  до  $E_F$ , что и служит критерием для входа в этот режим.

В рассмотренном примере требуемый изгиб зон в режиме инверсии составляет примерно половину от ситуации в массивной структуре MOS, поскольку в последней уровень Ферми в  $p$ -Si располагается вблизи валентной зоны, так что изгиб зоны проводимости должен покрыть почти всю запрещенную зону для того, чтобы уровень Ферми совместился с зоной проводимости.

#### *Подвижный электронный заряд выше порогового напряжения*

Уравнение (86) справедливо как ниже, так и выше порогового напряжения. В рассмотренном выше случае ( $V_G < V_T$ ) мы могли предполагать, что падение напряжения на оксидном слое незначительное и таким образом связать электронный заряд в подпороговом режиме с напряжением на затворе согласно (96). Выше порога,  $V_G > V_T$ , падение напряжения на оксидном слое становится значительным, и зависимость  $Q \propto V_G$  изменится.

Уравнение (86) дает зависимость  $Q \propto \psi_s$ , а уравнение (90) связывает  $\psi_s$  с  $V_G$ . Численное решение этих двух уравнений показывает, что заряд растет

приблизительно линейно с  $V_G$  при  $V_G > V_T$ , как и в случае массивной структуры MOS. Наклон этой линейной зависимости

$$C_G = \frac{dQ_M}{dV_G} = \frac{d(-Q_S)}{dV_G} = \frac{d(-Q)}{dV_G}. \quad (98)$$

Дифференцируя (90) по  $(-Q)$ , находим, что

$$\frac{1}{C_G} = \frac{1}{2C_{ox}} + \frac{1}{C_S}, \quad (99)$$

и записываем инверсионный заряд выше порогового напряжения в виде

$$Q(V_G) = -C_G(V_{GS} - V_T), \quad (100)$$

где  $C_G \approx \text{Const}$ . Для  $C_S \gg 2C_{ox}$ ,  $C_G \approx 2C_{ox}$ . Коэффициент двойка появляется из-за DG в этой модели ETSOI (рис. 13).

Поскольку ёмкость полупроводникового слоя конечная,  $C_G$  немного меньше, чем  $2C_{ox}$ , как это было и в случае массивной структуры MOS. Воспользовавшись (86), находим ёмкость полупроводникового слоя ETSOI

$$C_S^{inv} = \frac{d(-Q)}{d\psi_S} = \frac{-Q}{kT/q}. \quad (101)$$

Сравнивая этот результат с аналогичным (55) для массивной структуры MOS, видим, что ёмкость полупроводникового слоя ETSOI в два раза больше аналогичной величины для массивной структуры MOS.

Оба результата (101) и (55) предполагают невырожденную статистику электронов. А что получится в пределе вырождения? В этом случае  $E_F \gg E_{c0} + \varepsilon_1^c$  и уравнение (69) становится таким

$$Q(\psi_S) = -qn_S = -qN_{2D}^{1c} (E_F - E_{c0} - \varepsilon_1^c + q\psi_S) / kT, \quad (102)$$

так что ёмкость полупроводникового слоя ETSOI становится равной

$$C_S^{inv} = \frac{d(-Q)}{d\psi_S} = \frac{q^2 N_{2D}^{1c}}{kT} = q^2 \frac{m_D^*}{\pi \hbar^2} = q^2 D_{2D} = C_Q, \quad (103)$$

где  $C_Q$  есть квантовая ёмкость (97).

В более сложных ситуациях таких, как вдали от предела вырождения, при нескольких занятых подзонах, при утолщении полупроводникового слоя, когда учет изгиба зон становится существенным, ёмкость кремниевого слоя вынуждены рассчитывать численно. Однако, в любом случае ёмкость полупроводниковой подложки будет связана с плотностью состояний электронов. Для структур MOS, в которых используются полупроводники с легкими эффективными массами (напр., III – V полупроводники), можно ожидать, что ёмкость полупроводникового слоя и общая ёмкость затвора уменьшаться сравнительно с кремниевой подложкой.

Наконец, давайте перепишем уравнение (99) в таком виде

$$C_G = \frac{(2C_{ox})C_S}{(2C_{ox}) + C_S}. \quad (104)$$

Мы могли бы ожидать, что полная ёмкость затвора для DG ETSOI будет в два раза больше ёмкости затвора для ETSOI с одиночным затвором, однако, она все же несколько меньше. Чтобы увидеть почему, перепишем (104) следующим образом:

$$C_G = 2 \times \left[ \frac{C_{ox}(C_S/2)}{C_{ox} + (C_S/2)} \right]. \quad (105)$$

Ёмкости в скобках есть последовательное соединение  $C_{ox}$  и  $C_S/2$ . Ёмкость DG ETSOI распределяется между двумя затворами, так что каждый из двух затворов характеризуется ёмкостью, которая несколько меньше ёмкости однозатворного ETSOI [10].

Приведем численные оценки ёмкости инверсионного слоя и толщины эквивалентной ёмкости. К выше указанным исходным данным по кремнию добавим  $t_{ox} = 1.8 \text{ нм}$ ,  $\epsilon_{ox} = 4.0$ . Ответим на следующие пять вопросов:

1) Какова ёмкость полупроводникового слоя ETSOI при  $n_s = 1 \times 10^{13} \text{ см}^{-2}$  ?

Эта поверхностная плотность электронов типична для режима «ON» современного MOSFET. Из уравнения (101) находим

$$C_S^{inv} = \frac{-Q}{kT/q} = \frac{qn_s}{kT/q} = 61.6 \times 10^{-6} \Phi / \text{см}^2,$$

что в два раза больше ёмкости, найденной для массивной структуры MOSFET (59). По сравнению с  $C_{ox} = 2.0 \times 10^{-6} \Phi / \text{см}^2$  это очень большая, неприемлемая

ёмкость, величина которой существенно уменьшится, если обратиться к статистике Ферми – Дирака и учесть квантовые ограничения для структуры ETSOI. Оценка  $C_S$  в пределе вырождения по формуле (103) дает

$$C_S = C_Q = 25.4 \times 10^{-6} \Phi/\text{см}^2, \quad (106)$$

что меньше половины величины, полученной выше для невырожденной статистики.

2) Какова ёмкость затвора?

Согласно (104),

$$C_G = \frac{(2C_{ox})C_S}{(2C_{ox}) + C_S} = \frac{2C_{ox}}{1 + 2C_{ox}/C_S} = \frac{2C_{ox}}{1 + 4.0/25.4} = 0.86(2C_{ox}) = 3.44 \times 10^{-6} \Phi/\text{см}^2,$$

где мы воспользовались  $C_{ox}$  (32) и  $C_S$  (106). Как и ожидалось,  $C_G < 2C_{ox}$ .

3) Какова толщина эквивалентной ёмкости DG ETSOI?

Воспользуемся определением  $CET$  по (33) и учтем DG в ETSOI:

$$C_G/2 \equiv \frac{\varepsilon_{ox}}{CET} \Rightarrow CET = \frac{\varepsilon_{ox}}{C_G/2} = \frac{4.0 \times 8.854 \times 10^{-12} \times 10^{-2}}{(3.44/2) \times 10^{-6}} = 2.06 \text{ нм}.$$

Обращает на себя внимание, что  $CET$  в DG ETSOI несколько больше толщины оксидного слоя 1.8 нм и этот эффект превышения более существенный по сравнению с прежним результатом (61) и обязан он использованию статистики Ферми – Дирака и учету квантовых ограничений.

4) Каков поверхностный потенциал полупроводникового слоя DG ETSOI при  $n_s = 1 \times 10^{13} \text{ см}^{-2}$ ?

Из (79) имеем

$$\psi_s = \frac{kT}{q} \ln \frac{n_s}{n_{Si}} = 0.60 \text{ В}. \quad (107)$$

Вспомним, что ранее вхождение в режим инверсии мы определили происходящим при  $\psi_s^{inv} = 0.58$ , так что полученное значение  $\psi_s$  немного больше. В случае массивной структуры MOS поверхностный потенциал в режиме сильной инверсии был приблизительно на шесть  $kT/q$  больше чем  $2\psi_B$ . В нашем

случае поверхностный потенциал при сильной инверсии лишь на один  $kT/q$  больше чем  $\psi_s^{inv}$ . Это различие отчасти обязано тому обстоятельству, что в случае структуры ETSOI электронный заряд ведет себя как  $Q \propto \exp(q\psi_s/kT)$ , а в случае массивной структуры MOS – как  $Q \propto \exp(q\psi_s/2kT)$ , так что в последнем случае требуется бóльший изгиб зон, чтобы получить бóльший заряд в инверсионном слое.

5) Каков поверхностный потенциал полупроводникового слоя ETSOI при использовании статистики Ферми – Дирака?

Уравнение (70) связывает  $n_s$  и  $\psi_s$  в общем случае, а именно:

$$n_s = N_{2D}^{1c} \ln\left(1 + e^{(E_F - E_{C0} + q\psi_s - \varepsilon_f^c)/kT}\right).$$

Предположим, что  $\psi_s = 0$  при  $n_s = n_{Si}$  и полупроводник невырожденный при  $n_s = n_{Si}$ , тогда

$$n_{Si} = N_{2D}^{1c} e^{(E_F - E_{C0} - \varepsilon_f^c)/kT}.$$

Подставляя это выражение в предыдущее, получаем

$$n_s = N_{2D}^{1c} \ln\left(1 + \frac{n_{Si}}{N_{2D}^{1c}} e^{+q\psi_s/kT}\right),$$

откуда находим, что

$$\psi_s = \frac{kT}{q} \ln\left[\frac{N_{2D}^{1c}}{n_{Si}} \left(e^{n_s/N_{2D}^{1c}} - 1\right)\right],$$

а после подстановки  $n_s = 1 \times 10^{13} \text{ см}^{-2}$ ,  $n_{Si}$  (85) и  $N_{2D}^{1c}$  (83) окончательно получаем  $\psi_s = 0.64 \text{ В}$ , что на  $0.04 \text{ В}$  больше значения (107), полученного со статистикой Максвелла – Больцмана.

*Ход поверхностного потенциала с изменением напряжения на затворе*

Рисунок 5 иллюстрирует ход поверхностного потенциала с изменением напряжения на затворе для массивной структуры MOS. Ниже порогового напряжения поверхностный потенциал  $\psi_s = V_G/m$ , где  $m$  немного больше единицы. Выше порогового напряжения  $\psi_s$  растет медленно с увеличением  $V_G$ , поскольку  $m$  становится очень большим, так что бóльшая часть растущего

напряжения приходится на падение напряжения на оксидном слое, а не на полупроводниковой подложке. Можно ожидать аналогичные тенденции и для структуры ETSOI.

На рис. 17 сравнивается ход  $\psi_s \propto V_G$  для структур массивной MOS и ETSOI.

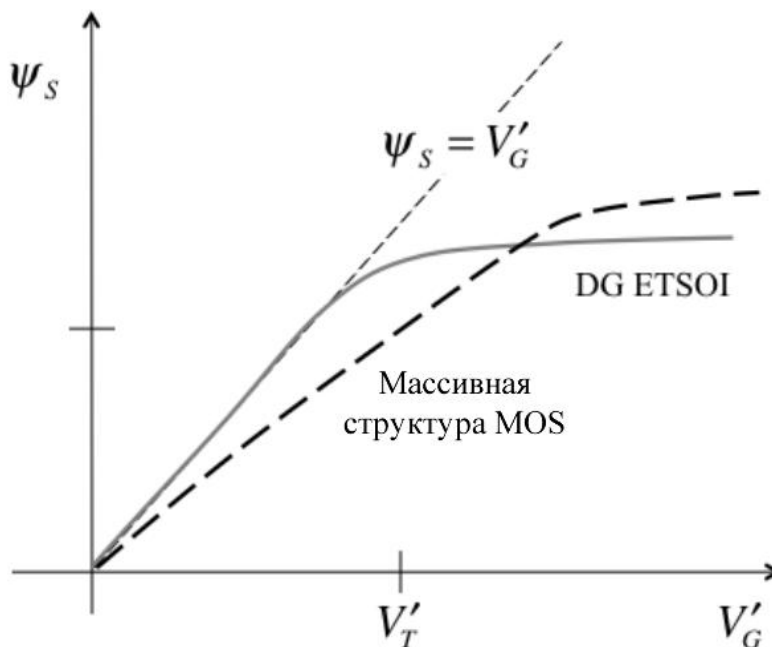


Рис. 17. Зависимость поверхностного потенциала от напряжения на затворе.

Сплошная линия относится к структуре ETSOI, а пунктирная – к массивной структуре MOS. Предполагается, что при  $\psi_s = 0$  уровень Ферми лежит близко к валентной зоне в обоих случаях.

В подпороговой области  $\psi_s = V_G$  (мелкопунктирная прямая), поскольку  $m = 1$  для структуры DG ETSOI. Выше порога  $\psi_s$  меняется медленно с  $V_G$  по той же причине, что и в случае с массивной структурой MOS. Замедление, однако, более слабое, поскольку в случае DG ETSOI инверсионный электронный заряд увеличивается более быстро с ростом  $\psi_s$ .

Итак, показано, что электронный заряд  $Q(\psi_s)$  для структуры ETSOI изменяется экспоненциально с потенциалом  $\psi_s$  как ниже, так и выше порога, уравнение (86). Аналогичный результат был получен и для массивной структуры MOS, уравнения (44) и (48).

Ниже порога  $Q(V_G)$  меняется экспоненциально с  $V_G$ , поскольку  $\psi_s = V_G$ , уравнение (96). Однако, выше порога зависимость  $Q(V_G)$  линейная, уравнение (100). Опять же, результаты аналогичны тем, которые были получены для массивной структуры MOS.



Сопоставим результаты для DG ETSOI

$$\begin{aligned} Q(V_G) &= -C_Q \frac{kT}{q} e^{q(V_G - V_T)/kT}, & (V_G \ll V_T) \\ Q(V_G) &= -C_G(V_G - V_T). & (V_G \gg V_T) \end{aligned} \quad (108)$$

с таковыми же (63) для массивной структуры MOS

$$\begin{aligned} Q(V_G) &= -(m-1)C_{ox} \frac{kT}{q} e^{q(V_G - V_T)/mkT}, & (V_G \ll V_T) \\ Q(V_G) &= -C_G(V_G - V_T). & (V_G \gg V_T) \end{aligned}$$

Организация канала проводимости (массивная структура MOS, ETSOI с двойным затвором или с одинарным и т. п.) качественно картину не меняет – заряд меняется экспоненциально с напряжением на затворе ниже порога и линейно – выше порога. В общем случае  $Q \propto V_G$  можно посчитать численно, вместе с тем, как будет показано в следующей главе, доступно эмпирическое выражение, обеспечивающее корректные результаты как ниже, так и выше порогового напряжения.

Подведем итоги. Физика процессов в полупроводниковой подложке MOSFET определяется изгибом зон, зависящим от поверхностного потенциала  $\psi_s$ , который в свою очередь определяется напряжением на затворе  $V_G$ . Получено достаточное общее уравнение (11), связывающее  $V_G$  с  $\psi_s$ . В условиях режима истощения получена более простая связь  $\psi_s$  с  $V_G$  (30), которой будем нередко пользоваться. Затем обсуждается поведение подвижного электронного заряда  $Q \propto \psi_s$  и  $Q \propto V_G$ : каким образом электронный заряд изменяется с поверхностным потенциалом и с напряжением на затворе в условиях до и после порогового напряжения. Правильные результаты в допороговой области напряжений и в режиме сильной инверсии получены не прибегая к численному решению уравнения Пуассона – Больцмана, вместе с тем численное решение этого уравнения покрывает как подпороговую область и область сильной инверсии, так и переходную область между ними. Рассмотрено поведение  $Q \propto \psi_s$  и  $Q \propto V_G$  для совершенно другой структуры MOS, структуры с исключительно тонкой кремниевой подложкой. Эта структура MOS характерна для нынешней тенденции миниатюризации транзисторов и носит название чрезвычайно тонкой КНИ-подложки (Extremely Thin Silicon-On-

Insulator/ETSOI). Тем не менее, мы убедились, что основные особенности структуры ETSOI подобны свойствам массивной структуры MOS. Получены зависимости  $Q \propto \psi_s$  и  $Q \propto V_G$  для структуры ETSOI как ниже, так и выше порога напряжений. Полезны, конечно, и численное решение или эмпирическая зависимость, охватывающие всю область напряжений на затворе, к чему мы еще вернемся. Полученные результаты говорят также о том, что 1D электростатика приемлема как для массивных структур MOS, так и для структур ETSOI. Однако, далее мы покажем как 2D электростатика объясняет почему структура DG ETSOI предпочтительна для очень коротких нанотранзисторов.

## Литература

1. R. F. Pierret, *Semiconductor Device Fundamentals* (New York: Addison-Wesley: 1996).
2. B. Streetman, S. Banerjee, *Solid State Electronic Devices* (New York: Prentice Hall: 2005).
3. Chenming Hu, *Modern Semiconductor Devices for Integrated Circuits* (London, UK: Pearson India: 2009).
4. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
5. Y. Tsividis, C. McAndrew, *Operation and Modeling of the MOS Transistor* (New York: Oxford Univ. Press: 2011).
6. Y. Taur, T. Ning, *Fundamentals of Modern VLSI Devices* (New York: Oxford Univ. Press: 2013).
7. M. Lundstrom, Xingshu Sun, *Notes on the Solution of the Poisson – Boltzmann Equation for MOS Capacitors and MOSFETs* (West Lafayette, Indiana: Purdue University, USA); [www.nanohub.org/resources/5338](http://www.nanohub.org/resources/5338).
8. M. Lundstrom, *ECE 612 Lecture 14:  $V_T$  Engineering* (West Lafayette, Indiana: Purdue University, USA); [www.nanohub.org/resources/5670](http://www.nanohub.org/resources/5670).
9. Ю. А. Кругляк, Нанoeлектроника «снизу – вверх» (Одесса: ТЭС: 2015).
10. D. Vasileska, D. K. Schroder, D. K. Ferry, *IEEE Trans. Electron Dev.*, **44**: 584 – 587 (1997).

## Глава 4. 2D электростатика MOS и модель виртуального истока

### 4.1. Введение

Рассмотрим 2D электростатику MOS и обусловленные ею эффекты, а затем сформулируем ставшую уже классической модель виртуального истока без учета возможности баллистического транспорта в явном виде.

Мы до сих пор обсуждали каким именно образом поверхностный потенциал полупроводниковой подложки изменяется с ростом напряжения на затворе, прибегая к рассмотрению лишь 1D электростатики. В нанотранзисторах с короткими каналами проводимости потенциалы, подаваемые на исток и сток, создают сильное электрическое поле вдоль направления канала. Совершенно очевидно, что электростатический потенциал в короткоканальном MOSFET будет сильно изменяться в обоих  $x$  и  $y$  направлениях (рис. 1).

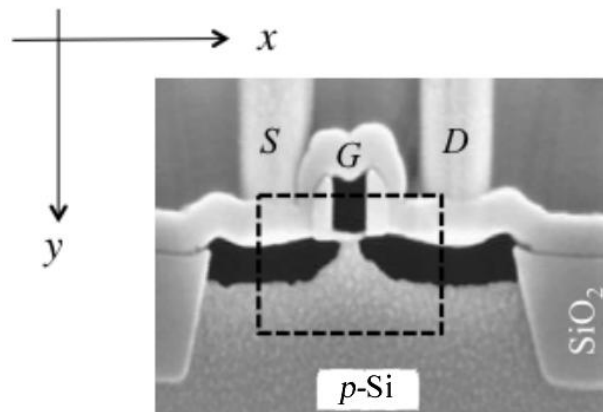


Рис. 1. Микрограмма сечения MOSFET, полученная на сканирующем электронном микроскопе (Texas Instruments, 1997), на которой хорошо видны все контакты и диэлектрическая прослойка, изолирующая затвор от канала проводимости.

Учет 2D электростатики, особенно в нанотранзисторах, влечет серьезные последствия для режимов работы MOSFET. Как показано на рис. 2 слева, приложение большого напряжения на стоке сдвигает характеристику  $\log_{10} I_{DS} \propto V_{GS}$  влево.

Этот сдвиг на рис. 2 связан с  $DIBL = -\Delta V_{GS} / \Delta V_{DS}$  (рис. 11/гл. 1), где  $\Delta V_{GS}$  есть такое изменение напряжения на затворе, которое необходимо, чтобы поддерживать ток стока неизменным при изменении напряжения на стоке на величину  $\Delta V_{DS}$ .

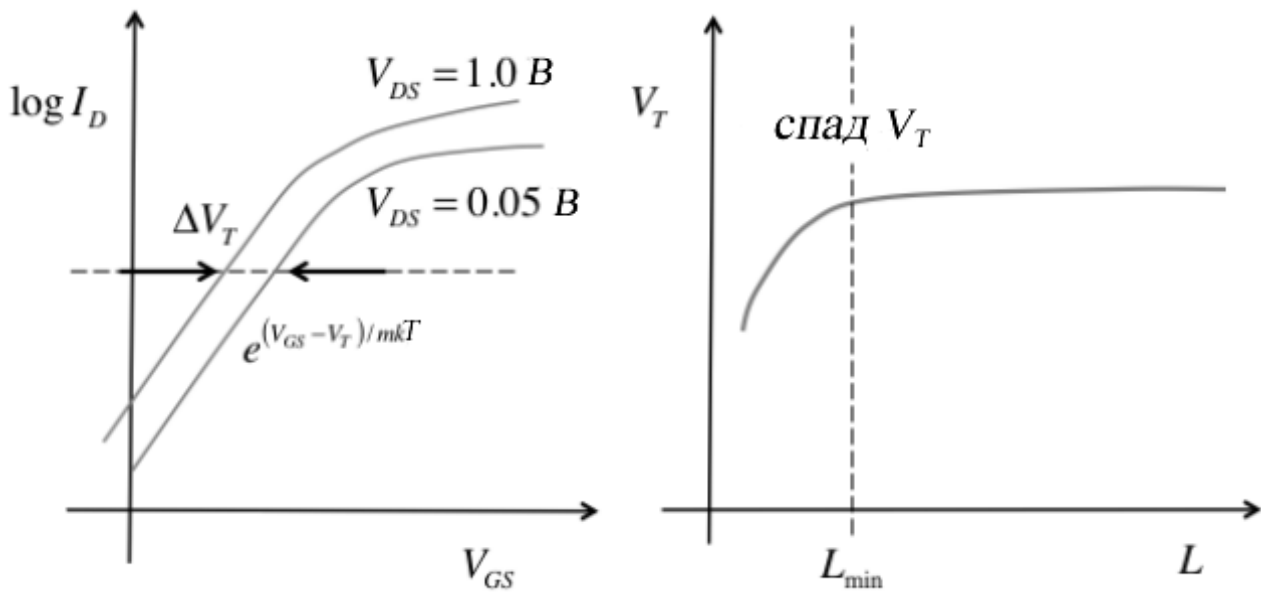


Рис. 2. Иллюстративные примеры того, как учет 2D электростатики сказывается на характеристиках короткоканальных транзисторов.

Слева: *DIBL*, сдвигающее передаточные характеристики  $\log_{10} I_{DS} \propto V_{GS}$ ; такое поведение может быть также интерпретировано как уменьшение порогового напряжения  $V_T$  с возрастанием напряжения на стоке.

Справа: спад  $V_T$  как результат снижения порогового напряжения у нанотранзисторов.

Если взять небольшой ток, соответствующий транзистору в режиме «ON» (пунктирная линия на рис. 2), то мы увидим, что большое напряжение на стоке уменьшает величину порогового напряжения на  $\Delta V_T$ . Однако, полученное ранее выражение для  $V_T$  (13/гл. 3) не содержит зависимости от напряжения на стоке, и причина этого кроется в неучете 2D электростатики. Другое проявление 2D электростатики состоит в зависимости  $V_T$  от длины канала проводимости (рис. 2/справа). Выходное сопротивление транзистора также связано с 2D электростатикой. Эти и другие вопросы, связанные с 2D электростатикой, обсудим в этой главе.

#### 4.2. 2D электростатика MOS

Учесть 2D электростатику можно путем численного решения двумерного уравнения Пуассона, а если речь пойдет о транзисторе с очень коротким каналом, то вынуждены будем решать и трехмерное уравнение Пуассона. Численное моделирование неизбежно в проектировании промышленного производства современных нанотранзисторов. Наша же задача все-таки состоит

в том, чтобы построить качественно правильную физическую картину процессов в MOSFET.

## 2D Уравнение Пуассона

Закон Гаусса для электрического смещения

$$\vec{\nabla} \cdot \vec{D}(x, y) = \rho(x, y), \quad (1)$$

где  $\vec{D}$  – вектор смещения,  $\rho$  – зарядовая плотность. Связь с электрическим полем дается соотношением

$$\vec{D}(x, y) = \varepsilon_s \vec{E}(x, y), \quad (2)$$

где  $\varepsilon_s$  – диэлектрическая проницаемость полупроводника или оксидной пленки, которые предполагаются пространственно однородными. Электростатический потенциал связан с электрическим полем:

$$\vec{E}(x, y) = -\vec{\nabla} \psi(x, y), \quad (3)$$

так что уравнение Пуассона в окончательном виде

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = -\frac{\rho(x, y)}{\varepsilon_s}. \quad (4)$$

Нас особенно интересует решение уравнения Пуассона в подпороговой области и в самом начале режима инверсии, когда учет 2D электростатики позволяет понять природу эффекта *DIBL* и спада  $V_T$ . В подпороговом режиме

$$\rho(x, y) \approx q[N_D^+(x, y) - N_A^-(x, y)] \approx -qN_A, \quad (5)$$

где последнее соотношение следует из предположения, что учитываются в канале проводимости только  $p$ -допанты и что они полностью ионизированы и однородно распределены.

Слой диэлектрика и электрод затвора являются частью канала проводимости и должны быть включены в рассмотрение при вычислении  $\psi(x, y)$ . У слоя диэлектрика другая диэлектрическая проницаемость и электронным зарядом в этом слое обычно пренебрегают, так что уравнением Лапласа для этого слоя будет

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0. \quad (6)$$

В общем случае нужно численно решать систему уравнений (4) и (6) для нахождения электростатического потенциала. Мы же далее обсудим некоторые качественные подходы к пониманию решений этих уравнений.

Особое внимание далее мы уделим короткоканальным транзисторам, эффекты 2D электростатики в которых существенны. В случае достаточно длинного канала проводимости потенциал меняется медленно вдоль направления канала, так что

$$\frac{\partial^2 \psi}{\partial x^2} \ll \frac{\partial^2 \psi}{\partial y^2} \quad (7)$$

и уравнение (4) сводится к одномерному уравнению Пуассона типа (57/гл. 2). Значительная и существенная часть современной традиционной модели MOSFET построена в предположении (7). Этот подход известен как плавноканальное приближение (gradual channel approximation). Стандартный подход к моделированию короткоканальных MOSFET сводится к разработке модели для достаточно длинного канала проводимости с дальнейшим учетом эффектов, вызванных 2D электростатикой [1 – 4].

#### *Спад порогового напряжения и DIBL*

Перепишем уравнение (4) с учетом (5) для режима истощения в виде

$$\frac{\partial^2 \psi}{\partial y^2} = \frac{qN_A}{\epsilon_s} - \frac{\partial^2 \psi}{\partial x^2}. \quad (8)$$

Для  $n$ -MOSFET электростатический потенциал растет от истока к стоку, так что  $d\psi/dx > 0$ . Электрическое поле,  $-d\psi/dx$ , также увеличивается от истока к стоку, так что кривизна потенциала  $d^2\psi/dx^2$  положительная. Это также следует из результатов моделирования нанотранзисторов, представленных на рисунках 16 и 17/гл. 1: как при низком, так и при высоком напряжении на стоке зависимость  $E_c(x)$  имеет отрицательную кривизну, так что кривизна электростатического потенциала положительная.

Учитывая положительную кривизну  $\psi(x)$ , перепишем (8) в виде

$$\frac{\partial^2 \psi}{\partial y^2} = \frac{qN_A|_{eff}}{\epsilon_s}, \quad (9)$$

где

$$N_A|_{eff} = \frac{qN_A}{\epsilon_s} - \frac{\partial^2 \psi}{\partial x^2} < N_A. \quad (10)$$

Уравнение (9) есть одномерное уравнение Пуассона для электростатического потенциала вглубь полупроводниковой подложки с эффективной плотностью допирования, меньшей фактической плотности допирования. Согласно (13/гл. 3), пороговое напряжение связано с плотностью допирования следующим образом:

$$V_T = V_{FB} + \frac{\sqrt{2q\epsilon_s N_A (2\psi_B)}}{C_{ox}} + 2\psi_B. \quad (11)$$

Поскольку учет 2D электростатики эффективно уменьшает  $N_A$  (10), можно ожидать понижения порогового напряжения. При укорочении длины канала кривизна потенциала  $d^2\psi/dx^2$  увеличивается, что уменьшает эффективную плотность допирования и понижает пороговое напряжение. В этом состоит объяснение почему  $V_T$  уменьшается с укорочением канала проводимости (спад  $V_T$ ). Прибегая к тем же самым аргументам, можно понять явление *DIBL* и уменьшение  $V_T$  с увеличением напряжения на стоке при фиксированной длине канала: с ростом  $V_{DS}$  кривизна потенциала  $d^2\psi/dx^2$  увеличивается, эффективная плотность допирования  $N_A|_{eff}$  и  $V_T$  уменьшаются.

Учет 2D электростатики проявляется также в следующем. Барьер между истоком и каналом критически важен для функционирования транзистора. В идеале высота этого энергетического барьера полностью контролируется напряжением на затворе, а напряжение на стоке никак на нее не влияет (рис. 3/верх).

В реальном транзисторе потенциал стока достаёт таки до вершины барьера и понижает его (рис. 3/низ). Через более низкий барьер проходит больше тока при том же напряжении на затворе. Можно сказать иначе: требуется меньшее напряжение на затворе для достижения нужного тока, поскольку барьер был понижен в результате влияния обоих потенциалов – и затвора и стока. Этот эффект *DIBL* сдвигает передаточную характеристику  $\log_{10} I_{DS} \propto V_{GS}$  влево (рис. 2).

Этот подход с понижением барьера также позволяет объяснить почему учет 2D электростатики понижает эффективное допирование. Для сильно допированных подложек зоны труднее изогнуть, а вот при учете 2D электростатики сток помогает затвору барьер понизить. Протяженность области канала проводимости, находящейся под влиянием стока, зависит от 2D геометрии транзистора, к обсуждению которой мы переходим.



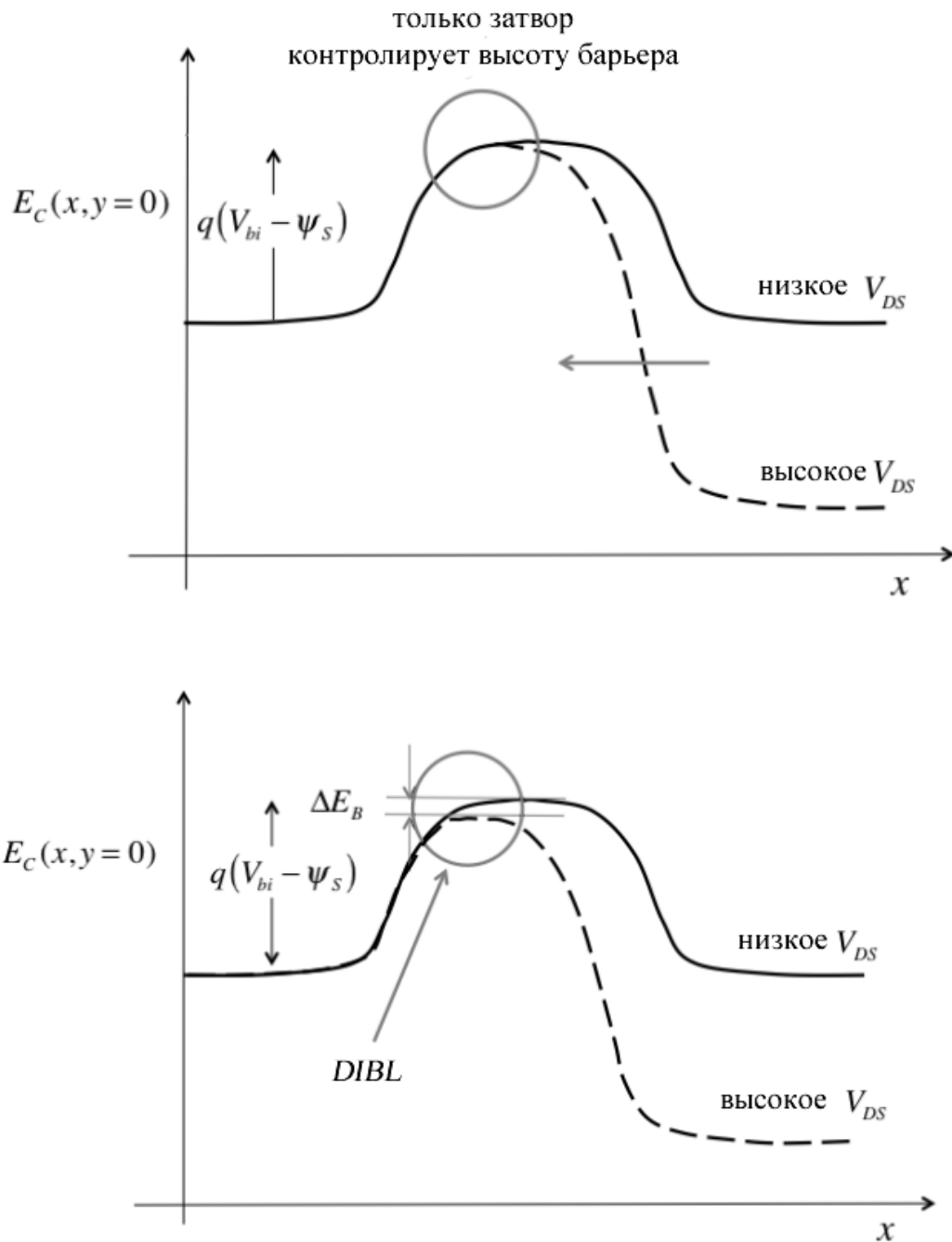


Рис. 3. Иллюстрация влияния напряжения на стоке  $V_{DS}$  на высоту барьера.

Верх: напряжение  $V_{DS}$  не влияет на высоту барьера.

Низ: напряжение  $V_{DS}$  немного понизило высоту барьера (эффект *DIBL*).

### Геометрическое экранирование

Экранирование зарядов характерно для металлов и полупроводников. Если в какой-либо среде образовался «эффективный заряд», подвижные носители тока перестраиваются таким образом, чтобы нейтрализовать (экранировать)

этот заряд. Характерное расстояние, на котором заряд экранируется, называют длиной экранирования, или иначе, дебаевской длиной  $L_D$ , а именно:

$$L_D = \sqrt{\frac{\epsilon_s kT}{q^2 n_0}}, \quad (12)$$

где  $n_0$  – электронная плотность.

В MOSFET реализуется иной вариант экранирования электрического поля. Рисунок 4 иллюстрирует так называемое «геометрическое экранирование».

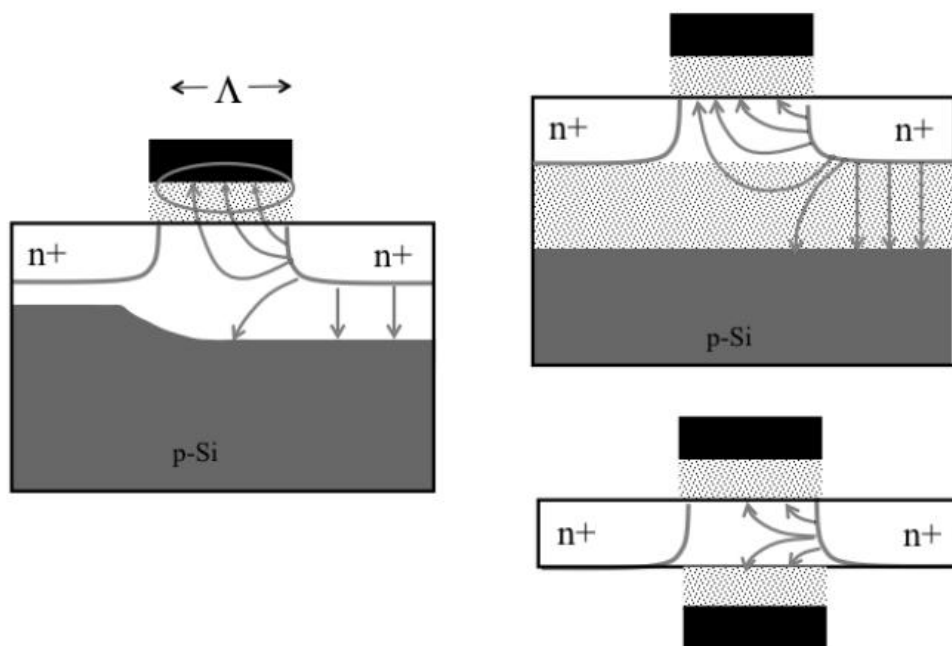


Рис. 4. Геометрическое экранирование в трех типах MOSFET.

Слева: массивная структура MOSFET. Справа наверху: однозатворный SG ETSOI. Справа внизу: DG ETSOI. В MOSFET высокого качества силовые линии электрического поля от стока проникают лишь на расстояние  $\approx \Lambda$  в канал проводимости, поскольку большая часть линий закорачивается на верхнем и нижнем затворах [8].

Показаны три структуры MOSFET. Для DG ETSOI силовые линии электрического поля стока закорачиваются на верхнем и нижнем электродах затворов, проникая в канал лишь на расстояние  $\Lambda$ . Если величина  $\Lambda$  меньше длины канала проводимости  $L$ ,  $\Lambda < L$ , электрическое поле стока не может «достать» начала канала проводимости и поэтому не может понизить барьер. В этом случае *DIBL*, если и есть, то очень небольшой. Точное значение длины геометрического экранирования определяется 2D геометрией транзистора, но интуитивно понятно, что чем плотнее окружен канал затворами, тем более

эффективным будет геометрическое экранирование. На рис. 4 наиболее сильное геометрическое экранирование реализуется в структуре DG ETSOI (наименьшее  $\Lambda$ ) и поэтому ETSOI с двойным затвором меньше всего подвержено эффектам, связанным с 2D электростатикой.

Хотя расчет величины  $\Lambda$  для произвольной геометрии затворов транзистора может оказаться весьма сложным [6, 7], эвристические соображения подсказывают отчего же может зависеть величина  $\Lambda$ . Прежде всего обратимся к одномерному уравнению Пуассона вглубь кремниевой подложки

$$\frac{d^2\psi}{dy^2} = \frac{qN_A}{\epsilon_s}. \quad (13)$$

Это уравнение можно переписать иначе, а именно:

$$\frac{d^2\psi}{dy^2} \approx \frac{V_G - \psi_s}{\Lambda^2}, \quad (14)$$

имея просто в виду, что если  $V_G > \psi_s$ , то  $d^2\psi/dy^2$  будет положительным, а коэффициент пропорциональности выбран из соображений размерности. Приравнявая эти оба уравнения, находим, что

$$\frac{V_G - \psi_s}{\Lambda^2} = \frac{qN_A}{\epsilon_s}. \quad (15)$$

Вспользуемся известным решением в режиме истощения

$$V_G = -\frac{Q(\psi_s)}{C_{ox}} + \psi_s = \frac{qN_A W_D}{C_{ox}} + \psi_s, \quad (16)$$

где  $W_D$  есть толщина слоя истощения. Из этих двух уравнений с использованием (5/гл. 3) находим

$$\Lambda = \sqrt{\frac{\epsilon_s}{\epsilon_{ox}} W_D t_{ox}}. \quad (17)$$

Из (8) и (4) получаем одномерное уравнение Пуассона вдоль канала проводимости:

$$\frac{d^2\psi_s(x)}{dx^2} = \frac{\psi_s - V_G}{\Lambda^2} + \frac{qN_A}{\epsilon_s}, \quad (18)$$

в котором мы сразу определились, что ищем поверхностный потенциал вдоль канала проводимости  $\psi_S(x) = \psi_S(x)|_{y=0}$ . Заменой переменных

$$\phi = \psi_S - V_G + \frac{qN_A}{\epsilon_s} \Lambda^2 \quad (19)$$

получаем простое дифференциальное уравнение

$$\frac{d^2\phi}{dx^2} - \frac{\phi}{\Lambda^2} = 0, \quad (20)$$

в решения которого  $\phi = \exp(\pm x/\Lambda)$  входит величина  $\Lambda$ , задаваемая по (17).

Итак, мы имеем характеристическую длину геометрического экранирования  $\Lambda$ , на протяжении которой возмущение, вызванное потенциалом стока, затухает. Если длина канала проводимости  $L > \Lambda$ , то такие эффекты как *DIBL* в короткоканальных транзисторах будут умеренными. Величины  $L \approx (1.5 - 2)\Lambda$  приемлемы для нынешних короткоканальных MOSFET. Как следует из (17), тонкие диэлектрические слои предпочтительны, так же как и тонкие области истощения. Как показано на рис. 4, в этих случаях силовые линии электрического поля стока скорее всего закоротятся либо на затворе, либо на незаряженной области подложки вместо того, чтобы достичь начала канала проводимости и понизить, пусть даже немного, барьер.

Итак, выше предложена эвристическая оценка характеристической длины геометрического экранирования  $\Lambda$ . Известны более строгие оценки  $\Lambda$  для различных структур MOS [6 – 8]. В целом,  $\Lambda_{BULK} > \Lambda_{DG\ ET\ SOI}$ . Транзистор тем лучше, чем меньше длина геометрического экранирования. Общая закономерность такова, что чем плотнее канал проводимости окружен проводящими электродами, в частности, затворными, тем меньше длина геометрического экранирования.

## 2D Электростатика: емкостная модель

На рис. 5 показан полезный метод учета 2D электростатики. Каждый конденсатор на этом рисунке моделирует электростатическую связь всех четырех терминалов MOSFET с вершиной барьера, виртуальным истоком  $V_S$ . Сама вершина барьера располагается вблизи середины канала проводимости при низких значениях  $V_{DS}$  и сдвигается к истоку по мере увеличения напряжения на стоке, так что параметры конденсаторов зависят от напряжения на стоке [9]. Нужно решать 2D уравнение Пуассона при заданной геометрии

MOSFET для оценки ёмкости каждого конденсатора, однако, емкостной анализ сети конденсаторов полезен сам по себе и приводит к некоторым полезным общим результатам. Рисунок 5 предназначен для массивной структуры MOS. Аналогичные сети конденсаторов могут быть предложены для любых других структур MOS, таких как SG и DG SOI MOSFET [9].

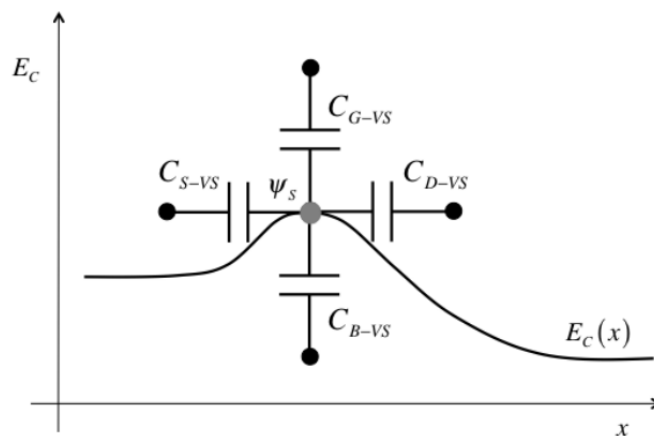


Рис. 5. Емкостная модель 2D электростатики для массивной структуры MOS. Каждый конденсатор моделирует электростатическую связь электродов истока, стока, затвора и подложки с вершиной энергетического барьера.

Для анализа простой сети конденсаторов на этом рисунке мы воспользуемся принципом суперпозиции конденсаторов и вначале предположим, что напряжение на терминалы не подается, однако, заряд на вершине барьера ненулевой ( $Q_s \neq 0$ ). Соответствующая этой ситуации сеть показана на рис. 6/слева.

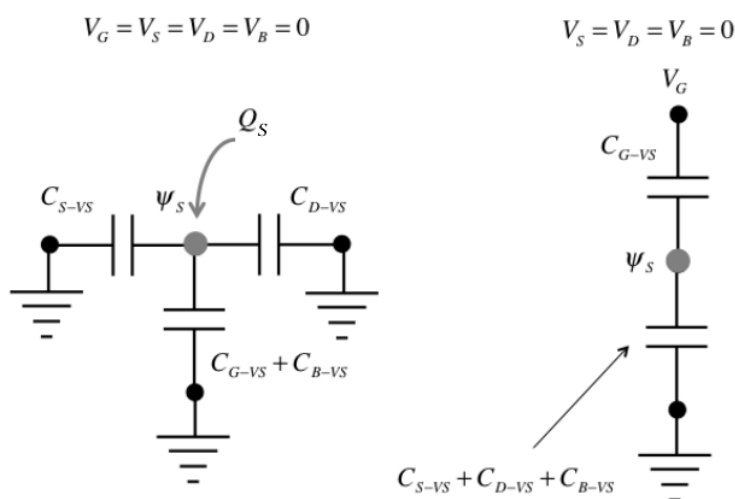


Рис. 6. Упрощенные сети конденсаторов без подачи напряжения.

Слева: с зарядом  $Q_s$  на виртуальном истоке (VS).

Справа: с нулевым зарядом на VS и ненулевым напряжением на затворе.

Суммарная ёмкость на VS есть

$$C_{\Sigma} = C_{G-VS} + C_{S-VS} + C_{D-VS} + C_{B-VS}, \quad (21)$$

а соответствующий потенциал в локации VS есть

$$\psi_S = \frac{Q_S}{C_{\Sigma}}. \quad (22)$$

Теперь предположим, что напряжение на затвор подается, но остальные терминалы заземлены. Соответствующая сеть показана на рис. 6. Вклад в потенциал в локации VS теперь

$$\psi_S = \left( \frac{C_{G-VS}}{C_{\Sigma}} \right) V_G; \quad (23)$$

аналогичные вклады получаем от всех других терминалов: подайте на нужный терминал напряжение, а остальные терминалы заземлите. После суммирования всех вкладов от четырех терминалов и потенциала от ненулевого заряда при нулевых напряжениях на всех терминалах имеем

$$\psi_S = \left( \frac{C_{G-VS}}{C_{\Sigma}} \right) V_G + \left( \frac{C_{S-VS}}{C_{\Sigma}} \right) V_S + \left( \frac{C_{D-VS}}{C_{\Sigma}} \right) V_D + \left( \frac{C_{B-VS}}{C_{\Sigma}} \right) V_B + \frac{Q_S}{C_{\Sigma}}. \quad (24)$$

Это уравнение сравним с аналогичным 1D результатом (11/гл. 3), из которого следует, что

$$\psi_S = V_G + \frac{Q_S}{C_{ox}}. \quad (25)$$

Можно показать, что 2D результат (24) сводится к 1D (25) в том случае, если ёмкость оксидного слоя затвора намного больше ёмкостей остальных терминалов. В этом случае потенциал на вершине барьера полностью контролируется напряжением на затворе, а остальные терминалы не оказывают влияния. Именно к такой ситуации стремятся все разработчики транзисторов: либо создают как можно бóльшую ёмкость затвора, утончая слой диэлектрика или подбирая диэлектрик с как можно более высокой диэлектрической проницаемостью, или максимально, насколько это возможно, уменьшают ёмкость остальных терминалов путем геометрического экранирования с целью электростатически изолировать остальные терминалы от влияния на виртуальный источник на вершине барьера.

Рассмотрим тот случай, когда прилагается напряжение только от затвора и стока, а остальные терминалы заземлены. В подпороговом режиме, когда заряд пренебрежимо мал, уравнение (24) упрощается до

$$\psi_S = \left( \frac{C_{G-VS}}{C_\Sigma} \right) V_G + \left( \frac{C_{D-VS}}{C_\Sigma} \right) V_D. \quad (26)$$

Только затвор и сток влияют на потенциал в локации VS:

$$\frac{\partial \psi_S}{\partial V_G} = \frac{C_{G-VS}}{C_\Sigma}, \quad \frac{\partial \psi_S}{\partial V_D} = \frac{C_{D-VS}}{C_\Sigma}. \quad (27)$$

Транзистор будет высокого качества, если контроль со стороны затвора намного эффективнее, чем со стороны стока:  $\partial \psi_S / \partial V_G \gg \partial \psi_S / \partial V_D$ . Также стремятся к тому, чтобы изменения потенциала соответствовали изменениям напряжения на затворе:  $\partial \psi_S / \partial V_G \approx 1$ . Таким образом, критерии транзистора высокого качества сводятся к выполнению следующих условий:

$$\begin{aligned} C_{G-VS} &\gg C_{D-VS}, \\ C_{G-VS} &\approx C_\Sigma. \end{aligned} \quad (28)$$

Утончение диэлектрического слоя увеличивает  $C_{G-VS}$ , а геометрическое экранирование уменьшает  $C_{D-VS}$ .

Конденсаторы в эквивалентной схеме можно напрямую связать с передаточными характеристиками MOSFET. Ток стока экспоненциально зависит от высоты барьера между истоком и каналом проводимости, например, по (7/гл. 1):

$$I_{DS} \propto e^{-E_{SB}/kT} = e^{q\psi_S/kT}. \quad (29)$$

Перепишем уравнение (26) иначе:

$$\psi_S = \frac{V_G}{m} + \frac{DIBL}{m} V_D, \quad (30)$$

$$m \equiv \frac{C_\Sigma}{C_{G-VS}},$$

где

$$DIBL \equiv \frac{C_{D-VS}}{C_{G-VS}}. \quad (31)$$

Теперь ток стока (29) можно с помощью (30) переписать иначе:

$$I_{DS} \propto e^{q\psi_S/kT} = e^{q(V_G + DIBL \times V_D)/mkT}. \quad (32)$$

Допороговый разброс при постоянном напряжении на стоке, согласно (4/гл. 1), определяется как

$$SS = \left[ \frac{\partial(\log_{10} I_{DS})}{\partial V_{GS}} \right]^{-1} = 2.3m kT \quad (33)$$

и дает такое изменение напряжения на затворе, которое приводит к изменению тока в 10 раз. Допороговый разброс контролируется величиной  $m \geq 1$ , так что  $SS \geq 60 \text{ мВ/декада}$ . Положив  $C_{G-VS} = C_{ox}$  и  $C_{B-VS} = C_D$ , ёмкости полупроводниковой подложки в режиме истощения, из (31) получим

$$m = 1 + \frac{C_D}{C_{ox}} + \frac{C_{S-VS} + C_{D-VS}}{C_{ox}}. \quad (34)$$

Это выражение сравним с уравнением (31/гл. 3) для коэффициента эффекта подложки в режиме истощения, которое было получено для 1D модели электростатики. Первое слагаемое, единица, соответствует идеальному допороговому разбросу. Второе слагаемое ответственно за 1D эффект (31/гл. 3), который обязан падению напряжения на затворе и подложке в режиме истощения. Это слагаемое отсутствует для структуры ETSOI в режиме полного истощения в отличие от массивной структуры MOSFET. Третье слагаемое обязано учету 2D электростатики. Мы видим, что учет 2D электростатики увеличивает коэффициент эффекта подложки и тем самым увеличивается подпороговый наклон. Этот эффект, показанный на рис. 7, нежелателен, и его всегда стараются уменьшить.

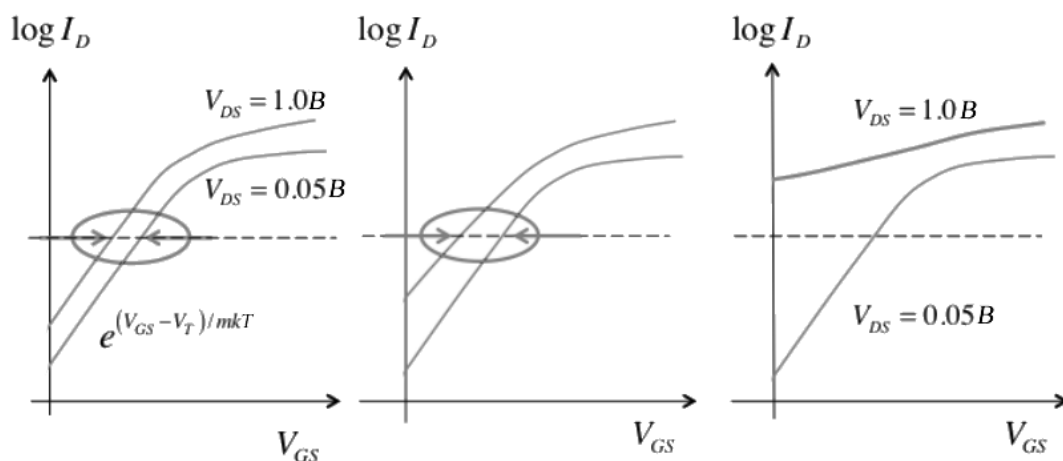


Рис. 7. Иллюстрация происхождения *DIBL* (слева) и увеличения допорогового разброса (по центру), а справа показан также эффект пробоя (punch through), который обсудим ниже.



Отметим также, что емкостная модель 2D электростатики описывает эффект *DIBL*. Согласно (32), если увеличить  $V_D$  на  $\Delta V_D$  с поддержанием тока стока постоянным, то необходимо уменьшить  $V_G$ . Требуемое уменьшение  $V_G$  будет

$$\Delta V_G = -DIBL \times \Delta V_D, \quad (35)$$

что фактически является определением эффекта *DIBL*.

Итак, емкостная модель фактически позволяет качественно понять эффекты, вызываемые строгим, путем численного решения уравнения Пуассона, учетом 2D электростатики. В качественных транзисторах роль ёмкости затвора должна доминировать во влиянии на виртуальный исток (вершину барьера) со стороны других терминалов, эквивалентные ёмкости которых увеличивают допороговый разброс и вызывают *DIBL*.

### 4.3. Масштабирование транзисторов

Последние более чем 50 лет отмечены невиданным прогрессом в области миниатюризации интегральных схем, скейлинга транзисторов и самих электронных устройств. Если скейлинг транзисторов выполнен корректно, производительность интегральных схем улучшается. Выполняя скейлинг транзисторов, основную задачу видят в избавлении у короткоканальных нанотранзисторов от эффектов 2D электростатики.

Глобальная цель скейлинга сформулирована на рис. 8: при уменьшении всех линейных размеров транзистора в  $\kappa$  раз напряжения и токи всех его характеристик также должны уменьшиться в  $\kappa$  раз.

Рисунок 2/справа иллюстрирует ожидаемое изменение порогового напряжения  $V_T$  с уменьшением длины канала проводимости MOSFET. Спад  $V_T$  в короткоканальных транзисторах вызван эффектами 2D электростатики. Ниже некоторой минимальной длины канала  $L_{min}$  напряжение  $V_T$  не только мало, но и слишком чувствительно к длине канала  $L$ . Ниже  $L_{min}$  подпороговый разброс *SS* и *DIBL* становятся неприемлемыми и транзистор может быть пробит. Эффект пробоя транзистора обсудим ниже. Задачу скейлинга в  $\kappa$  раз видят также в том, чтобы уменьшить  $L_{min}$  также в  $\kappa$  раз таким образом, чтобы масштабированный транзистор с длиной  $L = L_{min}/\kappa$  не был подвержен короткоканальным 2D эффектам.

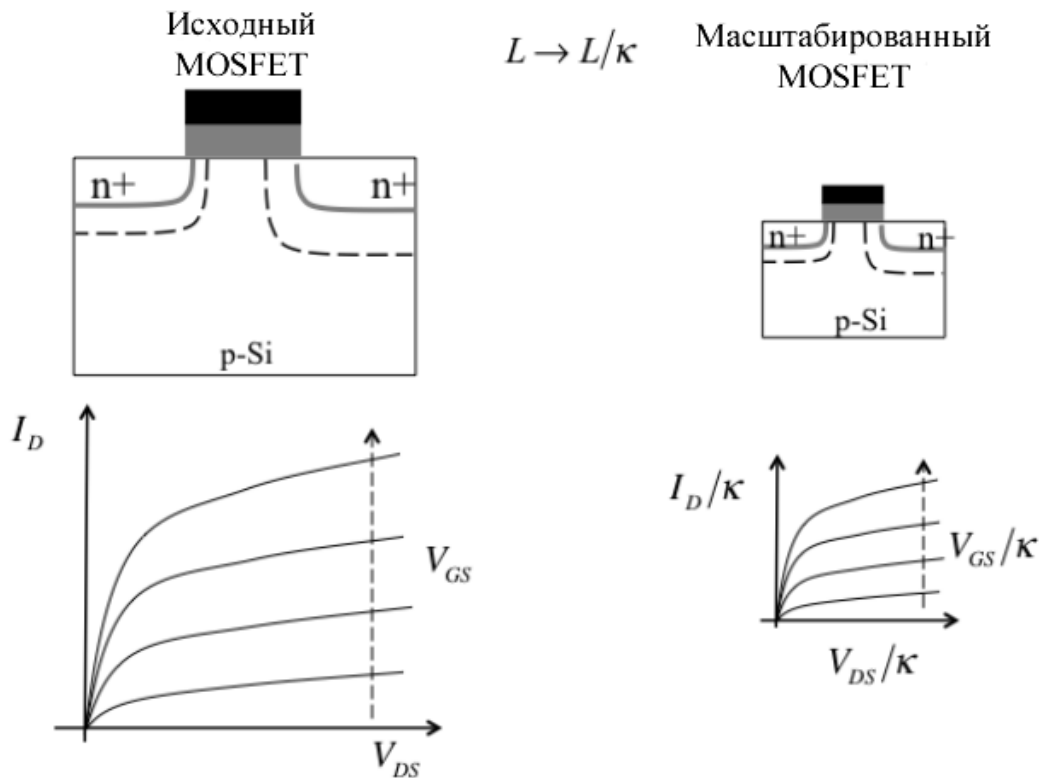


Рис. 8. Скейлинг транзистора. Слева показаны исходный транзистор и набор его характеристик. Справа показан масштабированный транзистор с коэффициентом скейлинга  $\kappa > 1$ . Если масштабирование выполнено корректно, тогда в масштабированных характеристиках все токи и напряжения должны измениться с тем же коэффициентом скейлинга.

Описанный выше подход к масштабированию транзисторов был предложен Робертом Деннардом более 40 лет назад [10] и многие десятилетия служит руководством по скейлингу не только транзисторов, но и интегральных схем и самих конечных электронных устройств. Основная идея скейлинга Деннарда заключается в том, чтобы все размеры уменьшить в  $\kappa$  раз, увеличить допинг в это же число раз и уменьшить напряжение источника питания в  $\kappa$  раз. Такая постановка задачи поддерживает электрическое поле в канале постоянным по ходу скейлинга.

Более подробно скейлинг Деннарда заключается в следующем:

1) Уменьшение всех размеров:

$$\begin{aligned}
 L, W &\Rightarrow (L, W) / \kappa, \\
 t_{ox} &\Rightarrow t_{ox} / \kappa, \\
 W_D &\Rightarrow W_D / \kappa, \\
 y_j &\Rightarrow y_j / \kappa;
 \end{aligned}
 \tag{36}$$

2) Увеличение допирования канала:

$$N_A \Rightarrow \kappa N_A; \quad (37)$$

3) Уменьшение напряжения источника питания:

$$V_{DD} \Rightarrow V_{DD}/\kappa. \quad (38)$$

Величина  $W_D$  это ширина области истощения, а  $y_j$  это глубина  $p$ - $n$ -перехода истока/стока.

Прикинем как все это работает, используя самые простые соображения. Прежде всего, электрическое поле  $E \approx V_{DD}/L$  и поэтому оба  $V_{DD}$  и  $L$  уменьшаются одинаково с тем, чтобы электрическое поле в исходном и в масштабированном канале было одним и тем же.

Низкополевая скорость электронов есть подвижность, помноженная на электрическое поле. Предполагая, что подвижность не изменяется при скейлинге, скорость носителей тока также не изменится. Деннард предположил, что высокополевая скорость совпадает с  $v_{sat}$ , которая обычно служит материальным параметром, не изменяемым при масштабировании. В итоге, скорость носителей при скейлинге не меняется.

Важно масштабировать толщину области истощения

$$W_D = \sqrt{\frac{2\varepsilon_s}{qN_A} (V_{bi} + V_{DD})},$$

где  $V_{bi}$  – встроенный потенциал  $p$ - $n$ -переходов. Если  $V_{DD} \gg V_{bi}$ , то рост допинга в  $\kappa$  раз и уменьшение  $V_{DD}$  в это же число раз приводит к уменьшению ширины области истощения в  $\kappa$  раз. Если  $t_{ox}$  и  $y_j$  так же масштабировать, то эффекты 2D электростатики в масштабированном транзисторе станут проявляться сильнее при длине канала проводимости в  $\kappa$  раз короче, чем в исходном транзисторе. В результате величина  $L_{min}$  уменьшится приблизительно в  $\kappa$  раз.

Ёмкость

$$C = \frac{\varepsilon A}{t}, \quad [\Phi]$$

где  $t$  – толщина оксидного слоя или же области истощения. Поскольку все толщины уменьшаются в  $\kappa$  раз, а площадь  $A$  – в  $\kappa^2$  раз, все ёмкости

уменьшатся в  $\kappa$  раз, однако,  $C_{ox}$ , как ёмкость оксидного слоя на единицу поверхности, увеличится в  $\kappa$  раз.

Рассмотрим влияние скейлинга Деннарда на некоторые важные величины. Заряд инверсионного слоя

$$Q = -C_{ox}(V_G - V_T).$$

Поскольку  $C_{ox}$  растёт в  $\kappa$  раз, а напряжения уменьшаются в это же число раз, то заряд инверсионного слоя на единицу поверхности не изменяется при масштабировании.

Теперь рассмотрим ток

$$I_{DS} = WQv.$$

Поскольку заряд и скорость не меняются при масштабировании, а ширина уменьшается в  $\kappa$  раз, то и ток уменьшится в  $\kappa$  раз.

Подводя итоги, приходим к следующему общему выводу из скейлинга Деннарда при сохранении постоянства электрического поля:

$$\begin{aligned} Q &\Rightarrow Q, \\ v &\Rightarrow v, \\ C &\Rightarrow C/\kappa, \\ C_{ox} &\Rightarrow \kappa C_{ox}, \\ I_{DS} &\Rightarrow I_{DS}/\kappa. \end{aligned} \tag{39}$$

Пользуясь скейлингом Деннарда, проверим работоспособность масштабированных сетей. Задержка сети – это время, необходимое для удаления заряда  $C \cdot V_{DD}$ , накопленного сетевыми ёмкостями,

$$\tau = \frac{C \cdot V_{DD}}{I_{DS}}.$$

Как видим, задержка сети  $\tau$  уменьшается в  $\kappa$  раз. Мощность  $P_D = V_{DD}I_{DS}$  уменьшается в  $\kappa^2$  раз. Плотность мощности в единицах  $Вт/м^2$  не изменяется при скейлинге. Поскольку размер каждого транзистора уменьшается в  $\kappa^2$  раз, то плотность упаковки транзисторов  $D$  увеличивается в  $\kappa^2$  раз. Наконец, такая важная метрика как произведение мощности на задержку сети  $P_D \cdot \tau$  уменьшается в  $\kappa^3$  раз.

Итак, суммируем результаты скейлинга Деннарда при сохранении постоянства электрического поля:

$$\begin{aligned}
 \tau &= C \cdot V_{DD} / I_{DS} \Rightarrow \tau / \kappa, \\
 P_D &= V_{DD} I_{DS} \Rightarrow P_D / \kappa^2, \\
 P_D / A &\Rightarrow P_D / A, \\
 D &= \frac{N_{FET}}{A} \Rightarrow D \times \kappa^2, \\
 P_D \cdot \tau &= C \cdot V_{DD}^2 \Rightarrow P_D \cdot \tau / \kappa^3.
 \end{aligned}
 \tag{40}$$

Соблюсти скейлинг по Деннарду не так легко, как кажется на первый взгляд, поскольку некоторые величины не масштабируются. Например, вспомните уравнение для порогового напряжения  $V_T$  (11). Напряжение плоских зон  $V_{FB}$  не масштабируется, а потенциал  $\psi_B$  относительно нечувствительный к скейлингу, так что в соответствии со сценарием Деннарда  $V_T \Rightarrow V_T / \kappa$ . Учтем, однако, что нередко применяются нетрадиционные режимы допинга [4], что может сказаться на результатах скейлинга.

Стоковая область истощения зависит от  $\sqrt{(V_{bi} + V_{DD}) / N_A}$ . Поскольку встроенный потенциал  $V_{bi}$  не масштабируется, то  $W_D \Rightarrow W_D / \kappa$  (36). Подпороговый разброс не чувствителен к скейлингу. Все эти факторы обнадеживают. Вместе с тем, разработчики транзисторов и чипов постоянно совершенствуют скейлинг с целью предложить устройства с еще большей экономией мощности и меньшей задержкой.

В настоящее время масштабирование устройств испытывает серьезные вызовы, а некоторые даже предрекают скорый конец эры скейлинга. Одна из проблем связана с оксидным слоем затвора, поскольку достигнут предел в утончении этого слоя вплоть до и даже меньше 1.5 нм, что соответствует нескольким молекулам  $\text{SiO}_2$  [11]; дальнейшее утончение ведет к чрезмерному протеканию тока. Это обстоятельство вынуждает переходить от плоских MOSFET к FinFET [12], которые обеспечивают лучший электростатический контроль при той же толщине оксидной пленки [13]. Еще один вызов дальнейшему масштабированию транзисторов связан с неудачными попытками скейлинга подпорогового разброса  $SS$ . Задавшись максимально разрешенным током в режиме «OFF», величина  $SS$  должна быть лишь немного больше чем 60 мВ/декада и ток должен расти линейно выше порога, что требует  $V_{DD} \approx 1V$  для достижения нужного значения тока  $I_{DS}$  в режиме «ON». В результате не

удается масштабировать напряжение питания. Предложено несколько новых архитектур транзисторов для решения подобных проблем [14].

#### 4.4. Пробой транзисторов

Усиление  $2D$  эффектов может привести к пробую, т. е. к закорачиванию стока на исток. В этой ситуации ток идет от стока к истоку в отсутствие или при минимальном влиянии напряжения на затворе (рис. 9).

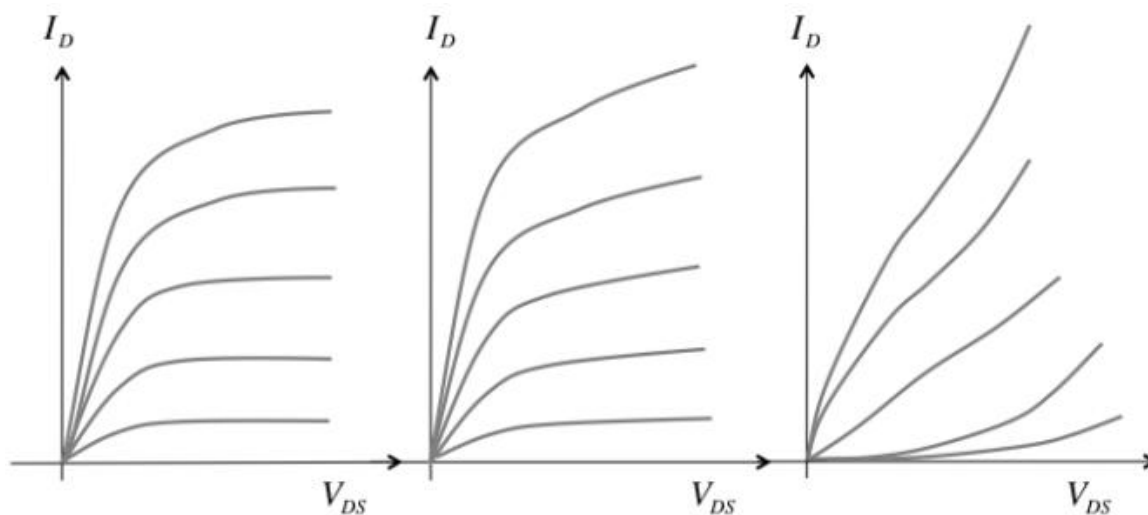


Рис. 9. Иллюстрация влияния  $2D$  эффектов на выходные характеристики MOSFET. Слева: достаточно длинный канал проводимости с очень большим выходным сопротивлением. По центру: короткий канал с низким сопротивлением. Справа: транзистор поврежден пробоем.

На рис. 9/слева влияние эффектов  $2D$  электростатики незначительное: подпороговый разброс лишь чуть больше  $60 \text{ мВ/декада}$  и  $DIBL$  приемлемый. Когда  $2D$  эффекты становятся заметными (рис. 9/по центру), подпороговый разброс существенно возрастает, а  $DIBL$  становится весьма большим, например, более  $100 \text{ мВ/В}$ . При доминировании  $2D$  эффектов (рис. 9/справа) наблюдается резкая деградация функциональности транзистора. Ток лишь слабо зависит от напряжения на затворе, а  $DIBL$  вообще не поддается определению, поскольку подпороговые выходные характеристики при низких и высоких напряжениях на стоке даже не параллельны оси напряжений.

Рисунок 9 показывает также как учет  $2D$  электростатики влияет на выходные характеристики транзисторов. При достаточно длинном канале проводимости ток стока в области насыщения постоянный и выходное сопротивление приближается к бесконечности (рис. 7/гл. 1). Для MOSFET с коротким каналом выходное сопротивление существенно уменьшается.

Причину этого легко понять: ток пропорционален  $(V_{GS} - V_T)$ , а  $V_T$  уменьшается с ростом напряжения на стоке из-за эффекта *DIBL*. Не сразу очевидно, что роль 2D электростатики в подпороговой области, где формируется и измеряется *DIBL* и где заряд в канале пренебрежимо мал, так же важна, как и выше порога, где подвижный заряд в канале велик. Показано, однако, что в качественно собранных транзисторах одни и те же эффекты *DIBL* 2D электростатики наблюдаются как ниже, так и выше порогового напряжения [8, 15]. Наконец, справа на рис. 9 показаны выходные характеристики транзистора с пробоем: напряжение на стоке очень сильно влияет на ток даже в «области насыщения».

Пробой происходит тогда, когда электрическое поле стока покрывает весь канал проводимости вплоть до истока: приблизительно тогда, когда область истощения стока смыкается с областью истощения истока (рис. 10/слева).

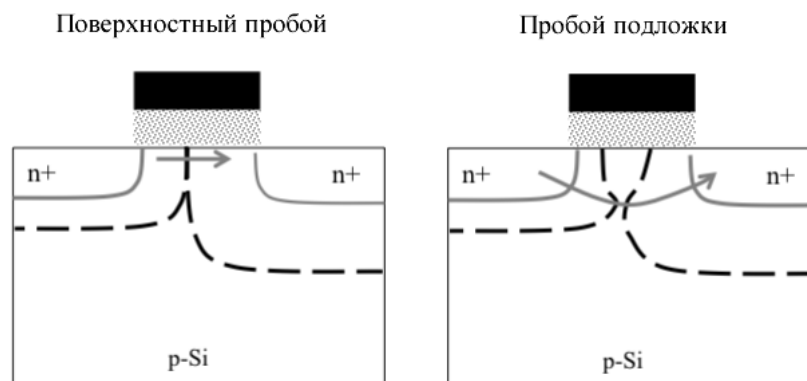


Рис. 10. Сечения MOSFET, иллюстрирующие границы истощения при поверхностном пробое (слева) и при пробое подложки (справа).

Как показано справа на рис. 10, границы областей истощения могут иметь сложные профили как в результате режимов 2D допирования, так и вследствие эффектов 2D электростатики. Итак, перекрытие областей истощения может происходить как вблизи поверхности раздела Si/SiO<sub>2</sub>, так и в подложке глубже, соответственно чему случается поверхностный пробой или же пробой подложки.

Критерий избежать пробоя  $L > W_s + W_d$  служит лишь очень грубой оценкой. Лучшее пояснение дает рис. 11.

Абсолютный пробой имеет место тогда, когда потенциал стока не просто покрывает весь канал проводимости и не просто понижает барьер, а ликвидирует барьер напрочь. В такой ситуации электроны перетекают из истока в сток, не нуждаясь в действии потенциала затвора.

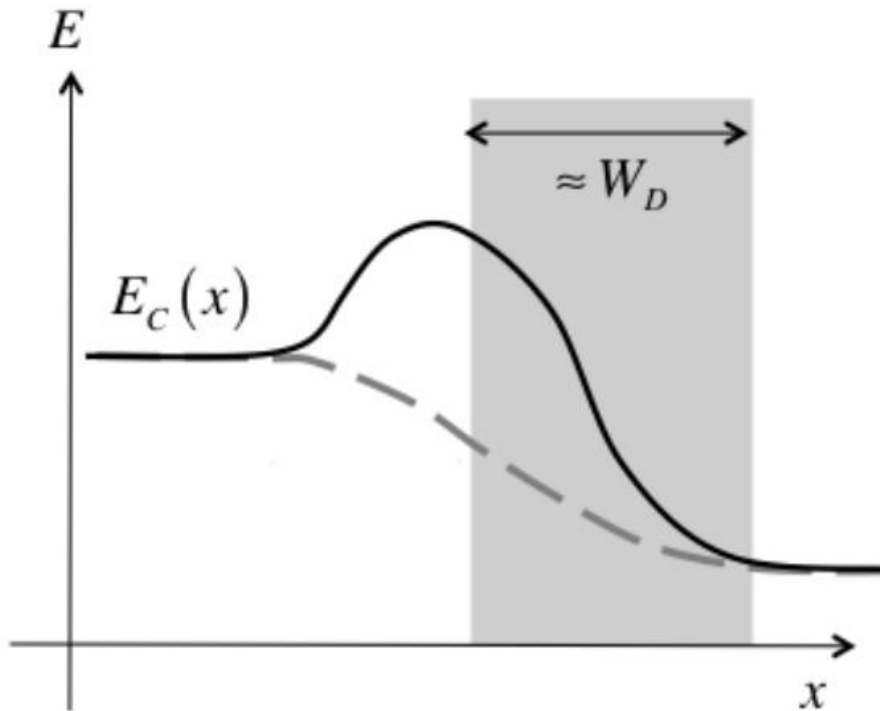


Рис. 11. Иллюстрация пробоя. Сплошная линия соответствует «хорошему» состоянию транзистора, пунктирная – пробоею. Серая зона показывает границы области истощения стока.

Пробой можно определить иначе: как ситуацию, когда контроль над током со стороны стока так же эффективен, как и контроль со стороны затвора. Согласно емкостной модели (27), пробой происходит когда  $C_{G-VS} = C_{D-VS}$ . Реальное напряжение на стоке, вызывающее пробой, можно оценить численно путем решения 2D уравнения Пуассона для транзистора конкретной структуры.

Итак, мы обсудили как 2D электростатика нарушает работу короткоканальных транзисторов. Как же функционирует электростатически хорошо сбалансированный MOSFET (рис. 12)? В таком транзисторе есть область вблизи начала канала проводимости, которая находится под сильным контролем со стороны затвора. В этой области поле  $dE_C/dx$  мало и она включает в себя вершину барьера, «экранированную» от влияния потенциала стока, что позволяет поддерживать незначительным *DIBL*.

Потенциал в районе вершины барьера контролирует высоту барьера, а тем самым и ток стока в транзисторе. В идеале этот потенциал должен определяться только затвором, как в уравнении (25). На практике напряжение на стоке всегда оказывает некоторое влияние на потенциал в области вершины барьера, как например, согласно (26); особенно это касается короткоканальных MOSFET.



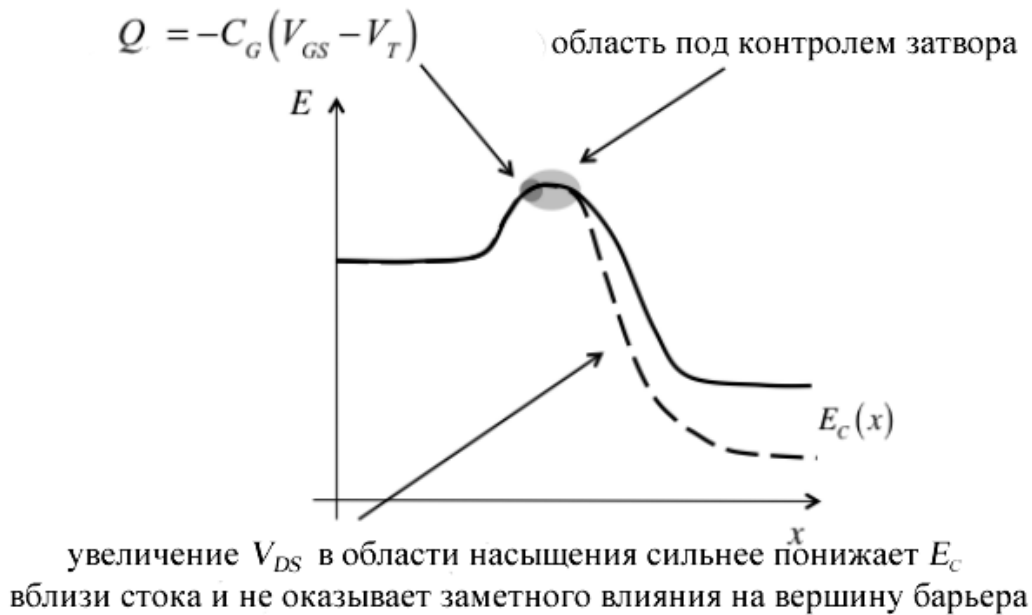


Рис. 12. Иллюстрация того, как работает электростатически хорошо сбалансированный MOSFET при большом напряжении как на стоке, так и на затворе. Как показывает пунктирная линия, рост напряжения на стоке увеличивает потенциал в области насыщения, не затрагивая барьер в начале канала, который находится под контролем затвора.

Глобальная задача, стоящая перед дизайнерами транзисторов, заключается в том, чтобы гарантировать, что заряд инверсионного слоя на вершине барьера подчиняется классическому 1D результату, а именно:

$$Q|_{x=0} = -C_G(V_{GS} - V_T), \quad (41)$$

где  $x=0$  соответствует вершине барьера. Физически корректно предполагать, что именно 1D электростатика MOS, учет которой приводит к (41), применима к вершине барьера, поскольку именно в этой локации кривизна потенциала  $d^2\psi/dx^2 = 0$  и 2D уравнение Пуассона сводится к 1D уравнению. Последнее утверждение не совсем корректно, поскольку при учете 2D электростатики пороговое напряжение оказывается зависящим от напряжения на стоке:

$$V_T = V_{T0} - \delta V_{DS}, \quad (42)$$

где параметр  $\delta$  учитывает *DIBL*.

Ток в MOSFET при высоком напряжении на стоке создается электронами, которые преодолевают барьер, диффундируют через небольшую область за вершиной барьера в начале канала, характеризуемую низким полем, затем

входят в область с высоким полем в стоковой части канала. Узким местом, которое ограничивает ток стока, является эта небольшая область с низким полем. Эта картина функционирования MOSFET фактически подобна тому, как работает биполярный транзистор, в котором роль истока играет эмиттер, низкополевая область в начале канала играет роль базы, а высокополевая область стока это коллектор. Аналогия между MOSFET и биполярным транзистором очень тесная [16].

При низком напряжении на стоке ток пропорционален  $V_{DS}$ , однако, при высоком стоковом напряжении ток в хорошо сбалансированном MOSFET зависит от  $V_{DS}$  в меньшей степени. В транзисторе с длинным каналом ток действительно насыщается. Такое поведение имеет место потому, что сильный контроль затвора экранирует потенциал в районе начала канала от влияния потенциала стока. Увеличение  $V_{DS}$  за пределы напряжения насыщения  $V_{DSAT}$  прежде всего увеличивает потенциал и электрическое поле в районе стокового конца канала проводимости. В хорошо сбалансированном MOSFET напряжение стока лишь несущественно влияет на потенциал в самом начале канала, что в итоге приводит к конечной выходной проводимости транзистора. Хотя и нет причин полагать, что роль 2D электростатики ниже порога такая же, как и выше порога, однако, опыт применения модели виртуального истока для хорошо сбалансированных MOSFET свидетельствует, что это действительно так [15] и численное моделирование подтверждает этот вывод [8].

Изложенная картина влияния электростатики в качественно собранных транзисторах будет использована позже для более глубокого понимания физики нанотранзисторов.

#### 4.5. Модель виртуального истока

Упрощенная модель виртуального истока, модель VS нулевого уровня была изложена в главе 2. Она построена с использованием простых и традиционных соображений в отличие от полноценной модели VS [15], которая была разработана целенаправленно с учетом физики нанотранзисторов. Сейчас мы уточним модель нулевого уровня и изложим исходную модель VS и физические принципы, лежащие в ее основе.

Ток стока MOSFET в общем случае

$$I_{DS} = W |Q|_{x=0}(V_{GS}, V_{DS}) | \langle v_x(V_{GS}, V_{DS}) \rangle |_{x=0}, \quad (43)$$

где  $x=0$  соответствует локации виртуального истока VS на вершине барьера. Ток непрерывен, так что мы вправе вычислять его там, где нам удобнее. В хорошо «темперированном» MOSFET на вершине барьера, как следует из анализа 1D электростатики,  $Q(V_{GS}, V_{DS}) \approx Q(V_{GS})$ . Нужна еще только небольшая поправка на *DIBL*. Далее мы уточним модель VS нулевого уровня, приняв во внимание эффекты электростатики MOS. Что же касается скорости электронов  $\langle v_x(V_{GS}, V_{DS}) \rangle|_{x=0}$  в применении к нанотранзисторам, нам придется отказаться от концепции «сверху – вниз», которую мы до сих пор исповедовали, и воспользоваться обобщенной моделью транспорта электронов Ландауэра – Датта – Лундстрема (ЛДЛ) [17 – 22], единственной полноценно пригодной для описания классического транспорта в наноэлектронике.

Модель VS нулевого уровня исходит из двух отдельных выражений (28/гл. 2) для тока стока в линейной области и в области насыщения:

$$\begin{aligned} I_{DLIN} &= \frac{W}{L} \mu |Q(V_{GS})| V_{DS}, \\ I_{DSAT} &= W v_{sat} Q(V_{GS}), \end{aligned} \quad (44)$$

которые показаны пунктирными прямыми на рис. 13. Реальная выходная характеристика, показанная на этом рисунке сплошной линией, получается плавным сочленением линейной области и области насыщения с помощью выражения для средней скорости, зависящей от напряжения на стоке, в виде

$$\begin{aligned} \langle v_x(V_{DS}) \rangle &= F_{SAT}(V_{DS}) v_{sat}, \\ F_{SAT}(V_{DS}) &= \frac{V_{DS} / V_{DSAT}}{\left[ 1 + (V_{DS} / V_{DSAT})^\beta \right]^{1/\beta}}, \end{aligned} \quad (45)$$

где  $F_{SAT}$  – функция насыщения тока стока, а напряжение насыщения на стоке

$$V_{DSAT} = \frac{v_{sat} L}{\mu}. \quad (46)$$

Ранее заряд на вершине барьера был описан следующим образом:

$$\begin{aligned} Q(V_{GS}) &= 0, \quad V_{GS} \leq V_T, \\ Q(V_{GS}) &= -C_{ox}(V_{GS} - V_T), \quad V_{GS} > V_T, \\ V_T &= V_{T0} - \delta V_{DS}. \end{aligned} \quad (47)$$

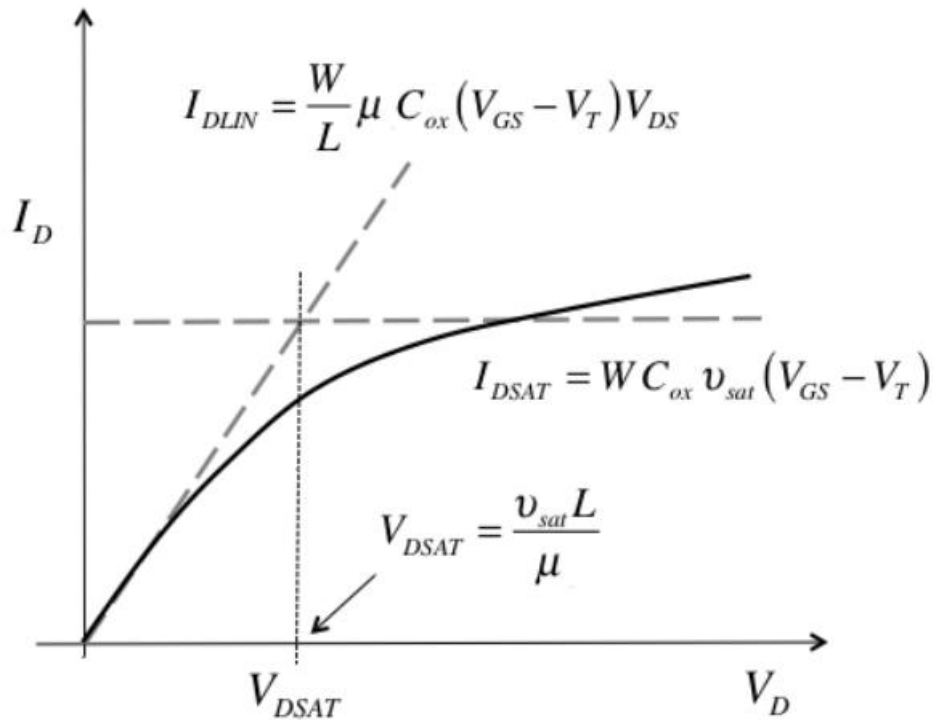


Рис. 13. Выходная характеристика  $n$ -MOSFET, включенного по схеме с общим истоком, при некотором фиксированном значении напряжения на затворе (сплошная линия). Пунктирные прямые дают ток в линейной области и в области насыщения согласно (44).

Мы теперь понимаем, что  $C_{ox}$  в (47) нужно заменить на емкость  $C_G$  (58/гл. 3), как последовательную комбинацию  $C_{ox}$  и ёмкости полупроводника в режиме инверсии  $C_s^{inv}$ , и что  $C_G < C_{ox}$ . Мы также понимаем теперь как описать  $Q(V_{GS}, V_{DS})$  ниже порога и таким образом можно уточнить модель VS, включив как подпороговую проводимость, так и надпороговую.

### Подпороговая область

При напряжении на затворе ниже порогового напряжения MOSFET находится в подпороговом режиме. Рисунок 14 иллюстрирует ход зависимости  $Q(V_{GS}) \propto V_{GS}$  в линейном масштабе и в логарифмическом. Ранее было показано (45/гл. 3), что поверхностная электронная плотность в массивном MOSFET в подпороговом режиме дается выражением

$$Q(V_G) = -(m-1)C_{ox} \frac{kT}{q} e^{q(V_G - V_T)/mkT}. \quad (48)$$

Аналогичный результат (96/гл. 3) был получен для ETSOI, только с  $m = 1$ . Ключевым для обоих случаев является экспоненциальная зависимость  $Q(V_G) \propto \exp[q(V_{GS} - V_T)/mkT]$  в подпороговом режиме.

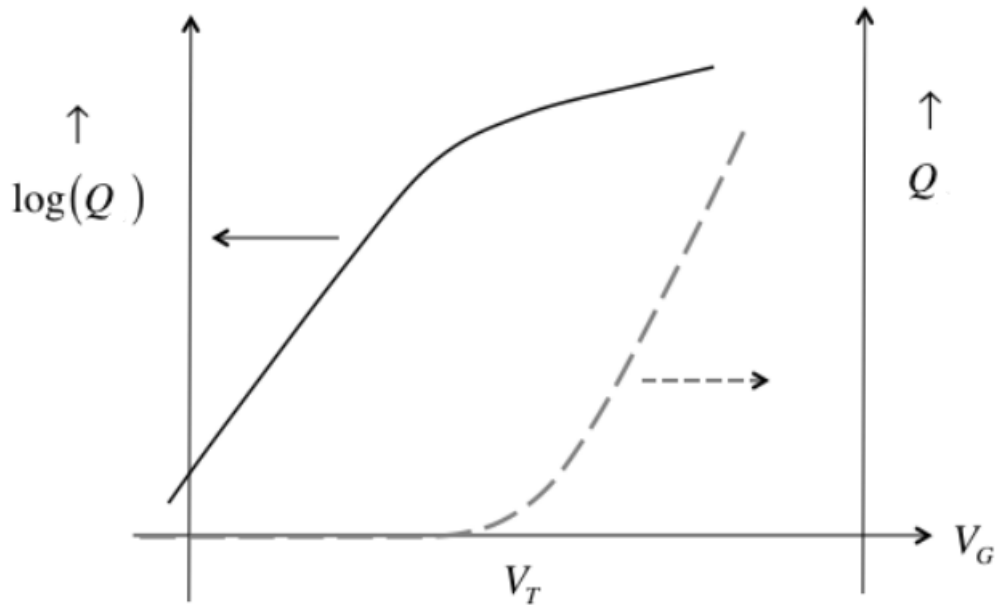


Рис. 14. Зарядовая плотность в инверсионном слое в зависимости от напряжения на затворе (пунктирная кривая справа), а также в логарифмическом масштабе (сплошная кривая слева).

Из (43) и (48) получаем уравнение для тока в массивном MOSFET в подпороговом режиме, а именно:

$$I_{DS} = W(m-1)C_{ox} \frac{kT}{q} e^{q(V_G - V_T)/mkT} \langle v_x \rangle_{x=0}. \quad (49)$$

Вспомним, что согласно емкостной модели (34),

$$m = 1 + \frac{C_\Sigma}{C_{ox}}, \quad (50)$$

где  $C_\Sigma$  есть суммарная ёмкость, влияющая на виртуальный исток VS. В случае массивного MOSFET это сумма ёмкостей всех четырех терминалов (21).

Для тока стока в подпороговом режиме знаем также подпороговый разброс (4/гл. 1)

$$SS = \left[ \frac{\partial(\log_{10} I_{DS})}{\partial V_{GS}} \right]^{-1} = 2.3m \frac{kT}{q}. \quad [B/\text{декада}] \quad (51)$$

Подпороговый разброс обычно приводится в *мВ/декада*;  $SS < 100$  *мВ/декада* считается приемлемым подпороговым разбросом.

Рисунок 15 показывает, почему подпороговый разброс является такой важной метрикой транзисторов.

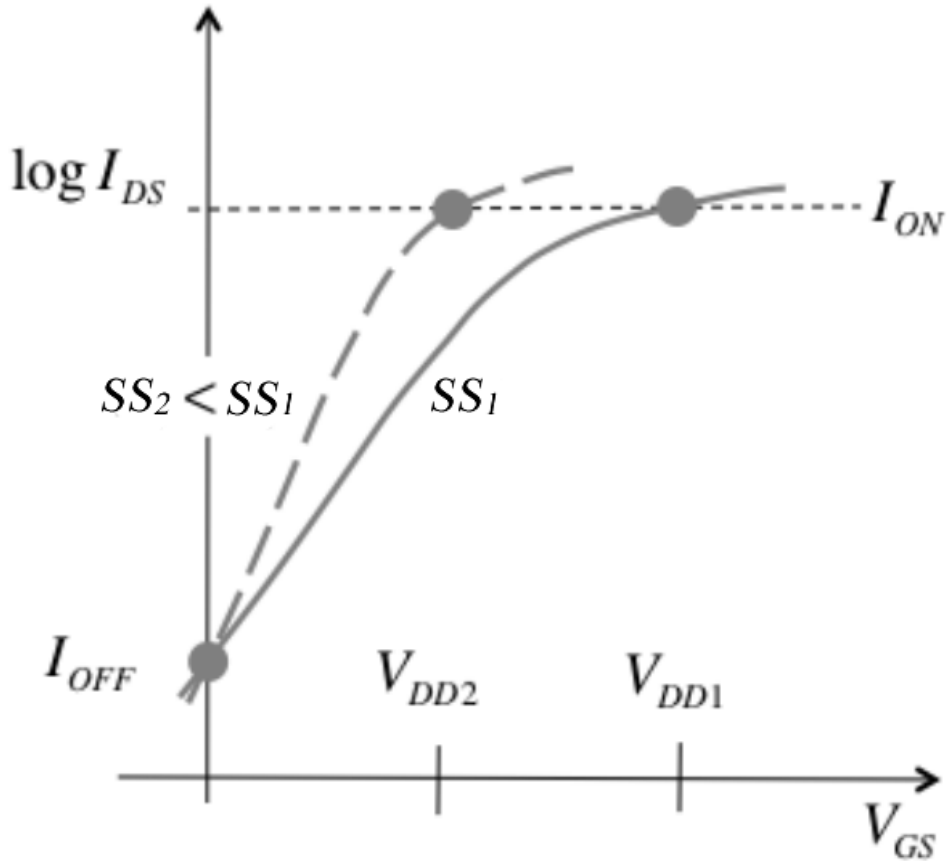


Рис. 15. Иллюстрация того, как подпороговый разброс  $SS$  связан с напряжением источника питания  $V_{DD}$ . Ток в режиме «ON» соответствует максимальному напряжению источника питания, поданному на затвор.

Низкие значения  $I_{OFF}$  предпочтительнее в любых устройствах, поскольку тогда электронная схема устройства не потребляет избыточную мощность в режиме ожидания. Для заданного значения  $I_{OFF}$  метрика  $SS$  определяет то напряжение источника питания  $V_{DD}$ , которое обеспечит нужный ток  $I_{ON}$ . Высокие значения последнего благоприятны для ускоренного функционирования электронной схемы, поскольку все ёмкости схемы могут заряжаться и разряжаться быстрее. У транзистора с меньшим  $SS$  нужное значение тока  $I_{ON}$  достигается при меньшем напряжении источника питания  $V_{DD}$ . Электронная схема будет работать с той же скоростью, но поскольку

мощность пропорциональна  $V_{DD}^2$ , то схема будет диссипировать меньшую мощность. Размещая в настоящее время несколько миллиардов транзисторов на одной интегральной схеме, потребляемая мощность приобретает критическое значение – как активная мощность в рабочем режиме пропорциональная  $V_{DD}^2$ , так и мощность в режиме ожидания, определяемая значением  $I_{OFF}$ .

Согласно (51), наименьший подпороговый разброс  $SS$  при комнатной температуре равен  $60 \text{ мВ/декада}$ . Полностью истощенные транзисторы такие, как ETSOI MOSFET, характеризуются значением  $m=1$  и наиболее благоприятны для достижения наименьших значений  $SS$ . Выше порога ток растет приблизительно линейно с увеличением напряжения на затворе, так что минимальный предел для  $SS$ , равный  $60 \text{ мВ/декада}$ , определяет и нижний предел напряжения источника питания  $V_{DD}$ . На практике сейчас этот нижний предел составляет  $1 \text{ В}$ . Напомним, что скейлинг Деннарда транзисторов при сохранении электрического поля постоянным требует масштабирования  $V_{DD}$  (38) с каждой новой технологической генерацией. Поскольку у подпорогового разброса  $SS$  есть нижний предел (при комнатной температуре!), то мощность источника питания масштабированию уже не поддается. Это обстоятельство и составляет суть критической ныне проблемы диссипации мощности интегральных схем [23].

Итак, у подпорогового разброса есть нижний предел с учетом того, что  $m \geq 1$ , однако, какова его природа, откуда он берется. Это поясняет рис. 16.

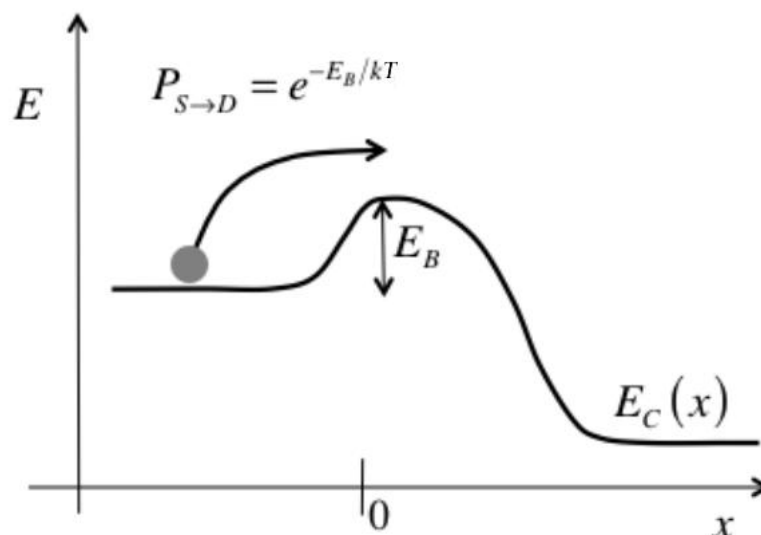


Рис. 16. Термоионный ток в MOSFET.

Ток стока образуют электроны, эмитированные истоком и преодолевшими барьер, которые далее скатываются к стоку. Вероятность этой термоионной эмиссии экспоненциально зависит от высоты барьера  $E_B$ , а именно:

$$P_{S \rightarrow D} = e^{-E_B/kT}. \quad (52)$$

Эта экспоненциальная зависимость влечет за собой экспоненциальную зависимость  $Q \propto V_{GS}$  и ограничивает подпороговый разброс  $SS \geq 60 \text{ мВ/декада}$  при комнатной температуре. Предложены физические подходы, позволяющие преодолеть это ограничение [24, 25].

Остается обсудить среднюю скорость электронов в канале проводимости  $\langle v_x \rangle|_{x=0}$  в подпороговом режиме; приведенное выше выражение (15) применимо лишь выше порога. В подпороговом режиме средняя скорость в локации виртуального истока (вершина барьера) определяется выражением

$$\langle v_x \rangle|_{x=0} = \frac{D}{L} \left( \frac{n_S(0) - n_S(L)}{n_S(0)} \right), \quad (53)$$

где  $D$  есть коэффициент диффузии электронов. Из простой модели термоионной эмиссии следует, что

$$n_S(L) / n_S(0) = e^{-qV_{DS}/kT},$$

так что

$$\langle v_x \rangle|_{x=0} = \frac{D}{L} \left( 1 - e^{-qV_{DS}/kT} \right) = \frac{kT}{q} \frac{\mu}{L} \left( 1 - e^{-qV_{DS}/kT} \right), \quad (54)$$

где было использовано соотношение Эйнштейна между коэффициентом диффузии электронов и их подвижностью. Подставляя (54) в (49), получаем стандартное выражение для тока в подпороговом режиме [4]:

$$I_{DS} = \mu C_{ox} \frac{W}{L} (m-1) \left( \frac{kT}{q} \right)^2 e^{q(V_{GS}-V_T)/mkT} \left( 1 - e^{-qV_{DS}/kT} \right). \quad (55)$$

Отметим еще один момент. Из рис. 16 совсем не очевидно, что электроны должны диффундировать через весь канал проводимости. Представляется физически более корректным, что электроны диффундируют через низкополевую область канала, а затем более сильное электрическое поле быстро пронесит их через оставшуюся часть канала. В соответствии с такой



моделью длину канала проводимости  $L$  надо бы заменить на меньшее значение  $l < L$ . Исходя даже из этого обстоятельства, на практике затруднительно определить предэкспоненциальный фактор, так что уравнение (55) обеспечивает удовлетворительное описание реальных транзисторов в целом [4].

*Из под порога в область выше порога*

Уравнение (48) дает зависимость  $Q(V_{GS})$  ниже порога, а в режиме сильной инверсии, согласно (58/гл.3),  $Q(V_{GS}) = -C_G^{inv}(V_{GS} - V_T)$ , однако, переход из подпорогового режима в режим сильной инверсии является постепенным, что необходимо учитывать. Это особенно важно при моделировании и при численном решении уравнений Пуассона – Больцмана для  $Q(\psi_s)$ , лежащем в основе так называемых моделей поверхностного потенциала [26].

Поведение  $Q \propto V_{GS}$  можно описать эмпирически. Одно из выражений было предложено таким [27]:

$$Q(V_{GS}) = -m C_G^{inv} \left( \frac{kT}{q} \right) \ln \left( 1 + e^{q(V_{GS} - V_T)/mkT} \right). \quad (56)$$

Для  $V_{GS} \ll V_T$  можно воспользоваться приближением  $\ln(x) \approx 1 + x$  и переписать (56) в виде

$$Q(V_{GS}) = -m C_G^{inv} \left( \frac{kT}{q} \right) e^{q(V_{GS} - V_T)/mkT}, \quad (57)$$

которое оказывается близким к уравнению (48). В практических расчетах предэкспоненциальный множитель не является существенным, так что эмпирическое выражение (57) вполне приемлемо.

При  $V_{GS} \gg V_T$  уравнение (56) дает уже известный правильный результат:

$$Q(V_{GS}) = -C_G^{inv}(V_{GS} - V_T). \quad (58)$$

Таким образом эмпирическое уравнение (56) пригодно для описания поведения  $Q \propto V_{GS}$  от подпороговой области до области с сильной инверсией. Модель виртуального истока VS использует несколько улучшенную версию уравнения (46), обеспечивая тем самым лучшее сочетание областей со слабой и сильной инверсией [18].

В заключение отметим тесную связь между токами  $I_{OFF}$  и  $I_{ON}$ . Мы видели, что  $I_{OFF} \propto \exp[(V_{GS} - V_T)/mkT]$ , а ток  $I_{ON} \propto V_{GS} - V_T$ , так что в итоге

$$\ln I_{OFF} \propto I_{ON}. \quad (59)$$

Системные дизайнеры могли бы уменьшить пороговое напряжение  $V_T$  с тем, чтобы увеличить  $I_{ON}$ , что улучшит быстродействие, однако, это влечет за собой экспоненциальный рост  $I_{OFF}$  и как следствие увеличение диссипации мощности в режиме ожидания. Рисунок 17 может служить примером такого компромисса применительно к технологии 65 нм NMOS. Этот фундаментальный компромисс является непосредственным следствием физических принципов функционирования MOSFET.

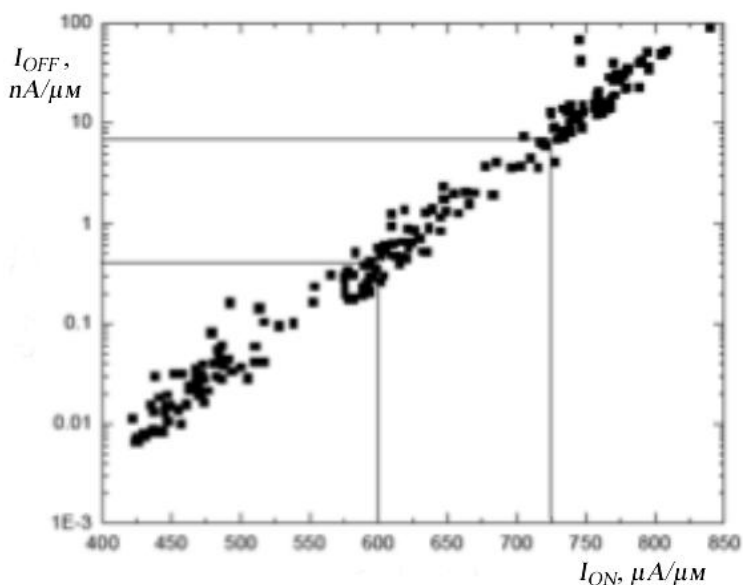


Рис. 17. График зависимости  $\log_{10} I_{OFF} \propto I_{ON}$  для технологии 65 нм NMOS [31].

Уравнение (46) дает значение напряжения на стоке  $V_{DSAT}$  в режиме сильной инверсии. Модель виртуального истока VS сводит  $V_{DSAT}$  в режиме сильной инверсии к  $kT/q$  в подпороговом режиме эмпирически. Характерная ошибка такой эвристической процедуры меньше 10% [15].

Как свидетельствует рис. 7/гл. 2, модель VS с высокой точностью вписывается в экспериментальные данные для нанотранзисторов. Это удивительно, поскольку такие параметры, как  $\mu$  и  $v_{sat}$ , имеют очевидный и ясный физический смысл в том случае, когда длина канала проводимости во много раз превосходит среднюю длину свободного пробега, но это не та ситуация, которая имеет место в нанотранзисторах. Остается пока рассматривать эти величины как эмпирические параметры. Далее будет развита теория транспорта электронов в нанотранзисторах и мы увидим, что этим параметрам может быть придан ясный физический смысл.

Подведем итоги. Подробно рассмотрена электростатика MOS. 1D электростатика изгибает зоны, понижает барьер и позволяет потоку электронов двигаться от истока к стоку. 2D электростатика деградирует транспорт электронов в полевых транзисторах увеличивая подпороговый разброс и вызывая *DIBL*, который в свою очередь увеличивает выходную проводимость и уменьшает пороговое напряжение в короткоканальных транзисторах. Количественный учет 2D электростатики требует численного подхода, вместе с тем все существенные эффекты физически понятны.

2D электростатика разрушает функционирование транзисторов и приводит к 1) подпороговому разбросу большему, чем фундаментальный предел в  $60 \text{ мВ/декада}$ , 2) сдвигу проходных характеристик  $\log_{10} I_{DS} \propto V_{GS}$  влево при увеличении напряжения на стоке (*DIBL*), 3) возникновению порога в зависимости от параметров затвора и напряжения на стоке, 4) низкому выходному сопротивлению. Когда эффекты 2D электростатики сильны, затвор теряет контроль над током и транзистор претерпевает пробой. Поскольку эти эффекты проявляются сильнее в короткоканальных транзисторах, их еще называют эффектами короткого канала. По мере того, как транзисторы становятся все меньше и меньше, основной вызов, с которым встречаются схемотехники, сводится к контролю короткоканальных эффектов. Как правило, требуется численное моделирование.

Возвращаясь к уравнению (43), видим, что зависимость  $Q|_{x=0}(V_{GS}, V_{DS})$  имеет вполне приемлемое физическое объяснение. Что же касается  $\langle v_x(V_{GS}, V_{DS}) \rangle|_{x=0}$  в этом уравнении, то далее будет построена адекватная физическая картина транспортных явлений в нанотранзисторах, исходя из обобщенной модели транспорта электронов ЛДЛ, и установлена связь ее с традиционным подходом «сверху – вниз».

## Литература

1. R. F. Pierret, *Semiconductor Device Fundamentals* (New York: Addison-Wesley: 1996).
2. B. Streetman, S. Banerjee, *Solid State Electronic Devices* (New York: Prentice Hall: 2005).
3. Y. Tsividis, C. McAndrew, *Operation and Modeling of the MOS Transistor* (New York: Oxford Univ. Press: 2011).
4. Y. Taur, T. Ning, *Fundamentals of Modern VLSI Devices* (New York: Oxford Univ. Press: 2013).
5. D. Frank, Yuan Taur, H.-S. Philip Wong, *IEEE Device Research Conf. Technical Digest*, 18 – 21 (1999).
6. D. Frank, Yuan Taur, H.-S. Philip Wong, *IEEE Electron Device Lett.*, **19**, 385 – 387 (1998).
7. Jing Wang, P. Solomon, M. Lundstrom, *IEEE Trans. Electron Dev.*, **51**: 1361 – 1365 (2004).
8. Qian Xie, Jun Xu, Yuan Taur, *IEEE Trans. Electron Dev.*, **59**: 1569 – 1579 (2012).
9. Risho Koh, Haruo Kato, H. Matsumoto, *Japanese J. Appl. Phys.*, **35**, 996 – 1000 (1996).
10. R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Ridout, E. Bassous, A. R. LeBlanc, *IEEE J. Solid-State Circuits*, **51**, 256 – 264 (1974).
11. Chenming Hu, *Modern Semiconductor Devices for Integrated Circuits* (London, UK: Pearson India: 2009).
12. Chenming Hu. *What Else Besides FinFET?*:  
[www.synopsys.com/community/resources/events/keynote-finfet.html](http://www.synopsys.com/community/resources/events/keynote-finfet.html).
13. X. Huang, W.-C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y.-K. Choi, K. Asano, V. Subramanian, T.-J. King, J. Bokor, Chenming Hu, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 67 – 70 (1999).
14. M. Jeong, B. Doris, J. Kedzierski, K. Rim, M. Yang, *Science*, **306**, 2057 – 2060 (2004).
15. A. Khakifirooz, O. M. Nayfeh, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **56**: 1674 – 1680 (2009).
16. E. O. Johnson, *RCA Review*, **34**: 80 – 94 (1973).
17. S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport* (Singapore: World Scientific: 2012).

18. M. Lundstrom, C. Jeong, *Near-equilibrium transport. Fundamentals and Applications* (Singapore: World Scientific: 2013).
19. Ю. О. Кругляк, М. В. Стріха, *Укр. Фіз. Журн. Огляди*, **10**, 3 – 32 (2015).
20. Ю. А. Кругляк, *Наноэлектроника «снизу – вверх»* (Одесса: ТЭС: 2015).
21. S. Datta, *Lessons from Nanoelectronics. Part A: Basic Concepts* (Singapore: World Scientific: 2017).
22. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
23. D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Yuan Taur, H.-S. Philip Wong, *Proc. IEEE*, **89**, 259 – 288 (2001).
24. J. Appenzeller, Y.-M. Lin, J. Knoch, Ph. Avouris, *Phys. Rev. Lett.*, **93**, 196805-1-4 (2004).
25. S. Salahuddin, S. Datta, *Nano Lett.*, **8**, 405 – 410 (2008).
26. G. Gilenblat, X. Li, W. Wu, H. Wang, A. Jha, R. van Langevelde, G. D. J. Smit, A. J. Scholten, D. B. M. Klassen, *IEEE Trans. Electron Dev.*, **53**: 1979 – 1993 (2006).
27. G. T. Wright, *Electron Lett.*, **21**, 221 – 222 (1985).
28. A. L. Steegen, R. Mo, R. Mann *et al*, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 64 – 67 (2005).

## Глава 5. Транспортная модель Ландауэра – Датта – Лундстрёма и баллистические MOSFET

### 5.1. Введение

Сначала рассмотрим обобщенную модель электронного транспорта Ландауэра – Датта – Лундстрёма (ЛДЛ) применительно к 2D каналам проводимости полевых транзисторов MOSFET. С более подробным изложением этой модели применительно к наноскопическим, так и микро- и макроскопическим проводникам произвольной размерности 1D, 2D и 3D и работающих в баллистическом, квазибаллистическом и диффузионном режимах можно ознакомиться в [1 – 6].

Ток стока MOSFET пропорционален произведению электронного заряда и скорости электронов. До сих пор мы обсуждали вопросы, связанные с зарядом. Сейчас обратимся к обсуждению скорости носителей тока. Для вычисления средней скорости нам необходима теория транспортных явлений в полупроводниках. Традиционно изложение этой теории начинается с уравнения дрейфа-диффузии [7]:

$$J_x = n_s q \mu E_x + qD \frac{dn_s}{dx}, \quad (1)$$

где  $J_x$  – плотность тока 2D электронов в тонком слое в единицах [A/м],  $n_s$  – поверхностная плотность электронов в единицах [м<sup>-2</sup>]. Это уравнение вполне пригодное и привычное в качестве исходного уравнения для обсуждения транспортных явлений в транзисторах с достаточно длинным каналом проводимости, однако, не является таковым, если речь идет о нанотранзисторах.

В применении к нанотранзисторам мы воспользуемся концепцией Ландауэра [8], развитой Датта [1, 5] и Лундстрёмом [2, 6], согласно которой ток определяется уравнением

$$I = \frac{2q}{h} \int_{-\infty}^{+\infty} T(E) M(E) (f_1(E) - f_2(E)) dE, \quad [A], \quad (2)$$

где  $T(E)$  – коэффициент прохождения,  $M(E)$  – число мод (каналов) проводимости,  $f_1(E)$  и  $f_2(E)$  – фермиевские функции контактов 1 и 2.

## 5.2. Обобщенная модель электронного транспорта

Схема наноканала проводимости в полевом транзисторе представлена на рис. 1. Предполагается, что контакты массивные и имеет место сильное неупругое электрон-фононное рассеяние, так что электроны в контактах находятся в термодинамическом равновесии. В состоянии равновесия вероятность того, что электронное состояние с энергией  $E$  занято, определяется функцией Ферми

$$f_{1,2}(E) = \frac{1}{1 + e^{(E - E_{F1,2})/kT}}, \quad (3)$$

которая в нашем контексте называется также электрохимическим потенциалом контактов 1 и 2.

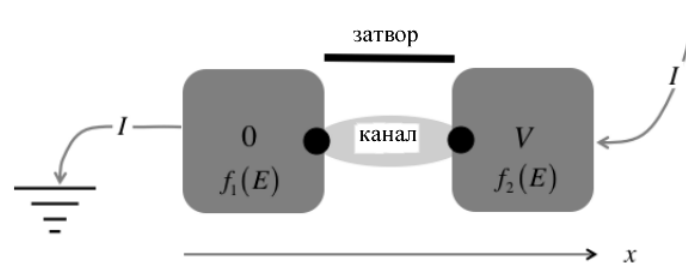


Рис. 1. Иллюстрация наноканала проводимости MOSFET с двумя массивными контактами, находящимися в термодинамическом равновесии. Если контакты имеют разную температуру или на них поданы разные потенциалы, то  $f_1(E) \neq f_2(E)$  и по каналу пойдет ток.

Если температура контактов одинакова и на них подано одинаковое напряжение, то в соответствии с (2) ток по каналу не идет. Это утверждение непосредственно следует из физического смысла фермиевской функции, поскольку в этой ситуации вероятности того, что состояние с энергией  $E$  на контактах 1 и 2 занято, одинаковы.

Теперь рассмотрим ситуацию когда  $f_1(E) \neq f_2(E)$ : ток идет. Согласно (3), в этом случае возможны два варианта возникновения тока. Во-первых, температура обоих контактов может быть различной, что вызовет термоэлектрические эффекты [1, 2, 4, 5, 9]. Во-вторых, на контакты может быть подано различное напряжение.

Пусть, например, левый контакт 1 заземлен, а на второй контакт подано напряжение  $V$  (рис. 1). Если подать положительное напряжение, то электрохимический потенциал правого контакта понизится на величину  $qV$ :

$$E_{F2} = E_{F1} - qV. \quad (4)$$

При этом предполагается, что даже при подаче напряжения вероятность того, что состояние занято, определяется равновесной функцией Ферми, которая принимает разные значения на контактах. Строго говоря, это не может быть так, поскольку при появлении тока система выходит из состояния равновесия, однако, мы предполагаем, что контакты настолько массивны, что подача небольшого напряжения вызывает лишь незначительное отклонение из состояния равновесия, которым можно пренебречь.

Когда имеет место разница в напряжениях, поданных на контакты, возникает различие  $f_1(E) \neq f_2(E)$  между фермиевскими функциями в некоторой области энергии, называемое окном Ферми. Концепция фермиевского окна проводимости иллюстрируется на рис. 2.

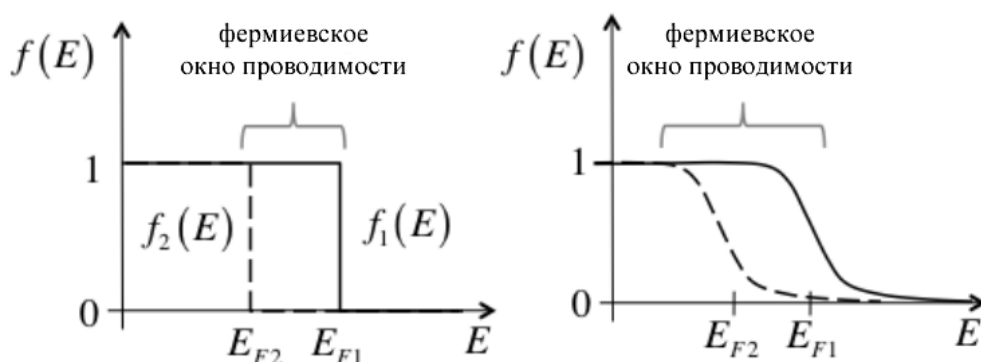


Рис. 2. Концепция фермиевского окна при подаче относительно большого напряжения: слева:  $T = 0$ ; справа:  $T > 0$  К.

Рисунок слева соответствует  $T = 0$ , а справа  $T > 0$  К:  $f_1(E) \neq f_2(E)$ . Согласно уравнению Ландауэра для тока (2), только электроны в канале проводимости, энергия которых попадает в окно Ферми, дают вклад в величину тока.

Различие между электрохимическими потенциалами контактов является причиной появления тока, однако, как видно из уравнения Ландауэра (2), ток при энергии  $E$  пропорционален произведению  $T(E)M(E)$ . Величина  $M(E)$  – это число мод (каналов) проводимости при энергии  $E$ . Число мод проводимости «напоминает» число полос автомобильного шоссе. Чем больше мод, тем сильнее ток, имея в виду, что проводят ток только те моды, энергия которых попадает в окно проводимости Ферми. Можно ожидать, что число мод зависит от плотности состояний при энергии  $E$  и от скорости электронов. Величина  $T(E)$  получила название коэффициента прохождения. Он равен вероятности того, что электрон, покинув контакт 1, пройдет весь канал проводимости без рассеяния назад и вернется в исходный контакт 1. Коэффициент прохождения



отличен от единицы при наличии рассеяния назад. Если электрон вышел из контакта 1 и рассеялся назад, он может вернуться в контакт 1. Вероятность рассеяния назад зависит от длины канала проводимости  $L$  и от среднего расстояния между соседними актами рассеяния назад  $\lambda$ , так называемой средней длины пробега между столкновениями назад (mean-free-path/mfp for backscattering). Когда  $L \ll \lambda$ , то  $T \rightarrow 1$ , а при  $L \gg \lambda$ , то  $T \rightarrow 0$ . В уравнении (2) предполагается, что вероятность прохождения электрона от контакта 1 до контакта 2 равна вероятности прохождения электрона от контакта 2 до контакта 1. Можно показать, что это имеет место при упругом рассеянии электронов, т. е. тогда, когда электроны при движении занимают разные, параллельные моды. В концепции Ландауэра предполагается, что в канале проводимости происходит упругое рассеяние электронов, а в контактах имеет место сильное неупругое рассеяние.

В итоге, уравнение Ландауэра (2) пригодно для описания транспорта электронов от баллистического предела без рассеяния с  $T=1$  и до диффузионного предела с интенсивным рассеянием с  $T \ll 1$ . Полный ток получается суммированием вкладов от всех мод, которые полагаются независимыми, поскольку неупругие столкновения, ведущие к взаимодействию мод, исключаются.

### *Большие и малые напряжения*

Разность электрохимических потенциалов  $f_1(E) - f_2(E) = qV$  играет важную роль в выражении Ландауэра для тока (2). Ситуации, когда эта разность велика или мала, различны. Когда большое напряжение прилагается к контакту 2, тогда для всех энергий  $f_1(E) \gg f_2(E)$  и (2) упрощается до

$$I = \frac{2q}{h} \int_{-\infty}^{+\infty} T(E) M(E) f_1(E) dE. \quad (5)$$

Будем пользоваться этим уравнением для области насыщения MOSFET.

Когда к контакту 2 прилагается небольшое напряжение, уравнение (2) также упрощается (рис. 3).

При  $T=0$  фермиевское окно выглядит как  $\delta$ -функция при энергии  $E = E_F$  (рис. 3/слева). При  $T > 0$  фермиевское окно  $f_1(E) - f_2(E)$  напоминает острый пик, центрированный при  $E = E_F$  (рис. 3/справа). При малых

напряжениях, отвечающих квазиравновесным состояниям, можно воспользоваться разложением  $f_2(E)$  в ряд Тейлора

$$f_2(E) \approx f_1(E) + \frac{\partial f_1}{\partial E_F} \delta E_F, \quad (6)$$

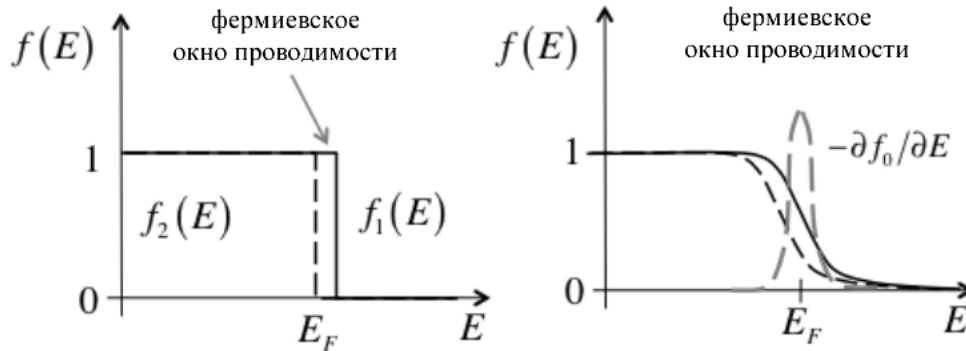


Рис. 3. Концепция фермиевского окна при подаче небольшого напряжения: слева:  $T = 0$ ; справа:  $T > 0 K$ .

откуда малая разница между электрохимическими потенциалами контактов есть

$$f_1(E) - f_2(E) = -\left(\frac{\partial f_1}{\partial E_F}\right) \delta E_F = -\left(-\frac{\partial f_1}{\partial E}\right) \delta E_F, \quad (7)$$

где мы воспользовались свойством функции Ферми

$$\partial f / \partial E_F = -\partial f / \partial E. \quad (8)$$

Поскольку  $\delta E_F = -qV$ , то

$$f_1(E) - f_2(E) = q \left(-\frac{\partial f_0}{\partial E}\right) V, \quad (9)$$

где мы воспользовались тем, что вблизи равновесия  $f_1(E) \approx f_2(E) \approx f_0(E)$ . Итак, вблизи равновесия после подстановки (9) в (2) получаем, что

$$I = GV, \quad [A]$$

$$G = \frac{2q^2}{h} \int T(E) M(E) \left(-\frac{\partial f_0}{\partial E}\right) dE. \quad [Cm] \quad (10)$$

Часто именно это уравнение для проводимости называют уравнением Ландауэра. Мы воспользуемся этим уравнением для вычисления тока в линейной области выходных характеристик MOSFET.

Наконец, по определению, фермиевское окно проводимости

$$W(E) \equiv \left( -\frac{\partial f_0}{\partial E} \right), \quad (11)$$

площадь под которым равна единице:

$$\int_{-\infty}^{+\infty} W(E) dE = 1, \quad (12)$$

что следует из

$$\int_{-\infty}^{+\infty} W(E) dE = \int_{-\infty}^{+\infty} \left( -\frac{\partial f_0}{\partial E} \right) dE = - \int_{-\infty}^{+\infty} df_0 = f_0(-\infty) - f_0(+\infty) = 1,$$

где последнее равенство вытекает из свойств функции Ферми (3). При стремлении температуры к нулю фермиевское окно  $W(E)$  стремится к  $\delta$ -функции при  $E = E_F$ . Поскольку площадь окна проводимости Ферми равна единице, то фермиевское окно для металлов можно рассматривать как  $\delta$ -функцию,

$$W(E) \approx \delta(E - E_F).$$

### *Коэффициент прохождения*

Рассмотрим следующую ситуацию: стационарный поток электронов  $\underline{F^+}|_{x=0}$  впрыскивается слева в однородный проводник в отсутствие электрического поля, а справа выходит поток  $\underline{F^+}|_{x=L}$  (рис. 4). Впрыскивания справа не происходит.

Коэффициент прохождения  $T$  определяется как отношение доли электронов, покинувших проводник, к впрыснутым в проводник:

$$T = \frac{\underline{F^+}|_{x=L}}{\underline{F^+}|_{x=0}}. \quad (0 \leq T \leq 1) \quad (13)$$

Часть электронов, впрыснутых в проводник, проходит через него, а часть электронов испытывает рассеяние назад и покидает проводник влево:

$$\underline{F^+}|_{x=0} = \underline{F^+}|_{x=L} + \underline{F^-}|_{x=0},$$

так что

$$\underline{F^-}|_{x=0} = (1-T)\underline{F^+}|_{x=0}.$$

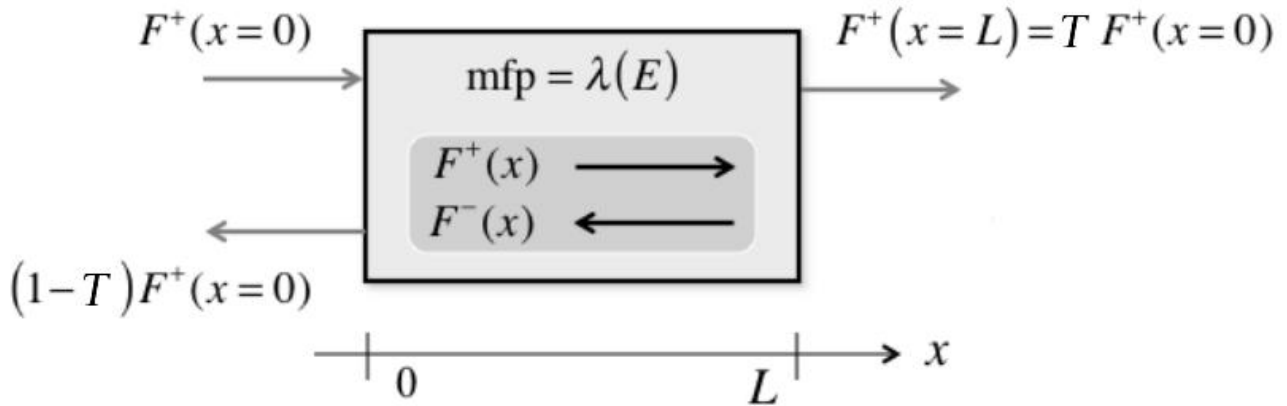


Рис. 4. Впрыскивание в проводник потока электронов  $\underline{F^+}|_{x=0}$  слева, доля которых  $T\underline{F^+}|_{x=0} = \underline{F^+}|_{x=L}$  покидает проводник справа. Предполагается, что впрыскивания в проводник справа не происходит, как и отсутствуют процессы генерации и рекомбинации электронов внутри проводника.

Ситуация, показанная на рис. 4, может иллюстрировать диффузию электронов через базу биполярного транзистора: поток электронов инжектируется эмиттером в базу при  $x = 0$  и выходит при  $x = L$  в коллектор. Если напряжение на коллекторе достаточно велико, то коллектор работает как поглощающий контакт: поглощаются все электроны и ни один электрон не возвращается в базу.

Рассмотрим сначала случай с тонкой базой,  $L \ll \lambda$ . Практически все инжектированные электроны покидают базу, так что  $F^+(L) = F^+(0)$  и  $F^-(0) = 0$ . Это баллистический предел и коэффициент прохождения

$$T_{ball} = 1. \quad (14)$$

В диффузионном пределе  $L \gg \lambda$ . Канал проводимости во много раз больше средней длины свободного пробега, и коэффициент прохождения небольшой. Эта ситуация типична для микроэлектроники.

Для вычисления коэффициента прохождения в диффузионном пределе пусть концентрация электронов в точке впрыскивания  $x=0$  есть  $\underline{n}|_{x=0}$ . Если канал проводимости достаточно длинный, то  $\underline{n}|_{x=L} \approx 0$ .

Суммарный поток электронов определяется законом диффузии Фика:

$$F = -D \frac{dn}{dx} = D \frac{n|_{x=0}}{L} = \underline{F^+}|_{x=L}. \quad (15)$$

В диффузионном пределе

$$\underline{F^+}|_{x=0} = \frac{n|_{x=0}}{2} v_T, \quad (16)$$

где двойка обязана тому, что в диффузионном пределе приблизительно половина электронов при входе в канал проводимости имеет положительную скорость, а другая половина – отрицательную, вызванную рассеянием назад. Скорость  $v_T$  – это средняя тепловая скорость электронов. В предположении статистики Максвелла – Больцмана это однонаправленная тепловая скорость

$$v_T = \sqrt{\frac{2kT}{\pi m^*}}. \quad (17)$$

Из (13), (15) и (16) находим

$$T = \frac{\underline{F^+}|_{x=L}}{\underline{F^+}|_{x=0}} = \frac{F}{\underline{F^+}|_{x=0}} = \frac{Dn|_{x=0}/L}{v_T n|_{x=0}/2} = \frac{2D}{v_T L}, \quad (18)$$

где коэффициент диффузии известным образом связан с тепловой скоростью и средней длиной свободного пробега [2, 4]

$$D = \frac{v_T \lambda}{2}, \quad [cm^2/c], \quad (19)$$

так что окончательно в диффузионном пределе

$$T_{diff} = \frac{\lambda}{L}. \quad (20)$$

Как и ожидалось, коэффициент прохождения в диффузионном пределе небольшой, поскольку  $L \gg \lambda$ .

Итак, получены выражения для коэффициента прохождения в двух предельных режимах – баллистическом и диффузионном. Современные нанотранзисторы работают в промежуточном квазибаллистическом режиме. В общем случае

$$T = \frac{\lambda_0}{\lambda_0 + L}, \quad (21)$$

$$T(E) = \frac{\lambda(E)}{\lambda(E) + L}.$$

Первое уравнение предполагает, что средняя длина свободного пробега не зависит от энергии, в результате чего коэффициент прохождения одинаков для всех мод проводимости. Во втором уравнении коэффициент прохождения зависит от энергии мод проводимости и охватывает как оба предельных режима – баллистический ( $L \ll \lambda$ ,  $T \rightarrow 1$ ) и диффузионный ( $L \gg \lambda$ ,  $T \ll 1$ ), так и промежуточный квазibalлистический режим ( $L \approx \lambda$ ,  $T < 1$ ). Это общее выражение для коэффициента прохождения следует из транспортного уравнения Больцмана [2 – 5].

Важно отметить, что фигурирующая в (21) величина  $\lambda$  есть средняя длина свободного рассеяния назад (mean-free-path for backscattering). Ее физический смысл связан с вероятностью (на единицу длины) того, что поток электронов вперед рассеивается назад. Иначе говоря, средняя длина свободного рассеяния есть просто среднее расстояние между двумя соседними по времени событиями рассеяния

$$\Lambda(E) \equiv v(E)\tau(E). \quad (22)$$

В случае 2D канала проводимости [2 – 5]

$$\lambda(E) \equiv \frac{\pi}{2} v(E)\tau_m(E), \quad (23)$$

где  $\tau_m(E)$  – время релаксации импульса, всегда большее чем среднее время между двумя последовательными актами рассеяния  $\tau(E)$ , так что  $\lambda > \Lambda$ : средняя длина свободного рассеяния назад всегда больше средней длины свободного рассеяния.

Однонаправленная тепловая скорость (17) играет важную роль в транспортных задачах. В состоянии равновесия средняя скорость электронов равна нулю, поскольку средняя скорость электронов, движущихся в направлении  $+x$ , есть величина положительная, равная по абсолютной величине средней скорости электронов, движущихся в направлении  $-x$ . Вычислим эту однонаправленную скорость для 2D полупроводника с параболической зонной структурой.

Сначала рассмотрим усреднение по углам (рис. 5). Скорость вдоль оси  $x$  есть  $v(E)\cos\theta$ . В 2D проводнике  $E = \hbar^2(k_x^2 + k_y^2)/2m^*$  и скорость  $v(E)$  не зависит от угла. Средняя скорость в направлении  $+x$

$$\langle v_x^+(E) \rangle = \frac{\int_{-\pi/2}^{+\pi/2} v(E)\cos\theta d\theta}{\pi} = \frac{2}{\pi} v(E), \quad (24)$$

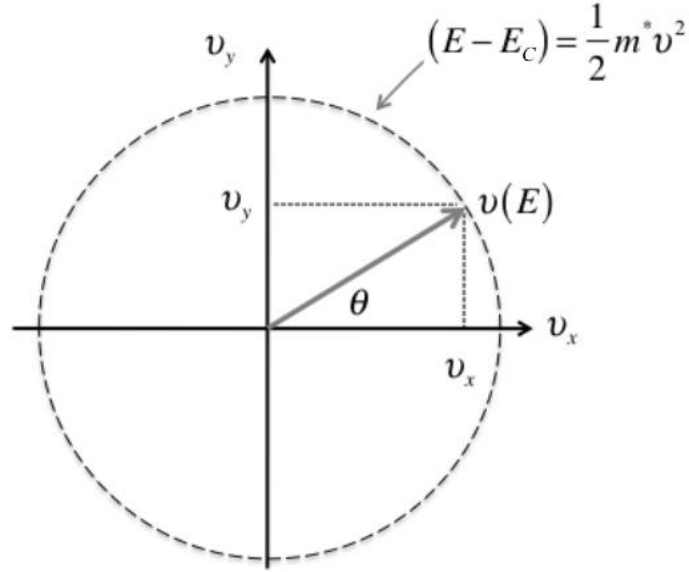


Рис. 5. Вектор скорости  $v(E)$  при энергии  $E$  в плоскости  $xу$ .  
Для параболической зонной структуры величина вектора скорости определяется его энергией и не зависит от направления.

где угловые скобки означают усреднение по углу в плоскости  $xу$  при энергии  $E$ .

Нас интересует скорость  $\langle\langle v_x^+(E) \rangle\rangle$ , где двойные угловые скобки означают усреднение как по углам, так и по энергии:

$$\langle\langle v_x^+(E) \rangle\rangle = \frac{\int_{E_C}^{\infty} \langle v_x^+(E) \rangle D_{2D}(E) f_0(E) dE}{\int_{E_C}^{\infty} D_{2D}(E) f_0(E) dE} = \frac{\frac{2}{\pi} \int_{E_C}^{\infty} v(E) D_{2D}(E) f_0(E) dE}{\int_{E_C}^{\infty} D_{2D}(E) f_0(E) dE}.$$

Для параболических зон скорость и плотность состояний [1 – 5]

$$v(E) = \sqrt{\frac{2(E - E_C)}{m^*}}, \quad D_{2D}(E) = g_v \frac{m^*}{\pi \hbar^2},$$

так что

$$\langle\langle v_x^+(E) \rangle\rangle = \frac{g_v \frac{m^*}{\pi \hbar^2} \frac{2}{\pi} \int_{E_C}^{\infty} \sqrt{\frac{2(E - E_C)}{m^*}} \frac{dE}{1 + e^{(E - E_F)/kT}}}{g_v \frac{m^*}{\pi \hbar^2} \int_{E_C}^{\infty} \frac{dE}{1 + e^{(E - E_F)/kT}}}, \quad (25)$$

а после замены переменных

$$\eta \equiv \frac{E - E_C}{kT}, \quad \eta_F \equiv \frac{E_F - E_C}{kT} \quad (26)$$

получим

$$\langle \langle v_x^+(E) \rangle \rangle = \sqrt{\frac{2kT}{\pi m^*}} \frac{2 \int_0^\infty \frac{\eta^{1/2} d\eta}{1 + e^{(\eta - \eta_F)}}}{\int_0^\infty \frac{d\eta}{1 + e^{(\eta - \eta_F)}}}.$$

В числителе имеем интеграл Ферми – Дирака порядка  $1/2$ , а в знаменателе – порядка 0 [4, 10]:

$$\langle \langle v_x^+(E) \rangle \rangle = \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{F}_{1/2}(\eta_F)}{\mathfrak{F}_0(\eta_F)}. \quad (27)$$

Ниже порога предполагается, что справедлива статистика Максвелла – Больцмана, так что интегралы Ферми – Дирака стремятся к  $\exp(\eta_F)$  [4, 10], и в результате получается уже приведенное выше уравнение (17).

### *Моды проводимости*

Распределение мод  $M(E)$  дает число каналов проводимости при энергии  $E$ , по которым может идти ток. Вывод и подробное обсуждение этого понятия для 1D, 2D и 3D проводников можно найти в [1 – 5]. Здесь же речь пойдет только о 2D проводниках, поскольку нас прежде всего интересуют каналы проводимости в нанотранзисторах.

Можно ожидать, что число мод проводимости должно быть связано с плотностью состояний и со скоростью электронов:

$$M(E) \propto \langle v_x^+(E) \rangle D(E) / 4, \quad (28)$$

где  $D(E)dE$  есть число состояний с энергией между  $E$  и  $E + dE$ . Множитель четверка включает в себя вырождение по спину двойку; нам нужна плотность состояний в пересчете на спин  $D(E)/2$ , поскольку вырождение по спину уже учтено в (2) в виде двойки в этой формуле. Еще один множитель двойка возникает потому, что только в половине состояний  $D(E)/2$  скорости электронов направлены так же, как и ток. Анализ размерности (28) показывает, что коэффициент пропорциональности должен иметь размерность постоянной Планка [Дж·с], так что окончательно

$$M(E) = \frac{h}{4} \langle v_x^+(E) \rangle D(E). \quad (29)$$



В применении к плоским MOSFET, в которых электроны движутся по 2D каналам,

$$M_{2D}(E) = \frac{\hbar}{4} \langle v_x^+(E) \rangle D_{2D}(E) \quad [M^{-1}] \quad (30)$$

дает число мод при энергии  $E$  на единицу ширины канала проводимости.

Для параболических зон 2D плотность состояний дается выражением [1 – 7]

$$D_{2D}(E) = g_v \left( \frac{m^*}{\pi \hbar^2} \right), \quad [Дж^{-1} \cdot м^{-2}], \quad (31)$$

где  $g_v$  – долиное вырождение. В случае MOSFET эти 2D состояния лежат в зоне проводимости или в валентной зоне при энергии соответственно выше  $E_C + \varepsilon_1$  или ниже  $E_V - \varepsilon_1$ , где  $\varepsilon_1$  это локализация энергии (confinement energy) для нижней подзоны (рис. 14/гл. 3). Аналогичным образом учитываются другие подзоны, если это необходимо.

Согласно (30), распределение мод в 2D канале проводимости с учетом (24)

$$M_{2D}(E) = g_v \frac{\sqrt{2m^*[E - (E_C + \varepsilon_1)]}}{\pi \hbar}, \quad (12.28)$$

что графически показано на рис. 6. Подобным же образом можно получить распределение мод в каналах 1D и 3D с параболической зонной структурой или, например, в графене, задавшись соответствующими выражениями для плотности состояний и скорости [1 – 5].

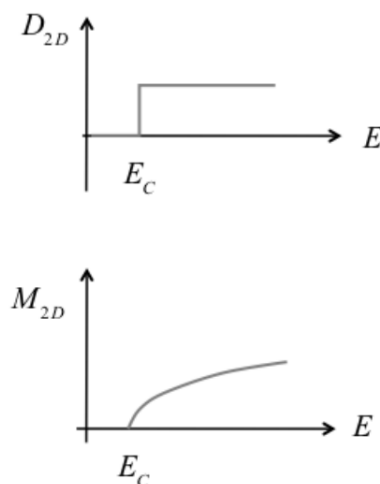


Рис. 6. Распределение мод проводимости с энергией и плотности состояний для 2D канала проводимости с параболической зонной структурой.

Для понимания физики транзисторов со структурой MOS нужно разобраться каким образом напряжение на затворе контролирует число электронов в канале проводимости и сам результирующий ток. Чтобы связать плотность электронов с электрохимическим потенциалом, воспользуемся плотностью состояний, а для увязки тока с электрохимическим потенциалом воспользуемся распределением мод в соответствии с (2). Такой подход аналогичен использованию двух различных «эффективных масс» в традиционной теории полупроводников – эффективной массы плотности состояний (density-of-states effective mass) и эффективной массы проводимости (conductivity effective mass).

### Квант проводимости

Рассмотрим проводимость 2D канала при  $T=0\text{ K}$ . Воспользуемся выражением (10) и тем, что при нулевой температуре фермиевское окно  $(-\partial f_0 / \partial E)$  ведет себя как  $\delta$ -функция при  $E = E_F$ . Тогда проводимость

$$\underline{G}|_{T=0} = \frac{2q^2}{h} T(E_F) M(E_F). \quad (33)$$

Если предположить, что транспорт баллистический,  $T(E_F) = 1$ , который сейчас несложно экспериментально продемонстрировать с нанопроводниками при низких температурах, тогда

$$\underline{G}|_{T=0} = \frac{2q^2}{h} M(E_F) = \frac{M(E_F)}{12.9\text{ k}\Omega}. \quad (34)$$

В коротких проводниках число мод невелико и вполне счетно. Итак, проводимость квантуется в единицах  $2q^2/h$ , что составляет  $1/12.9\text{ k}\Omega^{-1}$ .

Квантование проводимости – надежно установленный опытный факт. Рисунок 7 демонстрирует экспериментальный факт квантования проводимости электронного газа в интерфейсе AlAs/Al<sub>x</sub>Ga<sub>1-x</sub>As.

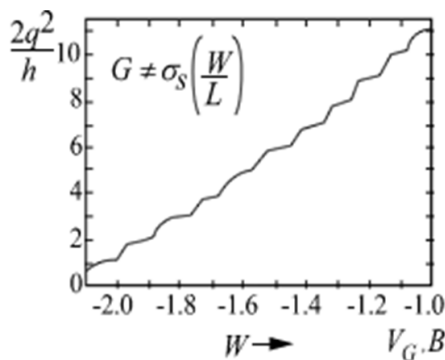


Рис. 7. Квантование проводимости электронного газа в интерфейсе AlAs/Al<sub>x</sub>Ga<sub>1-x</sub>As [15, 16].

В этом эксперименте ширина канала проводимости  $W$  изменялась электростатически путем подачи напряжения на затвор, в результате чего менялось число мод проводимости, что влекло за собой скачкообразное изменение проводимости в полном соответствии с (34). Этот эксперимент проводился при низкой температуре с тем, чтобы добиться баллистического транспорта электронов. Однако, современные нанотранзисторы имеют настолько короткий канал проводимости, что эффекты квантования проводимости проявляются уже при комнатной температуре.

Баллистический транспорт,  $T(E_F)=1$ , и квазibalлистический транспорт,  $T(E_F) \sim < 1$ , не являются необычными в современных нанотранзисторах. Все же в большинстве используемых ныне устройствах нанoeлектроники нанотранзисторы еще достаточно велики и дискретная природа их мод проводимости не проявляется. Распределение мод проводимости рассматривается как непрерывное свойство, пропорциональное ширине канала проводимости  $W$ . Мы далее будем предполагать, что

$$M(E) = W \cdot M_{2D}(E), \quad (35)$$

где  $M_{2D}(E)$  дается уравнением (32).

### *Плотность электронов*

Согласно традиционной теории полупроводников, плотность электронов в зоне проводимости дается интегралом от произведения плотности состояний при энергии  $E$  на вероятность того, что состояния при энергии  $E$  заняты [7]:

$$n_S = \int_{E_C}^{\infty} D_{2D}(E) f_0(E) dE. \quad [M^{-2}] \quad (36)$$

Для плотности состояний воспользуемся выражением (31), тогда

$$n_S = \int_{E_C}^{\infty} \left( g_v \frac{m^*}{\pi \hbar^2} \right) \frac{dE}{1 + e^{(E-E_F)/kT}} = \left( g_v \frac{m^* kT}{\pi \hbar^2} \right) \int_0^{\infty} \frac{d\eta}{1 + e^{\eta - \eta_F}}, \quad (37)$$

где  $\eta$  и  $\eta_F$  определены в (26). После взятия интеграла имеем

$$n_S = \left( g_v \frac{m^* kT}{\pi \hbar^2} \right) \ln(1 + e^{\eta_F}) = N_{2D} \ln(1 + e^{\eta_F}) = N_{2D} \mathfrak{F}_0(\eta_F), \quad (38)$$

где  $N_{2D}$  есть 2D эффективная плотность состояний.

Теперь вычислим плотность электронов в канале на рис. 1. Состояния в этом канале заняты электронами, поступающими из контакта 1 с электрохимическим потенциалом  $E_{F1}$  и из контакта 2 с электрохимическим потенциалом  $E_{F2}$ . Вероятность того, что состояние при энергии  $E$  занято электронами, поступившими из контакта 1, есть  $f_1(E_{F1})$ , а поступившими из контакта 2, есть  $f_2(E_{F2})$ . Теперь (36) можно переписать следующим образом:

$$n_S = \int_{E_C}^{\infty} \left( \frac{D_{2D}(E)}{2} f_1(E) + \frac{D_{2D}(E)}{2} f_2(E) \right) dE, \quad (39)$$

предполагая, что контакты идентичны. Воспользовавшись (38), получим

$$n_S = \frac{N_{2D}}{2} \mathfrak{F}_0(\eta_F) + \frac{N_{2D}}{2} \mathfrak{F}_0(\eta_F - qV/kT), \quad (40)$$

где  $\eta_F = (E_{F1} - E_C)/kT$ . Видим, что вычисленная неравновесная плотность электронов связана с плотностью состояний аналогично тому, как это имеет место и в случае равновесия. Нужно просто помнить, что при наличии тока есть два различных фермиевских уровня и две различных группы состояний в канале проводимости, одна группа связана в контактом 1, а другая – с контактом 2.

Оба рассмотренных понятия – моды проводимости  $M(E)$  и плотности состояний  $D(E)$  необходимы для моделирования MOSFET.

Подход Ландауэра – Датта – Лундстрема к транспорту электронов в наноразмерных каналах проводимости пригоден для всех режимов транспорта: от баллистического до диффузионного и для всех промежуточных квазибаллистических режимов. В качестве примера воспользуемся этим подходом для 2D проводника при низком поданном на него напряжении. Если длина и ширина проводника достаточно велики, то мы имеем стандартное выражение для проводимости

$$G = \sigma_s \frac{W}{L} = n_S q \mu \frac{W}{L}, \quad (41)$$

где  $\sigma_s$  – удельная поверхностная проводимость. Это выражение предполагает, что длина канала проводимости существенно превышает среднюю длину свободного пробега,  $L \gg \lambda$ , т. е. реализуется режим диффузионного предела. Посмотрим какие результаты дает подход Ландауэра – Датта – Лундстрема.

Начнем с уравнения (10) для режима диффузионного транспорта с  $T = \lambda/L$  и  $M(E)$  по (35). Для проводимости имеем

$$G = \frac{2q^2}{h} \int_{E_C}^{\infty} \left( \frac{\lambda(E)}{L} \right) \left( \frac{g_v W \sqrt{2m^*(E - E_C)}}{\pi \hbar} \right) \left( -\frac{\partial f_0}{\partial E} \right) dE. \quad (42)$$

Воспользуемся свойством функции Ферми (8) и вынесем производную из под знака интеграла:

$$G = \left[ \frac{2q^2}{h} \left( \frac{g_v \sqrt{2m^*}}{\pi \hbar} \right) \frac{\partial}{\partial E_F} \int_{E_C}^{\infty} \frac{\lambda(E) \sqrt{E - E_C}}{1 + e^{(E - E_F)/kT}} dE \right] \frac{W}{L}. \quad (43)$$

Пусть простоты ради средняя длина свободного пробега не зависит от энергии,  $\lambda(E) = \lambda_0$ , тогда после замены переменных (26) имеем:

$$G = \left[ \frac{2q^2}{h} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) \lambda_0 \frac{\partial}{\partial \eta_F} \int_0^{\infty} \frac{\eta^{1/2}}{1 + e^{\eta - \eta_F}} d\eta \right] \frac{W}{L} \quad (44)$$

или иначе через интеграл Ферми – Дирака

$$G = \left[ \frac{2q^2}{h} \frac{\sqrt{\pi}}{2} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) \lambda_0 \frac{\partial \mathfrak{F}_{1/2}(\eta_F)}{\partial \eta_F} \right] \frac{W}{L}. \quad (45)$$

Воспользуемся свойством интегралов Ферми – Дирака

$$d\mathfrak{F}_j/d\eta_F = \mathfrak{F}_{j-1}, \quad (46)$$

тогда

$$G = \left[ \frac{2q^2}{h} \frac{\sqrt{\pi}}{2} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) \lambda_0 \mathfrak{F}_{-1/2}(\eta_F) \right] \frac{W}{L}. \quad (47)$$

Опять для упрощения предположим невырожденную статистику электронов. В этом случае интегралы Ферми – Дирака сводятся к экспонентам, так что окончательно после некоторой перегруппировки сомножителей получаем:

$$G = \left[ \frac{q^2}{kT} \left( \frac{g_v m^* kT}{\pi \hbar^2} \right) e^{\eta_F} \sqrt{\frac{2kT}{\pi m^*}} \lambda_0 \right] \frac{W}{L}. \quad (48)$$

Теперь мы можем распознать некоторые сомножители, а именно:

$$n_S = \left( \frac{g_v m^* kT}{\pi \hbar^2} \right) e^{\eta_F} = N_{2D} e^{\eta_F},$$

$$v_T = \sqrt{\frac{2kT}{\pi m^*}},$$

$$D = \frac{v_T \lambda_0}{2},$$

$$\frac{D}{\mu} = \frac{kT}{q},$$

где  $v_T$  – однонаправленная тепловая скорость (17), а  $D$  – коэффициент диффузии (19), а уравнение (48) можно переписать в виде

$$G = n_S q \mu \frac{W}{L}, \quad (49)$$

где

$$\mu = \frac{v_T \lambda_0}{2kT/q}. \quad (50)$$

Уравнение (49) идентично привычному результату (41) в диффузионном пределе, а уравнение (50) дает подвижность, выраженную через не зависящую от энергии среднюю длину свободного пробега.

Эти выкладки показывают, что подход ЛДЛ приводит к традиционным результатам в диффузионном пределе, однако, его преимущество состоит в том, что он работает также и в баллистическом пределе. Убедимся в этом, вычислив проводимость. Можно воспользоваться уравнением (10) с  $T = \lambda(E)/L = 1$  или прямо из (42) для баллистической проводимости имеем:

$$G_B = \frac{2q^2}{h} \int_{E_C}^{\infty} \left( \frac{g_v W \sqrt{2m^*(E - E_C)}}{\pi \hbar} \right) \left( -\frac{\partial f_0}{\partial E} \right) dE. \quad (51)$$

Поступаем как и в случае диффузионного предела. Вместо (44) имеем:

$$G_B = \left[ \frac{2q^2}{h} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) \frac{\partial}{\partial \eta_F} \int_0^{\infty} \frac{\eta^{1/2}}{1 + e^{\eta - \eta_F}} d\eta \right] W. \quad (52)$$

В невырожденном пределе после интегрирования имеем:

$$G_B = q n_S \left( \frac{v_T}{2kT/q} \right) W. \quad (53)$$

Как и ожидалось, баллистическая проводимость не зависит от длины канала проводимости  $L$ . И тем не менее, мы можем записать баллистическую проводимость в традиционной форме (49), если (53) умножить и разделить на  $L$  и ввести понятие баллистической подвижности по Шуру [13]

$$\mu_B = \frac{v_T L}{2kT/q}. \quad (54)$$

Если в (49) вместо подвижности  $\mu$  использовать баллистическую подвижность  $\mu_B$ , то мы получим баллистическую проводимость. Последняя отличается от традиционной подвижности массивного проводника лишь заменой среднего пути свободного пробега  $\lambda_0$  на длину резистора  $L$ .

Почему так важно понятие о баллистической подвижности в канале проводимости, в котором не происходит рассеяние электронов? В массивном полупроводнике среднее расстояние между соседними по времени столкновениями есть  $\lambda_0$  и подвижность служит надежной характеристикой вещества. В баллистическом проводнике не происходят столкновения в канале проводимости, однако, в контактах, снабжающих канал проводимости электронами, происходит интенсивное рассеяние; таким образом, в таких проводниках расстояние между соседними актами рассеяния есть как раз длина канала проводимости. Представляется естественным в этой ситуации заменить среднее расстояние между актами рассеяния на длину канала проводимости, что и ведет к понятию баллистической подвижности. Это понятие оказывается просто параметром для записи выражения для проводимости в баллистическом канале в традиционной, привычной для диффузионного транспорта, форме (49) и одновременно имеющим очевидную физическую интерпретацию.

Современные нанотранзисторы обычно работают в режиме, промежуточном между баллистическим и диффузионным режимами, в так называемом квазибаллистическом режиме, когда время прохождения дается выражением (21). Как и прежде, проводимость в таком режиме можно вычислить, опираясь на уравнение (10). В этом случае вместо (42) имеем:

$$G = \frac{2q^2}{h} \int_{E_C}^{\infty} \left( \frac{\lambda(E)}{\lambda(E) + L} \right) \left( \frac{g_v W \sqrt{2m^*(E - E_C)}}{\pi \hbar} \right) \left( -\frac{\partial f_0}{\partial E} \right) dE. \quad (55)$$

Как и прежде, вычисления провести особенно легко, если предположить, что  $\lambda(E) = \lambda_0$ . Опять же, результат вычислений можно представить в традиционной

форме (49), если подвижность  $\mu$  заменить на кажущуюся подвижность  $\mu_{app}$ , определяемую следующим образом:

$$\frac{1}{\mu_{app}} = \frac{1}{\mu_B} + \frac{1}{\mu}. \quad (56)$$

Наименьшая из двух подвижностей, баллистическая или диффузионная, ограничат ток в нанопроводнике. По мере уменьшения длины канала проводимости баллистическая подвижность уменьшается согласно (54). Если  $\lambda \gg L$ , баллистическая подвижность в выражении (56) будет доминировать ( $\mu_B \ll \mu$ ) и кажущаяся подвижность численно приблизится к баллистической. Если мы воспользуемся уравнением (49) для вычисления проводимости резистора, длина которого меньше сравнительно с длиной свободного пробега, не учитывая при этом баллистическую подвижность, мы найдем проводимость выше баллистического предела. Баллистическая подвижность должна быть учтена в традиционном выражении для подвижности с тем, чтобы получить физически разумный результат для короткого проводника.

### 5.3. Баллистические MOSFET

Ранее в главах 1–3 транзисторы обсуждались как устройства с контролируемым барьером, в главе 4 изучалась электростатика структур MOS и транспорт электронов в таких структурах. Теперь мы намерены объединить все полученные сведения вместе и построить физическую модель процессов в нанотранзисторах. Начнем мы с баллистических транзисторов. Реальные короткоканальные MOSFET могут быть весьма сложными для исследования устройствами [14] и для детального понимания их работы необходимо как полуклассическое моделирование процессов в MOSFET [15], так и квантовомеханическое моделирование [16]. Наша задача – более скромная: понять основные физические принципы функционирования нанотранзисторов настолько, чтобы корректно интерпретировать экспериментальные данные и результаты численного моделирования как Si MOSFET и III–V MOSFET [17], так и нанопроводов и нанотрубок [18].

Примем к сведению, что зависимость  $Q(V_{GS}, V_{DS})$  известна как ниже, так и выше порогового напряжения, например, в виде полуэмпирического выражения (56/гл. 4). Сначала рассмотрим баллистический транспорт с  $T(E)=1$ , так что, согласно (2), для тока стока имеем:

$$I_{DS} = \frac{2q}{h} \int_{-\infty}^{+\infty} M(E)(f_S(E) - f_D(E))dE. \quad (57)$$



Когда напряжение на стоке велико, то  $f_s(E) \gg f_D(E)$ , и ток насыщения

$$I_{DSAT} = \frac{2q}{h} \int_{-\infty}^{+\infty} M(E) f_s(E) dE. \quad (58)$$

В линейной области ток стока небольшой,  $f_s(E) \approx f_D(E)$ , и ток в этих условиях можно определить из (10) в виде

$$I_{DLIN} = G_{ch} V_{DS} = \frac{V_{DS}}{R_{ch}}, \quad [A]$$

$$G_{ch} = \frac{1}{R_{ch}} = \frac{2q^2}{h} \int M(E) \left( -\frac{\partial f_0}{\partial E} \right) dE, \quad [C_M] \quad (59)$$

где  $G_{ch}$  – проводимость канала. Выполнив вычисления по этим уравнениям, получим баллистический ток в квазиравновесном режиме, найдем баллистический ток  $I_{ON}$ , а также баллистический ток от  $V_{DS} = 0$  до  $V_{DS} = V_{DD}$ . Однако, все приведенные выше уравнения были выведены для наноканала проводимости, прокомментированного и показанного на рис. 1. Какие у нас есть основания рассматривать нанотранзистор как электронное устройство именно этого типа?

### *MOSFET как устройство с наноканалом проводимости*

Ранее было введено и подробно описано понятие наноканала проводимости (рис. 1). Рисунок 8 показывает как нанотранзистор MOSFET можно рассматривать как устройство с наноканалом проводимости. MOSFET использует напряжение на затворе для модуляции высоты энергетического барьера. На рис. 8 показан ход  $E_c(x) \propto x$  от истока до стока с соответствующими значениями электрохимических потенциалов  $E_{FS}$  и  $E_{FD}$ . Ток стока определяется высотой барьера и коэффициентом прохождения через небольшую область длины  $\ell < L$  на вершине барьера. Если электроны, инжектированные истоком, рассеяться назад в этой небольшой области (узкое место наноканала проводимости), то они вернуться назад в исток и не дадут вклада в ток стока. Если же электроны пройдут эту небольшую область, то почти наверняка они достигнут стока. Уверенность в таком развитии событий обусловлена сильным электрическим полем со стороны стока, которое просто пронесет электроны в контакт стока. Эта область высокого электрического поля работает как коллектор электронов. Предполагается, что коэффициент

прохождения этой небольшой области длины  $\ell$  на вершине барьера равен единице. Проводимость этого узкого места баллистическая, тогда как проводимость всего канала может не быть баллистической.

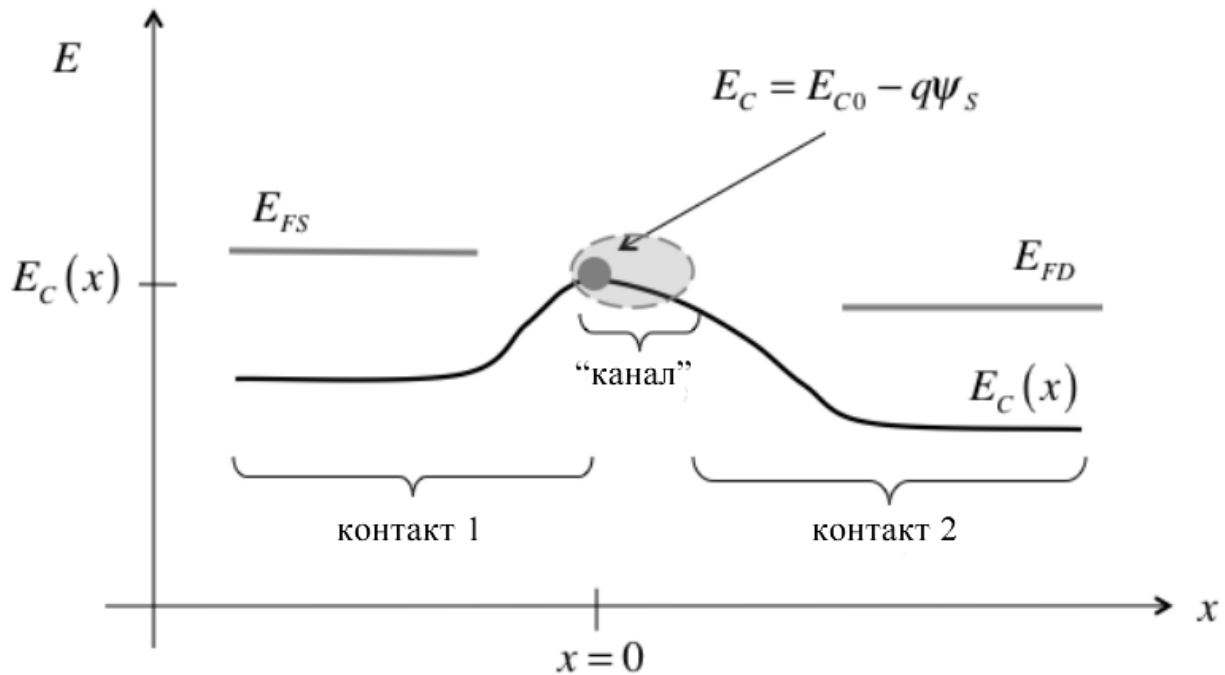


Рис. 8. Иллюстрация того, как MOSFET можно рассматривать как наноканал проводимости типа показанного на рис. 1. Небольшая область длины  $\ell$  является критическим узким местом для потока электронов, начинающегося с вершины барьера. Эта небольшая область и есть наноканал проводимости.

Применяя подход, развитый Ландауэром – Датта – Лундстроном применительно к MOSFET, нужно вычислять плотность электронов на вершине барьера из локальной плотности состояний на вершине  $LDOS = D_{2D}(E)|_{x=0}$ , а ток вычислять из распределения числа мод  $M(E)|_{x=0}$  на вершине барьера с учетом коэффициента прохождения  $T(E)$  через критическую область длины  $\ell$ . Мы не будем количественно определять ход зависимости  $E_C(x)$  и тем самым пространственно разделять вычисления, как это можно было бы сделать путем численного решения транспортных уравнений дрейфа-диффузии, транспортного уравнения Больцмана или квантовых уравнений совместно с уравнением Пуассона. Такие модельные расчеты необходимы при инженерном проектировании современных нанотранзисторов, мы же хотим понять физическую суть процессов в нанотранзисторах, сосредоточившись на критически важной небольшой области на вершине барьера.

### Линейная область

Для вычисления тока в квази-равновесной линейной области выходной характеристики воспользуемся уравнением (59) для баллистической проводимости. Распределение мод в (59) вычисляем по уравнениям (35) и (32):

$$M(E) = W \cdot M_{2D}(E) = W \cdot g_v \frac{\sqrt{2m^*[E - E_C(0)]}}{\pi\hbar}, \quad (60)$$

где  $E_C(0)$  – дно первой подзоны проводимости. Для функции Ферми в нашем случае  $E_F \approx E_{FS} \approx E_{FD}$ . После вычисления интеграла получим результат, похожий на (47) с заменой коэффициента прохождения  $\lambda_0/L$  на единицу, поскольку мы рассматриваем баллистический предел, а именно:

$$I_{DLIN} = G_{ch} V_{DS} = \left[ W \frac{2q^2}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{2\pi\hbar} \right) \mathfrak{F}_{-1/2}(\eta_F) \right] V_{DS}, \quad (61)$$

$$\eta_F = \frac{E_{FS} - E_C(0)}{kT}, \quad (62)$$

где  $E_C(0)$  – дно зоны проводимости на вершине барьера.

Уравнение (61) справедливо для линейной области баллистического MOSFET, хотя внешне оно мало похоже на привычное традиционное выражение типа (5/гл. 2)

$$I_{DLIN} = \frac{W}{L} |Q(V_{GS})| \mu V_{DS}. \quad (63)$$

Соответствие между традиционной и баллистической моделями MOSFET обсудим позже.

### Область насыщения

Ток в области насыщения вычисляем аналогичным образом по (58):

$$I_{DSAT} = W \frac{2q}{h} \left( \frac{g_v \sqrt{2m^* kT}}{\pi\hbar} \right) kT \frac{\sqrt{\pi}}{2} \mathfrak{F}_{1/2}(\eta_F). \quad (64)$$

Уравнение (64) справедливо для области насыщения баллистического MOSFET, хотя внешне оно мало похоже на привычное традиционное выражение типа (7/гл. 2)

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| v_{sat}. \quad (65)$$

Позже мы вернемся к обсуждению соответствия между традиционной и баллистической моделями MOSFET.

*От линейной области к насыщению*

Выше мы получили порознь выражения для тока в линейной области (низкие значения  $V_{DS}$ ) и в области насыщения (высокие значения  $V_{DS}$ ). Модель виртуального стока VS описывает всю область изменения  $V_{DS}$  благодаря использованию эмпирической функции насыщения тока стока (45/гл. 4), объединяющей эти обе области выходных характеристик MOSFET. Такой подход обсудим позже. В случае же баллистических MOSFET несложно получить общее выражение для тока от низких и до высоких значений  $V_{DS}$ .

Баллистический ток стока при произвольном значении напряжения на стоке вычисляем аналогичным образом по уравнению (57):

$$I_{DS} = W \frac{q}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{\pi \hbar} \right) kT [\mathfrak{F}_{1/2}(\eta_{FS}) - \mathfrak{F}_{1/2}(\eta_{FD})], \quad (66)$$

где

$$\eta_{FS} = [E_{FS} - E_C(0)]/kT, \quad (67)$$

$$\eta_{FD} = [E_{FD} - E_C(0)]/kT = [E_{FS} - qV_{DS} - E_C(0)]/kT.$$

*Выражения для тока через электронный заряд*

Выражение для тока (66) справедливо для баллистических MOSFET при произвольном значении  $V_{DS}$ , но в этом выражении не фигурирует инверсионный заряд  $Q$ . Для вычисления заряда нужно учесть электроны, инжектированные истоком с положительно направленными скоростями  $+v_x$  и заселяющими состояния на вершине барьера, а также учесть электроны, инжектированные стоком с отрицательно направленными скоростями  $-v_x$  и также заселяющими состояния на вершине барьера. Для произвольного значения  $V_{DS}$  из (39) находим инверсионный заряд

$$Q = -qn_S = -q \frac{N_{2D}}{2} [\mathfrak{F}_0(\eta_{FS}) + \mathfrak{F}_0(\eta_{FD})]. \quad (68)$$

Из (66) и (68) для тока и заряда имеем

$$I_{DS} = W |Q(V_{GS}, V_{DS})| v_{inj}^{ball} \left[ \frac{1 - \mathfrak{F}_{1/2}(\eta_{FD}) / \mathfrak{F}_{1/2}(\eta_{FS})}{1 + \mathfrak{F}_0(\eta_{FD}) / \mathfrak{F}_0(\eta_{FS})} \right],$$

$$Q(V_{GS}, V_{DS}) = -q \frac{N_{2D}}{2} [\mathfrak{F}_0(\eta_{FS}) + \mathfrak{F}_0(\eta_{FD})], \quad (69)$$

$$v_{inj}^{ball} = \langle \langle v_x^+ \rangle \rangle = v_T \frac{\mathfrak{F}_{1/2}(\eta_{FS})}{\mathfrak{F}_0(\eta_{FS})} = \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{F}_{1/2}(\eta_{FS})}{\mathfrak{F}_0(\eta_{FS})},$$

$$\eta_{FD} = \eta_{FS} - qV_{DS} / kT.$$

Это уравнение и есть уравнение выходной характеристики для баллистических MOSFET для всей области изменения напряжения на стоке. Уравнения такого типа были впервые получены в [19] и позже уточнены в [20].

Процедура вычисления выходной характеристики для баллистических MOSFET может быть следующей. Сначала вычисляем электронный заряд  $Q(V_{GS}, V_{DS})$  из уравнений электростатики MOS, например, из полуэмпирического уравнения (56/гл. 4). Задавшись вычисленным значением заряда, определяем положение уровня Ферми истока из второго уравнения в (69). Затем определяем баллистическую скорость из третьего уравнения. Наконец, находим ток в каждой точке  $(V_{GS}, V_{DS})$  по первому уравнению в (69). Вычисленные таким образом характеристики для параметров ETSOI  $n$ -MOSFET из [21] приведены на рис. 9.

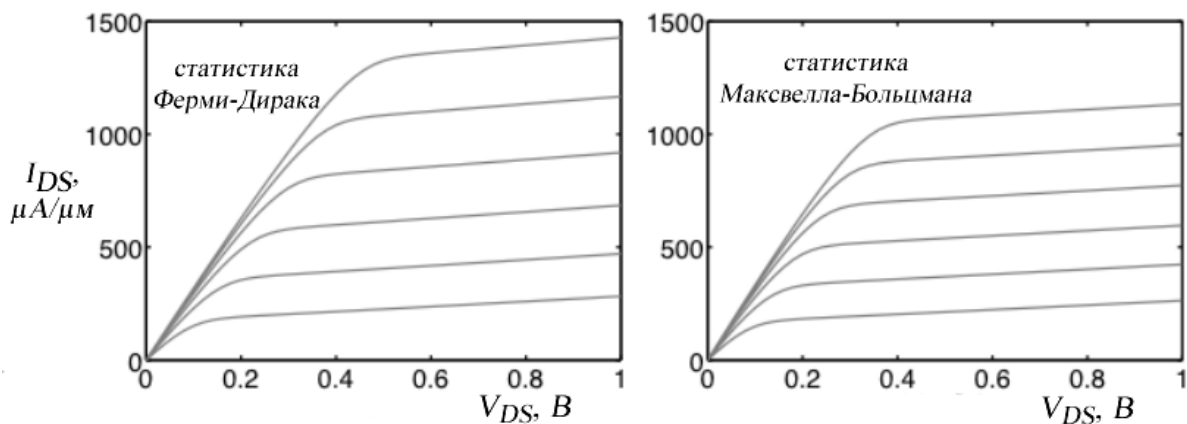


Рис. 9. Моделирование выходных характеристик баллистических MOSFET [6]. Использовались реальные параметры ETSOI  $n$ -MOSFET из [21]. Значение  $I_{OFF}$  было выбрано равным  $100 \text{ нА}/\mu\text{м}$ , что привело к  $V_T = 0.44 \text{ В}$ . Учитывались терминальные сопротивления  $R_{SD} = R_S + R_D = 260 \Omega \cdot \mu\text{м}$ , а  $V_{GS} = 0.5(0.1)1.0 \text{ В}$ .

Покажем, что общее баллистическое уравнение (69) дает правильные значения токов  $I_{DLIN}$  (61) и  $I_{DSAT}$  (64).

Сначала рассмотрим линейную область. Поскольку  $V_{DS}$  мало, т. е.  $\eta_{FS} \approx \eta_{FD}$ , то знаменатель в (69) равен двойке:

$$I_{DLIN}^{ball} = W |Q(V_{GS}, V_{DS})| v_{inj}^{ball} \left[ \frac{1 - \mathfrak{F}_{1/2}(\eta_{FD}) / \mathfrak{F}_{1/2}(\eta_{FS})}{2} \right].$$

После умножения и деления на  $\mathfrak{F}_{1/2}(\eta_{FS})$  получим

$$I_{DLIN}^{ball} = W |Q(V_{GS}, V_{DS})| v_{inj}^{ball} \frac{1}{\mathfrak{F}_{1/2}(\eta_{FS})} \left[ \frac{\mathfrak{F}_{1/2}(\eta_{FS}) - \mathfrak{F}_{1/2}(\eta_{FD})}{2} \right].$$

Далее умножаем и делим на  $\eta_{FS} - \eta_{FD} = qV_{DS}/kT$ :

$$I_{DLIN}^{ball} = W |Q(V_{GS}, V_{DS})| \frac{v_{inj}^{ball}}{2kT/q} \frac{1}{\mathfrak{F}_{1/2}(\eta_{FS})} \left[ \frac{\mathfrak{F}_{1/2}(\eta_{FS}) - \mathfrak{F}_{1/2}(\eta_{FD})}{\eta_{FS} - \eta_{FD}} \right] V_{DS}.$$

Поскольку разница  $\eta_{FS} - \eta_{FD}$  незначительная, то выражение в квадратных скобках есть производная от интеграла Ферми – Дирака [4, 10], т. е.

$$\left[ \frac{\mathfrak{F}_{1/2}(\eta_{FS}) - \mathfrak{F}_{1/2}(\eta_{FD})}{\eta_{FS} - \eta_{FD}} \right] \approx \frac{\partial \mathfrak{F}_{1/2}(\eta_{FS})}{\partial \eta_{FS}} = \mathfrak{F}_{-1/2}(\eta_{FS}),$$

так что

$$I_{DLIN}^{ball} = W |Q(V_{GS})| \frac{v_{inj}^{ball}}{2kT/q} \frac{\mathfrak{F}_{-1/2}(\eta_{FS})}{\mathfrak{F}_{1/2}(\eta_{FS})} V_{DS}, \quad (70)$$

что совпадает с (61).

В области насыщения напряжение на стоке велико, так что  $\eta_{FD} \ll 0$  и все интегралы с аргументом  $\eta_{FD}$  сводятся к экспонентам. Тогда вместо (69) имеем:

$$I_{DSAT}^{ball} = W |Q(V_{GS}, V_{DS})| v_{inj}^{ball} \left[ \frac{1 - e^{\eta_{FS} - qV_{DS}/kT} / \mathfrak{F}_{1/2}(\eta_{FS})}{1 + e^{\eta_{FS} - qV_{DS}/kT} / \mathfrak{F}_0(\eta_{FS})} \right].$$

При больших значениях  $V_{DS}$  дробь в квадратных скобках стремится к единице, так что для баллистического тока насыщения имеем

$$I_{DSAT}^{ball} = W |Q(V_{GS}, V_{DS})| v_{inj}^{ball},$$

что совпадает с (64).

Далее получим выражения для выходных характеристик, аналогичные (69), но для nanowire (NW) MOSFET [22]. В таких транзисторах 1D канал проводимости (проволока/NanoWire) окружен сплошным затвором (gate-all-around/GAA), что обеспечивает более надежное управление электростатикой, понижает *DIBL* и благоприятствует скейлингу. Предположим, что диаметр нанопроволоки достаточно мал настолько, что электроны движутся в 1D канале лишь с одной нижней занятой подзоной. Наша задача – получить выходную характеристику для 1D MOSFET и сравнить их с таковой для 2D (69).

Как и для 2D MOSFET, начинаем с уравнения (57), но вместо (60) для  $M(E)$  нам нужно распределение мод для 1D канала. Согласно (29), для 1D канала имеем [1 – 5]

$$M(E) = M_{1D}(E) = \frac{\hbar}{4} \langle v_x^+(E) \rangle D_{1D}(E),$$

где плотность состояний [2, 4, 18]

$$D_{1D}(E) = g_v \frac{\sqrt{2m^*}}{\pi\hbar} \frac{1}{\sqrt{E - E_C}}, \quad (71)$$

где  $E_C$  есть дно нижней подзоны проводимости.

Усреднения по углам для 1D канала не требуется, так что

$$\langle v_x^+(E) \rangle = v(E).$$

Окончательно для распределения мод имеем:

$$M_{1D}(E) = \begin{cases} 0, & (E < E_C) \\ g_v, & (E > E_C) \end{cases} \quad (72)$$

т. е. распределение мод для 1D канала есть константа при  $E > E_C$  [2 – 4].

Интегрируя (57) вместе с (72), получаем 1D аналог уравнения (66) для 2D канала, а именно:

$$I_{DS} = \frac{q}{h} kT [\mathfrak{I}_0(\eta_{FS}) - \mathfrak{I}_0(\eta_{FD})].$$

Теперь нам нужно ток стока выразить через электронный заряд. Для этого воспользуемся уравнением типа (39), но для 1D плотности состояний, а именно:

$$n_L = \frac{N_{1D}}{2} [\mathfrak{I}_{-1/2}(\eta_{FS}) + \mathfrak{I}_{-1/2}(\eta_{FD})], \quad m^{-1} \quad (73)$$

где эффективная 1D плотность состояний

$$N_{1D} = \sqrt{\frac{2m^*kT}{\pi\hbar^2}} \cdot M^{-1}. \quad (74)$$

Из этих двух уравнений находим удельный электронный заряд как 1D аналог уравнения (68) для 2D канала:

$$Q = -qn_L = -q \frac{N_{1D}}{2} [\mathfrak{F}_{-1/2}(\eta_{FS}) + \mathfrak{F}_{-1/2}(\eta_{FD})]. \quad [K/M]$$

Осталось выразить ток стока через электронный заряд. По аналогии с (69) для 2D канала, имеем:

$$I_{DS}^{1D} = |Q(V_{GS}, V_{DS})| v_{inj}^{ball} \left[ \frac{1 - \mathfrak{F}_0(\eta_{FD}) / \mathfrak{F}_0(\eta_{FS})}{1 + \mathfrak{F}_{-1/2}(\eta_{FD}) / \mathfrak{F}_{-1/2}(\eta_{FS})} \right],$$

$$Q(V_{GS}, V_{DS}) = -q \frac{N_{1D}}{2} [\mathfrak{F}_{-1/2}(\eta_{FS}) + \mathfrak{F}_{1/2}(\eta_{FD})], \quad (75)$$

$$v_{inj}^{ball} = \langle \langle v_x^+ \rangle \rangle = v_T \frac{\mathfrak{F}_0(\eta_{FS})}{\mathfrak{F}_{-1/2}(\eta_{FS})} = \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{F}_0(\eta_{FS})}{\mathfrak{F}_{-1/2}(\eta_{FS})},$$

$$\eta_{FD} = \eta_{FS} - qV_{DS}/kT.$$

Обратим внимание на то обстоятельство, что однонаправленная тепловая скорость (17) в невырожденном пределе не зависит от размерности канала проводимости:  $\langle \langle v_x^+ \rangle \rangle = v_T$ , а в вырожденных условиях  $\langle \langle v_x^+ \rangle \rangle > v_T$ .

Остается еще обсудить каким образом вычислить заряд  $Q(V_{GS}, V_{DS})$ . Можно выписать выражения, аналогичные (108/гл. 3) для ETSOI MOSFET, или же если мы удовлетворены простым подходом выше порога, то заряд в единицах  $[K/M]$  можно взять из

$$Q = 0, \quad V_{GS} \leq V_T$$

$$Q = -C_{ins}(V_{GS} - V_T), \quad V_{GS} > V_T,$$

$$C_{ins} = \frac{2\pi\epsilon_{ins}}{\ln\left(\frac{2t_{ins} + t_{wire}}{t_{wire}}\right)}, \quad [\Phi/M]$$

где  $t_{wire}$  – диаметр нанопроволоки, а  $t_{ins}$  и  $\epsilon_{ins}$  – толщина и диэлектрическая проницаемость затвора ГАА.

Проведенные выше вычисления свидетельствуют о том, что вывод выражений для выходных характеристик NW MOSFET аналогичен с



получением уравнений для плоских MOSFET и окончательные выражения весьма схожи.

Итак, уравнение (69) описывает выходные характеристики баллистических MOSFET. Вспомним, что ток стока MOSFET всегда можно записать в виде произведения заряда на скорость

$$I_{DS} = W |Q(V_{GS}, V_{DS})| \langle v_x(V_{GS}, V_{DS}) \rangle \Big|_{x=0}. \quad (76)$$

Приравнявая (76) с током стока (69), получим выражение для средней скорости электронов на вершине барьера

$$\begin{aligned} \langle v_x(V_{GS}, V_{DS}) \rangle \Big|_{x=0} &= v_{inj}^{ball} \left[ \frac{1 - \mathfrak{F}_{1/2}(\eta_{FD}) / \mathfrak{F}_{1/2}(\eta_{FS})}{1 + \mathfrak{F}_0(\eta_{FD}) / \mathfrak{F}_0(\eta_{FS})} \right], \\ v_{inj}^{ball} &= \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{F}_{1/2}(\eta_{FS})}{\mathfrak{F}_0(\eta_{FS})}. \end{aligned} \quad (77)$$

Позже мы вернемся к этой скорости и объясним почему скорость насыщается при высоких напряжениях на стоке в баллистических MOSFET.

Интегралы Ферми – Дирака во всех этих выкладках создают видимость сложности полученных выражений и тем самым вуалируют фактическую простоту функционирования баллистических MOSFET. Рассмотрим, например, невырожденный случай, когда уравнения упрощаются. Для невырожденного полупроводника

$$\begin{aligned} E_F &\ll E_C, \\ \eta_F &= (E_F - E_C) / kT \ll 0, \end{aligned}$$

а интегралы Ферми – Дирака любого порядка  $j$  сводятся к экспонентам:

$$\mathfrak{F}_j(\eta_F) \rightarrow e^{\eta_F}.$$

Таким образом, в невырожденном пределе уравнение (69) упрощается до

$$\begin{aligned} I_{DS} &= W |Q(V_{GS}, V_{DS})| v_T \left( \frac{1 - e^{-qV_{DS}/kT}}{1 + e^{-qV_{DS}/kT}} \right), \\ v_T &= \sqrt{\frac{2kT}{\pi m^*}}. \end{aligned} \quad (78)$$

Уравнение (78) допускает простую физическую интерпретацию в терминах термоионной эмиссии поверх барьера (рис. 10).

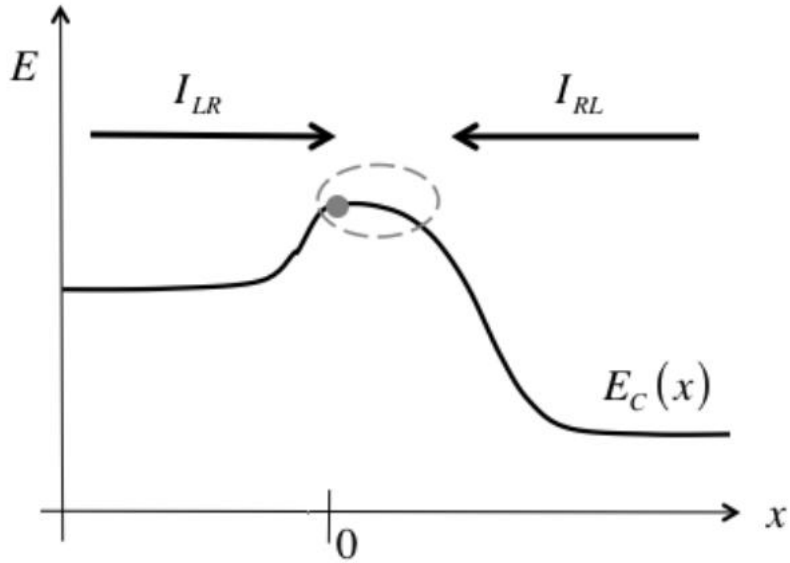


Рис. 10. Иллюстрация двух токов в баллистических MOSFET:  $I_{LR}$  со стороны истока и  $I_{RL}$  со стороны стока. Суммарный ток  $I_{DS} = I_{LR} - I_{RL}$ . В транзисторе хорошего качества электростатика MOS предполагает, что заряд на вершине барьера  $Q|_{x=0}$  должен быть независим от отношения этих двух токов.

Суммарный ток  $I_{DS}$  соответствует разности двух токов  $I_{LR}$  и  $I_{RL}$ . Простая трактовка термоионной эмиссии приводит к уравнению (20/гл. 1), идентичному уравнению (78). Выкладки, проведенные выше на основе подхода Ландауэра – Датта – Лундстрёма, показывают как вычислять  $v_T$  и допускают расширение на невырожденную статистику электронов, например, уравнение (78) по сравнению с (69). Ток стока насыщается когда  $I_{RL}$  становится незначительным по сравнению с  $I_{LR}$ . Это имеет место, когда  $V_{DS}$  становится больше нескольких  $kT/q$  в невырожденных условиях или при несколько бóльшем напряжении в случае статистики Ферми – Дирака (рис. 9).

Подведем итоги. Рассмотрена транспортная модель ЛДЛ, которая далее будет использоваться для построения теории MOSFET при низких и высоких напряжениях на стоке, в квазиравновесных и в далеких от равновесия условиях. Для достаточно длинных каналов проводимости результаты совпадут с привычными традиционными результатами, однако, мы также сможем достоверно построить физику нанотранзисторов, работающих в баллистическом или квазibalлистическом режимах.

Использовался подход ЛДЛ для расчета выходных характеристик баллистических MOSFET. Для этой цели применялось уравнение Ландауэра

(57) с ограничениями, которые накладывает электростатика MOS. В результате получена довольно простая модель баллистических MOSFET в виде уравнения (69). В случае невырожденной статистики эта модель упрощается до уравнения (78), аналогичного тому, которое уже получалось ранее в модели термоионной эмиссии.

Для MOSFET в подпороговом режиме можно пользоваться невырожденной статистикой и уравнением (78). В режиме выше порога зона проводимости на вершине барьера близка или даже лежит ниже уровня Ферми, так что нужно использовать уравнение (69). И тем не менее, стало обычной практикой в теории структур MOS предполагать невырожденную статистику Максвелла – Больцмана, поскольку использование ее просто упрощает расчеты и делает теорию более наглядной. Более того, на практике мы как правило не знаем значения некоторых параметров с нужной точностью, так что стало привычным пользоваться невырожденной статистикой с привлечением эмпирических параметров, чтобы вписаться в экспериментальные данные.

## Литература

1. S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport* (Singapore: World Scientific: 2012).
2. M. Lundstrom, C. Jeong, *Near-equilibrium transport. Fundamentals and Applications* (Singapore: World Scientific: 2013).
3. Ю. О. Кругляк, М. В. Стріха, *Укр. Фіз. Журн. Огляди*, **10**, 3 – 32 (2015).
4. Ю. А. Кругляк, *Нанoeлектроника «снизу – вверх»* (Одесса: ТЭС: 2015).
5. S. Datta, *Lessons from Nanoelectronics. Part A: Basic Concepts* (Singapore: World Scientific: 2017).
6. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
7. R. F. Pierret, *Advanced Semiconductor Fundamentals* (Upper Saddle River, N.J., USA: Prentice Hall: 2003).
8. R. Landauer, *IBM J. Res. Dev.*, **1**, № 3: 223 – 231 (1957).
9. Ю. А. Кругляк, *Термоэлектричество*, №6, 7 – 47 (2014).
10. R. Kim, M. Lundstrom, *Notes on Fermi – Dirac Integrals*, (West Lafayette, Indiana: Purdue University, USA); [www.nanohub.org/resources/5475](http://www.nanohub.org/resources/5475).
11. B. J. Van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, C. T. Foxon, *Phys. Rev. Lett.*, **60**, 848 – 850 (1988).
12. D. F. Holcomb, *Am. J. Phys.*, **67**, 278 – 297 (1999).
13. M. S. Shur, *IEEE Electron Device Lett.*, **23**, 511 – 513 (2002).
14. M. V. Fischetti, T. P. O'Regan, N. Sudarshan, C. Sachs, S. Jin, J. Kim, Y. Zhang, *IEEE Trans. Electron Dev.*, **54**: 2116 – 2136 (2007).
15. D. Frank, S. Laux, M. V. Fischetti, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 553 – 556 (1992).
16. Z. Ren, R. Venugopal, S. Goasguen, S. Datta, M. Lundstrom, *IEEE Trans. Electron Dev.*, **50**: 1914 – 1925 (2003).
17. Jesus A. del Alamo, *Nature*, **479**: 317 – 323 (2011).
18. M. Lundstrom, Jung Guo, *Nanoscale Transistors: Physics, Modeling, and Simulations* (New York: Springer: 2006).
19. K. Natori, *J. Appl. Phys.*, **76**: 4879 – 4890 (1994).
20. A. Rahman, J. Guo, S. Datta, M. Lundstrom, *IEEE Trans. Electron Dev.*, **50**: 1853 – 1864 (2003).
21. A. Majumdar, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **61**: 351 – 358 (2014).
22. Y. Guerfi, G. Larrieu, *Nanoscale Res. Lett.*, **11**, 210 – 217 (2016).



## Глава 6. Баллистическая скорость впрыскивания и объединение баллистической модели с моделью виртуального истока

### 6.1. Введение

Сначала введем понятие о баллистической скорости впрыскивания и изучим ее особенности, что далее позволит объединить баллистическую модель MOSFET с моделью виртуального истока.

### 6.2. Баллистическая скорость впрыскивания

Ток стока MOSFET есть произведение заряда на скорость в высшей точке барьера:

$$I_{DS} = W \left| \underline{Q(V_{GS}, V_{DS})} \right|_{x=0} \left| \langle v_x(V_{GS}, V_{DS}) \rangle \right|_{x=0}. \quad (1)$$

Из сравнения (1) с уравнением для баллистического тока стока (69/гл. 5) было получено выражение для средней скорости электронов на вершине барьера (77/гл. 5):

$$\begin{aligned} \left. \langle v_x(V_{GS}, V_{DS}) \rangle \right|_{x=0} &= \langle \langle v_x^+ \rangle \rangle \left[ \frac{1 - \mathfrak{I}_{1/2}(\eta_{FD}) / \mathfrak{I}_{1/2}(\eta_{FS})}{1 + \mathfrak{I}_0(\eta_{FD}) / \mathfrak{I}_0(\eta_{FS})} \right], \\ \langle \langle v_x^+ \rangle \rangle &= v_{inj}^{ball} = \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{I}_{1/2}(\eta_{FS})}{\mathfrak{I}_0(\eta_{FS})}. \end{aligned} \quad (2)$$

Уравнение (2) справедливо для 2D электронов в канале проводимости плоского MOSFET, тогда как для 1D электронов в канале проводимости NW MOSFET уравнение для скорости иное, а именно (75/гл. 5).

Средняя скорость на вершине барьера  $\left. \langle v_x(V_{GS}, V_{DS}) \rangle \right|_{x=0}$ , иначе скорость впрыскивания, зависит как от напряжения на затворе, так и от напряжения на стоке. Важно понять причину насыщения скорости при высоком напряжении на стоке в баллистических MOSFET и каким образом вычислить предельное значение скорости. Как показывают модельные расчеты баллистических MOSFET (рис. 9/гл.5), вычисленные вольт-амперные характеристики свидетельствуют о насыщении скорости – ток насыщения изменяется приблизительно линейно с величиной  $V_{GS} - V_T$ , однако, совершенно ясно, что причиной такого насыщения в баллистических MOSFET не может быть рассеяние, которое ограничивает скорость  $v_{sat}$ , как это подробно обсуждалось ранее в главе 2. Мы убедимся, что скорость действительно насыщается в баллистических MOSFET, однако, по причинам совершенно отличным от тех,

которые насыщают скорость в массивных полупроводниках под действием сильного электрического поля.

### Зависимость скорости от напряжения на стоке

Упростим общее выражение (2), перейдя к невырожденной статистике Максвелла – Больцмана:

$$\langle v_x(V_{DS}) \rangle|_{x=0} = v_T \left[ \frac{1 - e^{qV_{DS}/kT}}{1 + e^{qV_{DS}/kT}} \right], \quad (3)$$

$$v_T = \langle \langle v_x^+ \rangle \rangle = v_{inj}^{ball} = \sqrt{\frac{2kT}{\pi m^*}}.$$

Зависимость  $\langle v_x(V_{DS}) \rangle|_{x=0}$  от  $V_{DS}$  графически показана на рис. 1 вместе с ходом зонной энергии при низких и высоких значениях напряжения на стоке. При низких напряжениях скорость пропорциональна  $V_{DS}$ , а при высоких напряжениях скорость насыщается до значения  $v_T$ .

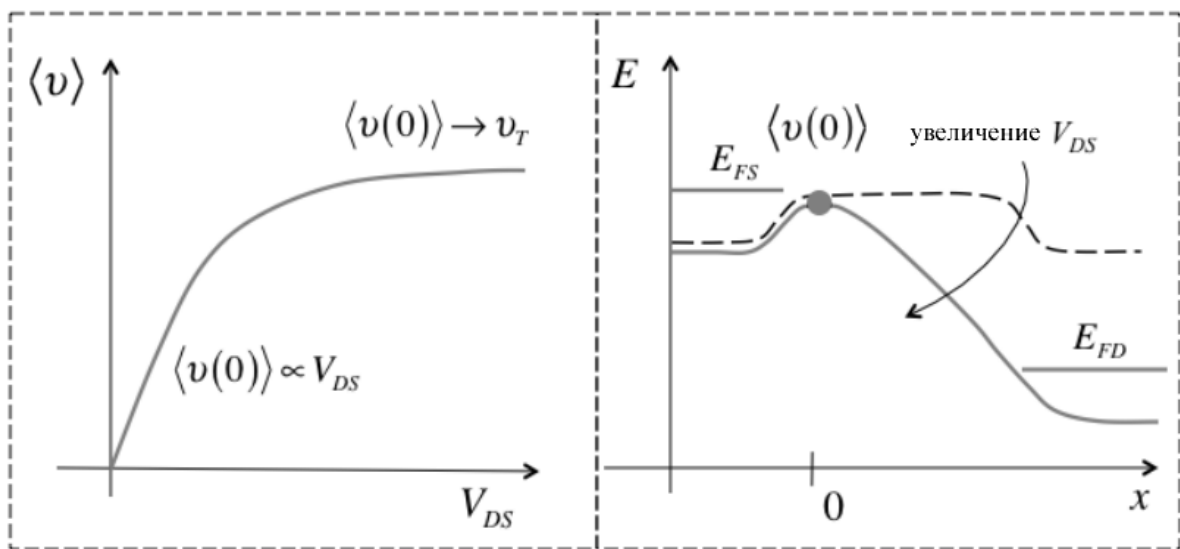


Рис. 1. Слева: зависимость средней скорости (3)  $\langle v_x(V_{DS}) \rangle|_{x=0} \equiv \langle v(0) \rangle$  от  $V_{DS}$ .

Справа: Иллюстрация изменения зонной энергии при низких и высоких значениях  $V_{DS}$ .

Казалось бы из рис. 1, что скорость в баллистическом MOSFET ведет себя так же, как и в массивном транзисторе: сначала при низком напряжении на стоке пропорционально  $V_{DS}$ , а затем насыщается при высоком напряжении. Заметим, однако, что речь идет о скорости на самой вершине барьера  $\langle v_x(V_{DS}) \rangle|_{x=0}$ : скорость в баллистическом MOSFET насыщается в канале

проводимости на самой границе истока  $S$ , на вершине барьера, где электрическое поле нулевое, а не на конечной границе стока канала проводимости, где электрическое поле очень велико.

Чтобы понять пропорциональность скорости напряжению  $V_{DS}$  при его малых значениях, разложим экспоненты в (3) при малых значениях показателей экспонент и получим:

$$\langle v_x(V_{DS}) \rangle|_{x=0} = \frac{v_T}{2kT/q} V_{DS}. \quad (4)$$

Умножим и разделим полученное выражение на длину канала проводимости:

$$\langle v_x(V_{DS}) \rangle|_{x=0} = \frac{v_T L}{2kT/q} \frac{V_{DS}}{L}. \quad (5)$$

В первом сомножителе узнаем баллистическую подвижность  $\mu_B$  (54/гл. 5), а второй сомножитель это электрическое поле в канале проводимости  $E_x = V_{DS}/L$ :

$$\langle v_x(V_{DS}) \rangle|_{x=0} = \mu_B E_x. \quad (6)$$

Таким образом, в баллистических MOSFET при малых напряжениях на стоке скорость записывается так же, как и для массивных проводников через  $\mu E_x$ , но только с заменой обычной подвижности  $\mu$  на баллистическую подвижность  $\mu_B$ .

### *Насыщение скорости в баллистических MOSFET*

Согласно (3), средняя однонаправленная скорость на вершине барьера насыщается при высоких значениях напряжения на стоке. Увидеть поточнее как это в действительности происходит могло бы помочь распределение скоростей в плоскости  $xu$  канала проводимости, в плоскости уходящей вглубь подложки вдоль канала проводимости. Сначала вспомним как скорости распределены в невырожденном массивном полупроводнике в состоянии равновесия. В этом случае функция Ферми упрощается до

$$f_0(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \rightarrow e^{(E_F-E)/kT}. \quad (7)$$

Для электронов в параболической зоне проводимости

$$E = E_C + \frac{m^* v^2}{2}, \quad (8)$$



так что невырожденная функция Ферми

$$f_0(v) = e^{(E_F - E_C)/kT} \times e^{-m^*v^2/2kT}. \quad (9)$$

Движение электронов в нашем случае происходит в плоскости  $xy$  с

$$v^2 = v_x^2 + v_y^2, \quad (10)$$

и невырожденная функция Ферми сводится к известному максвелловскому распределению скоростей

$$f_0(v_x, v_y) = e^{(E_F - E_C)/kT} \times e^{-m^*(v_x^2 + v_y^2)/2kT}. \quad (11)$$

Уравнение (11) дает распределение скоростей в невырожденном полупроводнике в состоянии равновесия и графически представлено на рис. 2. Как и ожидалось, все скорости в положительном направлении осей скомпенсированы скоростями в противоположном направлении, так что средняя скорость в состоянии равновесия равна нулю.

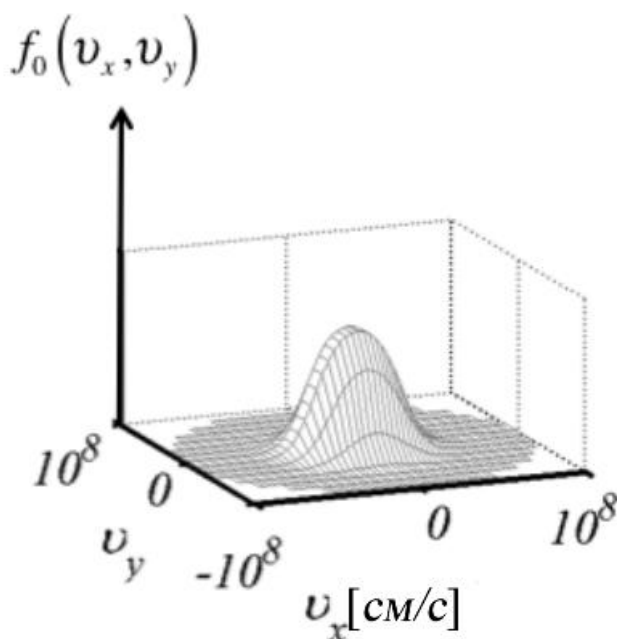


Рис. 2. Максвелловское распределение скоростей электронов в невырожденном полупроводнике в состоянии равновесия [1].

Можно ожидать, что в баллистических MOSFET при высоких напряжениях на стоке, далеких от состояния равновесия, распределение скоростей будет сильно отличаться от распределения в состоянии равновесия.

На рис. 3 показаны результаты численного решения транспортного уравнения Больцмана для 10-ти нм баллистического MOSFET.

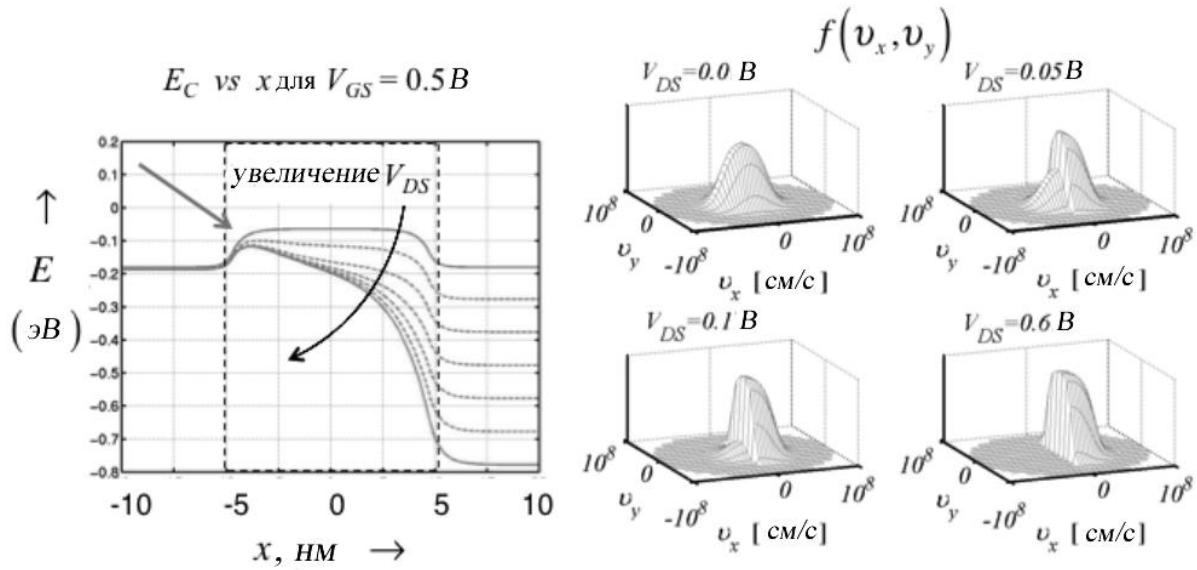


Рис. 3. Результаты численного моделирования баллистического MOSFET с длиной канала проводимости 10 нм. Слева: Ход  $E_C(x) \propto x$  при высоком напряжении на затворе и  $V_{DS} = 0.0(0.1)0.6 B$ . Справа: Распределение  $(v_x, v_y)$  на вершине барьера при разных значениях  $V_{DS}$  [1].

Напряжение на затворе велико, так что барьер на границе с истоком небольшой. Мы хотим понять показанное справа на этом рисунке распределение скоростей на вершине барьера при изменении напряжения на стоке от  $V_{DS} = 0$  до  $V_{DS} = V_{DD} = 0.6 B$ .

В баллистических MOSFET распределение скоростей на вершине барьера состоит из двух компонент – положительных скоростей электронов, впрыснутых истоком, и отрицательных скоростей электронов, впрыснутых стоком:

$$\begin{aligned} f^+(v_x > 0, v_y) &= e^{(E_{FS} - E_C(0))/kT} \times e^{-m^*(v_x^2 + v_y^2)/2kT}, \\ f^-(v_x < 0, v_y) &= e^{(E_{FD} - E_C(0))/kT} \times e^{-m^*(v_x^2 + v_y^2)/2kT}, \end{aligned} \quad (12)$$

где  $E_{FS}$  есть электрохимический потенциал истока,  $E_{FD} = E_{FS} - qV_{DS}$  это электрохимический потенциал стока. С ростом напряжения на стоке  $V_{DS}$  величина  $f^-(v_x, v_y)$  уменьшается.

На рис. 3/справа распределение скоростей показано при четырех различных напряжениях на стоке – от нулевого до  $V_{DD} = 0.6 B$ . Сначала рассмотрим ситуацию при  $V_{DS} = 0$  (рис. 2). В этом случае распределение

скоростей равновесное и  $\langle v(0) \rangle|_{x=0} = 0$ . Поскольку  $V_{DS} = 0$ , тока нет и MOSFET находится в состоянии равновесия, так что равновесное распределение скоростей на рис. 2 как бы очевидно. Однако, интересно каким же образом установилось это равновесие, наступающего обычно в результате электрон-фононного рассеяния, поскольку в баллистических MOSFET акты рассеяния не происходят. Ответ может быть следующим. На вершине барьера все электроны с  $v_x > 0$  инжектированы истоком, равновесие в котором поддерживается сильным электрон-фононным взаимодействием. На этой же вершине все электроны с  $v_x < 0$  инжектированы стоком, равновесие в котором также поддерживается сильным электрон-фононным взаимодействием. Поскольку при  $V_{DS} = 0$  электрохимические потенциалы одинаковы ( $E_{FD} = E_{FS}$ ), то положительные и отрицательные компоненты потоков электронов в локации вершины барьера одинаковы, результатом чего является равновесное максвелловское распределение скоростей электронов на вершине барьера (рис. 2).

Рассмотрим далее небольшое повышение напряжения на стоке до  $V_{DS} = 0.05V$  (рис. 3/справа). В этом случае величина компоненты с отрицательными скоростями в (12) меньше при сохранении величины компоненты с положительными скоростями, так что суммарная скорость в направлении  $v_x > 0$  положительная. Мы ранее уже видели, что при таких малых  $V_{DS}$  средняя скорость впрыскивания растет линейно с напряжением на стоке. При  $V_{DS} = 0.1V$  отрицательная компонента скорости еще меньше, так что суммарная скорость в направлении  $v_x > 0$  становится еще больше. Наконец, при  $V_{DS} = 0.6V$  отрицательная компонента скорости становится нулевой, поскольку электрохимический потенциал на стоке понизился настолько, что вероятность заселения вершины барьера электронами, прибывшими со стока, становится пренебрежимо малой. Положительная компонента скоростей электронов в направлении  $v_x > 0$  достигает максимально возможного значения. Дальнейшее увеличение напряжения на стоке не увеличит положительную компоненту: скорость достигла насыщения.

Графические результаты на рис. 3 объясняют поведение скорости на вершине барьера согласно уравнению (3), но есть одна тонкость, которая заслуживает обсуждения. Внимательное рассмотрение половинки максвелловского распределения при  $V_{DS} = 0.6V$  обнаруживает, что она больше и

выше аналогичной равновесной положительной половинки при  $V_{DS} = 0$ . Совершенно очевидно (рис. 3), что половинка максвелловского распределения, соответствующая  $v_x > 0$ , увеличивается с ростом  $V_{DS}$  хотя электрохимический потенциал истока  $E_{FS}$  не изменялся. Причина видна на левой части рис. 3: увеличение  $V_{DS}$  понижает вершину барьера  $E_c(0)$ . К такому результату приводит электростатика MOS у высококачественных транзисторов. У таких транзисторов электронный заряд на вершине барьера существенно или даже просто зависит только от напряжения на затворе и сколь-нибудь серьезно не изменяется с ростом напряжения на стоке (низкое значение *DIBL*). По мере того, как отрицательная компонента скорости электронов уменьшается с ростом  $V_{DS}$ , все больше электронов с положительной скоростью должны быть впрыснуты истоком чтобы сбалансировать электронный заряд. Поскольку электрохимический потенциал истока не изменяется, уравнение (12) свидетельствует, что  $E_c(0)$  должно уменьшиться с тем, чтобы увеличить электронный заряд за счет истока.

Наконец, отметим, что результирующий профиль распределения скоростей электронов при  $V_{DS} > 0$  (рис. 3) существенно отличается от равновесного профиля, но каждая из половинок распределения скоростей напоминает равновесный профиль. Лишь рассеяние могло бы вернуть результирующий профиль к равновесному виду, но такового в баллистических MOSFET нет. Баллистический канал проводимости очень далек от равновесия, но каждая из половинок распределения скоростей находится в равновесии с одним из двух контактов.

### *Баллистическая скорость впрыскивания*

Баллистическая скорость впрыскивания  $v_{inj}^{ball} = \langle\langle v_x^+ \rangle\rangle$  является важным параметром нанотранзистора; она играет роль  $v_{sat}$  в традиционной модели насыщения скорости. Как подробно рассматривалось выше,  $\langle\langle v_x^+ \rangle\rangle$  есть средняя скорость электронов, впрыснутых истоком в направлении  $v_x > 0$ , с максвелловским распределением скоростей (или ферми-дираковским) на вершине барьера при высоком напряжении на стоке. Это усредненная по углам скорость в направлении  $v_x > 0$  при определенной энергии  $E$ , а затем усредненная еще по энергии (2). Получено оно непрямым путем из выражения для тока, записанного в виде произведения заряда на скорость, а прямым путем получено ранее в виде (27/гл. 5). График зависимости баллистической скорости

впрыскивания от поверхностной плотности электронов в кремнии показан на рис. 4.

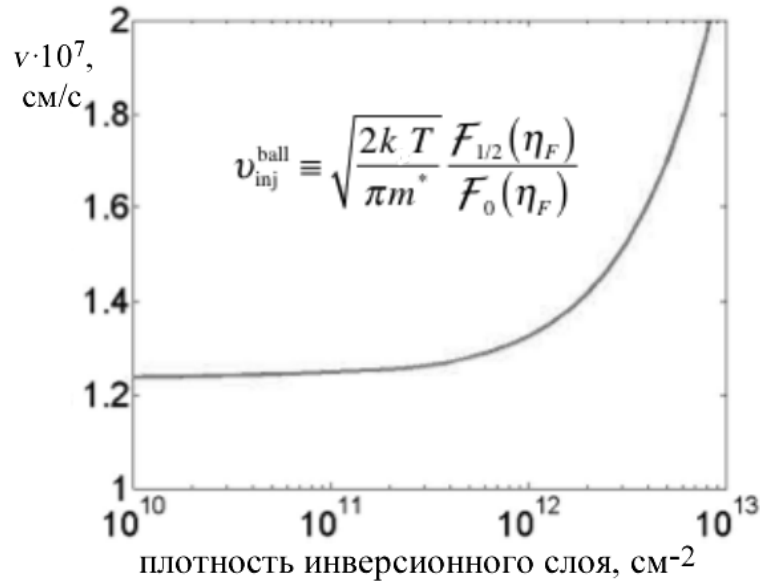


Рис. 4. Баллистическая скорость впрыскивания  $v_{inj}^{ball}$  в зависимости от поверхностной плотности 2D электронов при  $T = 300 K$  в DG SOI MOSFET [2].

Предполагалось, что только нижняя подзона проводимости занята, так что соответствующая эффективная масса  $m^* = 0.19m_0$  и долинное вырождение  $g_v = 2$ .

При высоком напряжении на стоке плотность инверсионного слоя на вершине барьера, согласно (40/гл. 5),

$$n_s|_{x=0} = \frac{N_{2D}}{2} \mathfrak{F}_0(\eta_F) = g_v \frac{m^* kT}{2\pi\hbar^2} \mathfrak{F}_0(\eta_F). \quad (13)$$

Для данного значения плотности решаем уравнение (13) относительно  $\eta_F$ , которое затем используется для вычисления  $v_{inj}^{ball}$  по (2). При  $T = 300 K$  2D плотность состояний для Si (100) имеет значение

$$N_{2D} = 2.05 \times 10^{12} \text{ см}^{-2}. \quad (14)$$

При  $n_s < N_{2D}$  полупроводник можно рассматривать невырожденным, а при  $n_s > N_{2D}$  важна уже статистика Ферми – Дирака.

Как видно из рис. 4, при  $n_s \ll 10^{12} \text{ см}^{-2}$  полупроводник невырожденный: интегралы Ферми – Дирака в (2) сводятся к экспонентам, так что

$$\langle\langle v_x^+ \rangle\rangle = v_{inj}^{ball} \rightarrow v_T = \sqrt{\frac{2kT}{\pi m^*}} = 1.2 \times 10^7 \text{ см/с}. \quad (15)$$

При  $n_s > 10^{12} \text{ см}^{-2}$  полупроводник становится вырожденным и  $v_{inj}^{ball}$  растет. Так получается когда состояния вблизи дна зоны проводимости заняты, энергия фермиевского уровня растет и заселяются состояния с большей скоростью. Увеличение скорости впрыскивания (рис. 4) объясняет почему выходные характеристики баллистических MOSFET на рис. 9/гл. 5 демонстрируют бóльшие токи для статистики Ферми – Дирака, поскольку именно она ответственна за бóльшие токи. Зависимость скорости впрыскивания от напряжения на затворе слабая, так что обе статистики – Максвелла – Больцмана и Ферми – Дирака на рис. 9/гл. 5 показывают насыщение тока, растущего почти линейно с  $V_{DS}$ . Таким образом, в обоих случаях мы можем сделать вывод о том, что речь идет о MOSFET с насыщением скорости. Как мы убедимся позже, рассеяние в реальных транзисторах уменьшает скорость впрыскивания, так что баллистическая скорость впрыскивания есть верхний предел для MOSFET.

Найдем баллистическую скорость впрыскивания в полностью вырожденном пределе. При  $T = 0 \text{ K}$   $f_0(E) = 1$  при  $E < E_F$  и  $f_0(E) = 0$  при  $E > E_F$ . Согласно (25/гл. 5) имеем:

$$v_{inj}^{ball} = \langle \langle v_x^+ \rangle \rangle = \frac{2}{\pi} \frac{\int_{E_C}^{\infty} \sqrt{2(E - E_C)/m^*} f_0(E) dE}{\int_{E_C}^{\infty} f_0(E) dE} \equiv \frac{2}{\pi} \frac{NUM}{DEN}.$$

Для числителя и знаменателя имеем:

$$NUM = \sqrt{\frac{2}{m^*}} \int_{E_C}^{E_F} \sqrt{E - E_C} dE = \sqrt{\frac{2}{m^*}} \left( \frac{2}{3} (E_F - E_C)^{3/2} \right), \quad DEN = E_F - E_C,$$

так что окончательно

$$v_{inj}^{ball} = \langle \langle v_x^+ \rangle \rangle = \frac{4}{3\pi} \sqrt{\frac{2}{m^*}} (E_F - E_C)^{1/2}.$$

Полученный результат удобно выразить через фермиевскую скорость, скорость электронов на фермиевском уровне энергии, исходя из соотношения

$$\frac{1}{2} m^* v_F^2 = E_F - E_C,$$

так что фермиевская скорость

$$v_F = \sqrt{\frac{2(E_F - E_C)}{m^*}}. \quad (16)$$

Окончательно, баллистическая скорость впрыскивания через фермиевскую скорость

$$v_{inj}^{ball} = \frac{4}{3\pi} v_F. \quad (17)$$

Как и ожидалось, баллистическая скорость впрыскивания меньше фермиевской скорости, поскольку последняя есть скорость, усредненная по всем электронам ниже уровня Ферми.

В заключение оценим баллистическую скорость впрыскивания реального MOSFET. Возьмем *n*-Si MOSFET при  $T = 300\text{ K}$  в состоянии  $I_{ON}$  с плотностью инверсионного слоя  $n_s = 1 \cdot 10^{13} \text{ см}^{-2}$ . Пусть в Si (100) занята лишь нижняя подзона проводимости. Какова баллистическая скорость впрыскивания?

Если полупроводник невырожден, что мало правдоподобно при таком большом значении  $n_s$ , получим уже известный результат (15). В нашем случае корректное выражение для баллистической скорости впрыскивания дается выражением (2). Однако, необходимо знать  $\eta_F$ , которое можно получить из (13), откуда

$$\eta_F = \log\left(e^{n_s/(N_{2D}/2)} - 1\right) = 9.76.$$

Подставляя полученное значение в (2), находим

$$v_{inj}^{ball} = \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{F}_{1/2}(\eta_{FS})}{\mathfrak{F}_0(\eta_{FS})} = 1.2 \times 10^7 \times \frac{\mathfrak{F}_{1/2}(9.76)}{\mathfrak{F}_0(9.76)} = 1.2 \times 10^7 \times \frac{23.2}{9.8} = 2.4 \times 10^7 \text{ см/с},$$

что в два раза превышает результат для невырожденной статистики.

Как показывают эти вычисления и рис. 4, переход к вырожденной статистике электронов существенно увеличивает баллистическую скорость впрыскивания. Для типичных Si MOSFET, однако, реальные баллистические скорости впрыскивания меньше, поскольку могут быть заселены несколько подзон проводимости и некоторым из них могут соответствовать бóльшие значения эффективных масс, а также в результате квантовых ограничений, которые увеличивают эффективные массы из-за непараболичности высоких подзон проводимости. Если нужны высокоточные количественные результаты, нужно обратить внимание на зонную структуру.

При вычислении баллистической скорости впрыскивания в Si (100) эффективная масса предполагалась равной  $0.19 m_0$ , а долинное вырождение бралось равным  $g_v = 2$ . У зоны проводимости кремния шесть эквивалентных

долин, а их эллипсоидальные поверхности постоянной энергии описываются эффективными массами  $m_l^* = 0.91m_0$  и  $m_t^* = 0.19m_0$ . Квантовые ограничения снимают вырождение. Две нижние подзоны вырождены с  $g_v = 2$  и  $m^* = 0.19m_0$ . При вычислении баллистической скорости впрыскивания было предположено, что заселена только нижняя нештрихованная подзона (рис. 16/гл.3), для которой эффективная масса  $m^* = m_l^*$  в направлении локализации, а эффективная масса в плоскости  $xu$  есть  $m^* = m_t^*$ . Если заселены более высокие подзоны, нужно учесть соответствующие эффективные массы вдоль осей  $x$  и  $y$ . Суммарная поверхностная плотность электронов учитывает вклады от всех заселенных подзон и баллистическая скорость впрыскивания есть скорость, усредненная по всем заселенным подзонам.

### 6.3. Объединение баллистической модели с моделью виртуального истока

Нам предстоит объединить баллистическую модель MOSFET (69/гл. 5) и модель виртуального истока (36/гл. 2).

Начнем как всегда с общего выражения для тока в виде произведения электронного заряда на скорость электронов:

$$I_{DS} = W \left| \frac{Q(V_{GS}, V_{DS})}{q} \right|_{x=0} \times v_x(V_{GS}, V_{DS}) \Big|_{x=0}. \quad (18)$$

Для вычисления заряда воспользуемся, например, полуэмпирическим выражением (46/гл. 4):

$$Q(V_{GS}, V_{DS}) = -m C_G^{inv} \left( \frac{kT}{q} \right) \ln \left( 1 + e^{q(V_{GS} - V_T)/mkT} \right), \quad (19)$$

$$V_T = V_{T0} - \delta V_{DS}.$$

Далее нужно определиться с вычислением средней скорости на вершине барьера, которая вычисляется по разному в баллистической модели и в модели виртуального истока.

#### *Резюме баллистической модели*

Ток в этой модели дается уравнением (18). Заряд определяется электростатикой MOS, например, уравнением (19). Для вычисления скорости нужно сначала определить положение уровня Ферми, если привлекается



статистика Максвелла – Больцмана. Энергия уровня Ферми определяется из выражения для инверсионного заряда (68/гл. 5):

$$Q(V_{GS}, V_{DS}) = -q \frac{N_{2D}}{2} [\mathfrak{F}_0(\eta_{FS}) + \mathfrak{F}_0(\eta_{FD})], \quad (20)$$

где

$$\eta_{FS} = (E_{FS} - E_C(0))/kT, \quad \eta_{FD} = \eta_{FS} - qV_{DS}/kT. \quad (21)$$

Далее определяем баллистическую скорость впрыскивания (2)

$$v_{inj}^{ball} = \sqrt{\frac{2kT}{\pi m^*} \frac{\mathfrak{F}_{1/2}(\eta_{FS})}{\mathfrak{F}_0(\eta_{FS})}}, \quad (22)$$

а затем среднюю скорость при заданных значениях напряжений на стоке и на затворе (77/гл. 5):

$$\langle v_x(V_{GS}, V_{DS}) \rangle \Big|_{x=0} = v_{inj}^{ball} \left[ \frac{1 - \mathfrak{F}_{1/2}(\eta_{FD}) / \mathfrak{F}_{1/2}(\eta_{FS})}{1 + \mathfrak{F}_0(\eta_{FD}) / \mathfrak{F}_0(\eta_{FS})} \right]. \quad (23)$$

Окончательно получаем ток стока в нужной точке  $(V_{GS}, V_{DS})$  выходной характеристики из (18). На практике важно учесть терминальные сопротивления. Именно по этому алгоритму были посчитаны характеристики на рис. 9/гл. 5.

### *Резюме модели виртуального стока*

Вычисление тока в этой модели также начинается с уравнений (18) и (19), но затем скорость вычисляется иначе, через функцию насыщения тока стока  $F_{SAT}$  согласно (33/гл. 2), а именно:

$$\langle v_x(V_{GS}, V_{DS}) \rangle \Big|_{x=0} = F_{SAT}(V_{DS}) v_{sat}, \quad (24)$$

$$F_{SAT}(V_{DS}) = \frac{V_{DS} N_{DSAT}}{\left[ 1 + (V_{DS} N_{DSAT})^\beta \right]^{1/\beta}}, \quad (25)$$

где

$$V_{DSAT} = \frac{v_{sat} L}{\mu}. \quad (26)$$

Итак, ток стока в модели виртуального истока в нужной точке  $(V_{GS}, V_{DS})$  выходной характеристики вычисляется по уравнению (18), используя для

вычисления заряда (19), а для вычисления скорости (24) – (26). Еще нужно учесть терминальные сопротивления.

Модель виртуального истока это полуэмпирическая модель, позволяющая гарантировано вписаться в экспериментальные данные. Используя эту модель, мы на самом деле лучше чувствуем физику процессов в MOSFET, поскольку параметры этой модели  $C_G^{inv}, V_T, m, \mu, v_{sat}, L$  имеют вполне конкретный физический смысл. Что касается параметра  $\beta$  в (25), то он изменяется в очень узких пределах для данного класса транзисторов. Чтобы вписаться в экспериментальные характеристики нанотранзисторов, такие параметры для микротранзисторов как  $\mu$  и  $v_{sat}$ , нужно адаптировать известным образом:

$$\begin{aligned} \mu &\rightarrow \mu_{app}, \\ v_{sat} &\rightarrow v_{inj}, \end{aligned} \quad (27)$$

что мы обсудим сейчас и убедимся в том, что кажущаяся подвижность  $\mu_{app}$  и скорость впрыскивания  $v_{inj}$  имеют вполне конкретный физический смысл.

### Объединенная модель

На рис. 5 показаны результаты расчета характеристик баллистических MOSFET, вычисленных по (78/гл. 5) в предположении статистики Максвелла – Больцмана. Параметры MOSFET взяты из [3], включая терминальные сопротивления. Модель виртуального истока по сути эмпирическая была вписана в эти расчетные характеристики MOSFET.

Подгонка параметров модели VS к расчетным характеристикам приводит к  $\mu_{app} = 654 \text{ см}^2/\text{В} \cdot \text{с}$  и  $v_{inj} = 1.24 \times 10^7 \text{ см}/\text{с}$ . Подгоночный параметр в (25)  $\beta$  оказался равным 2.9; характерные его значения лежат в пределах 1.6 – 2.0 для реальных MOSFET, работающих ниже баллистического предела. Физический смысл параметра  $\beta$  не ясен, он служит просто подгоночным параметром в функции насыщения тока стока  $F_{SAT}$  согласно (25) с целью описать переход от линейной области выходной характеристики к ее области насыщения. Параметры же  $\mu_{app}$  и  $v_{inj}$  имеют вполне ясный физический смысл. Чтобы убедиться в этом, нужно модель виртуального истока соотнести с транспортной моделью Ландауэра – Датта – Лундстрема. Сначала сравним токи в линейной части обеих моделей, затем перейдем к области насыщения и, наконец, получим цельные характеристики.

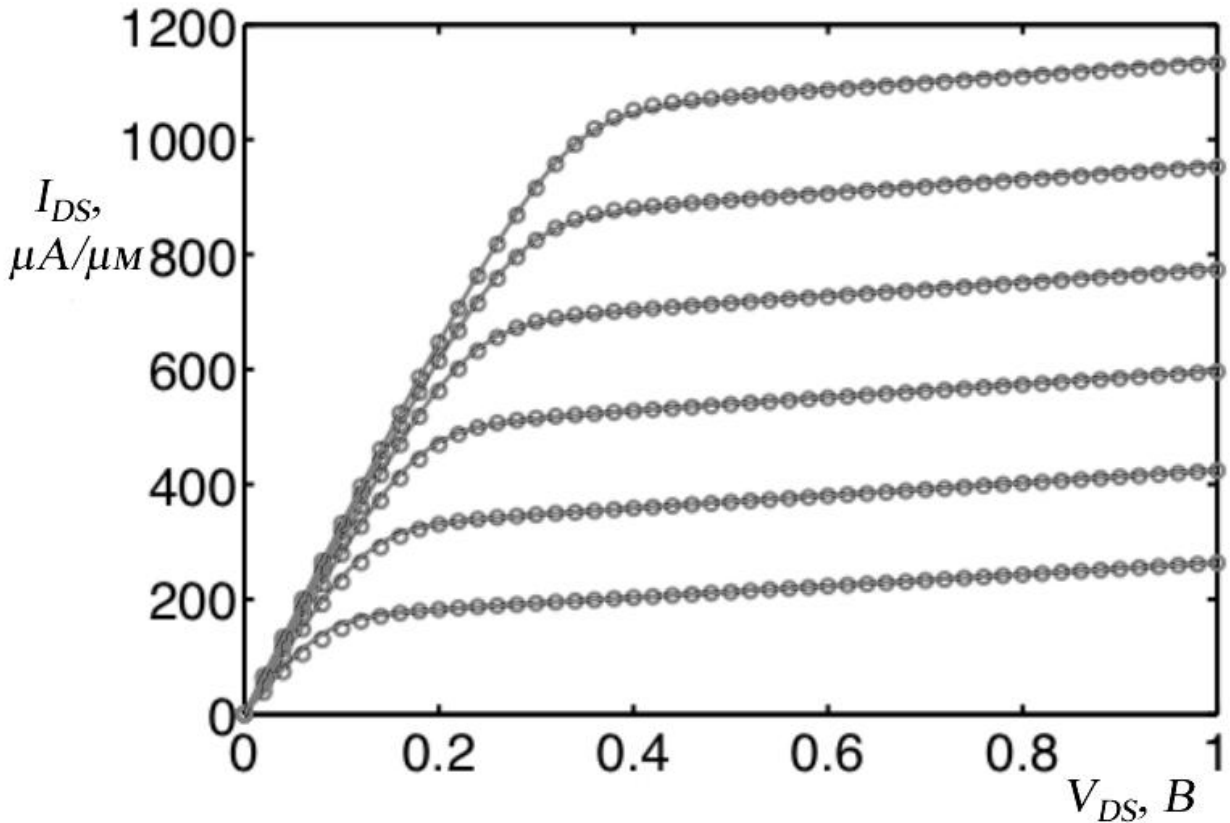


Рис. 5. Сплошными линиями показаны баллистические MOSFET согласно (78/гл. 5). Реальные параметры ETSOI MOSFET, включая терминальные сопротивления  $R_{SD} = R_S + R_D = 260 \Omega \cdot \mu\text{m}$ , взяты из [3]. Предполагалось, что  $I_{OFF} = 100 \text{ нА}/\mu\text{m}$ , что приводит к  $V_T = 0.44 \text{ В}$ . Напряжение на затворе  $V_{GS} = 0.5(0.1)1.0 \text{ В}$ . Результаты модели виртуального истока вписаны в расчетные баллистические характеристики (кружки) [4].

*Линейная область: баллистическая модель и модель виртуального истока*

Линейная область тока в баллистической модели дается уравнением (61/гл. 5):

$$I_{DLIN}^{ball} = \left[ W \frac{2q^2}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{2\pi\hbar} \right) \mathfrak{F}_{-1/2}(\eta_F) \right] V_{DS}, \quad (28)$$

$$\eta_F = \frac{E_{FS} - E_C(0)}{kT}, \quad (29)$$

где  $E_C(0)$  – дно зоны проводимости на вершине барьера.

Для небольших напряжений на стоке

$$F_{SAT} \rightarrow V_{DS} / V_{DSAT},$$

$$\underline{v_x(V_{GS}, V_{DS})}_{x=0} \rightarrow \mu V_{DS} / L.$$

Из (18) линейный ток в модели VS

$$I_{DLIN} = \frac{W}{L} |Q(V_{GS})| \mu V_{DS}, \quad (30)$$

что совпадает с результатом, вытекающим из традиционной теории MOSFET. Чтобы привести в соответствие ток в модели VS (30) с током в баллистической модели (28), нужно вместо традиционной подвижности  $\mu$  ввести в (28) кажущуюся подвижность  $\mu_{app}$  и тем самым уравнять эти оба выражения.

Только на первый взгляд кажется, что оба уравнения для тока (28) и (30) различны. Так, например, мы ожидаем, что квази-равновесный ток должен зависеть от заряда инверсионного слоя  $Q(V_{GS})$ , как того требует электростатика MOS. Это очевидно для традиционного выражения (30), но не сразу очевидно из выражения Ландауэра для баллистического тока (28). Обратите внимание на то, что величина электронного заряда  $Q$  определяет положение электрохимического потенциала  $\eta_F$  (29), а  $\eta_F$  фигурирует в (28), так что неявная зависимость тока от  $Q$  налицо в (28); необходимо только сделать эту зависимость явной.

В линейной области тока зависимость между инверсионным зарядом и электрохимическим потенциалом дается выражением:

$$Q = -qn_S = -qN_{2D} \mathfrak{F}_0(\eta_F) = -q \left( g_v \frac{m^* kT}{\pi \hbar^2} \right) \mathfrak{F}_0(\eta_F). \quad (31)$$

Это то же самое выражение, что и (68/гл. 5) с  $\eta_{FS} \approx \eta_{FD} = \eta_F$ , а  $N_{2D}$  есть 2D эффективная плотность состояний (38/гл. 5). Теперь мы можем записать баллистический ток  $I_{DLIN}$  как

$$I_{DLIN}^{ball} = Q \left[ \frac{G_{ch}}{Q} \right] V_{DS} = |Q| \left[ \frac{W \frac{2q^2}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{2\pi \hbar} \right) \mathfrak{F}_{-1/2}(\eta_F)}{q \left( g_v \frac{m^* kT}{\pi \hbar^2} \right) \mathfrak{F}_0(\eta_F)} \right] V_{DS}, \quad (32)$$

которое можно переписать следующим образом:

$$I_{DLIN}^{ball} = W |Q(V_{GS})| \left[ \frac{v_{inj}^{ball} \mathfrak{F}_{-1/2}(\eta_F)}{2(kT/q) \mathfrak{F}_{+1/2}(\eta_F)} \right] V_{DS}, \quad (33)$$

где  $v_{inj}^{ball}$  дается уравнением (22). Уравнение (33) идентично (28), в нем только зависимость от заряда показана в явном виде.

Уравнение (33) все еще внешне отличается от привычного выражения (30). Разделим и умножим его на  $L$ . Тогда

$$I_{DLIN}^{ball} = \frac{W}{L} |Q(V_{GS})| \left[ \frac{v_{inj}^{ball} L}{2(kT/q)} \frac{\mathfrak{S}_{-1/2}(\eta_F)}{\mathfrak{S}_{+1/2}(\eta_F)} \right] V_{DS}. \quad (34)$$

Размерность выражения в квадратных скобках есть размерность подвижности [ $m^2/B \cdot c$ ]. Это и есть обобщение подвижности по Шуру (54/гл. 5) на статистику Ферми – Дирака:

$$\mu_B \equiv \frac{v_{inj}^{ball} L}{2kT/q} \frac{\mathfrak{S}_{-1/2}(\eta_F)}{\mathfrak{S}_{+1/2}(\eta_F)}. \quad (35)$$

Наконец, линейный ток в баллистическом пределе

$$I_{DLIN}^{ball} = \frac{W}{L} |Q(V_{GS})| \mu_B V_{DS}, \quad (36)$$

что в точности совпадает с привычным выражением для квази-равновесного тока (30) с заменой лишь традиционной подвижности на баллистическую подвижность.

Итак, было показано, что баллистический ток в линейной области выходной характеристики (28) может быть записан традиционным образом как в модели виртуального истока (30), если только заменить привычную подвижность  $\mu$ , лимитированную рассеянием, на баллистическую подвижность  $\mu_B$  (35), как в уравнении (36). Физический смысл последней уже подробно обсуждался ранее.

*Область насыщения: баллистическая модель и модель виртуального истока*

Нам уже известно выражение тока насыщения в баллистическом пределе (64/гл. 5), а именно:

$$I_{DSAT}^{ball} = W \frac{2q}{h} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) kT \frac{\sqrt{\pi}}{2} \mathfrak{S}_{1/2}(\eta_F). \quad (37)$$

Оно выглядит совершенно иначе по сравнению с традиционным выражением для тока насыщения (32/гл. 2), а именно:

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| v_{sat}. \quad (38)$$

Эти оба выражения совмещаются, если под  $v_{sat}$  понимать  $v_{inj}$ . Действительно, мы вправе ожидать, что ток  $I_{ON}$  должен быть пропорционален заряду  $Q$ , что дает нам право переписать (37) в виде:

$$I_{DSAT}^{ball} = W|Q| \left[ \frac{\frac{2q}{h} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) kT \frac{\sqrt{\pi}}{2} \mathfrak{F}_{1/2}(\eta_F)}{Q} \right]. \quad (39)$$

Остается величину заряда в квадратных скобках выразить через  $\eta_F$ , как это мы уже делали в (31), и убедиться, что выражение в квадратных скобках есть ничто иное как  $v_{inj}^{ball} = \langle\langle v_x^+ \rangle\rangle$  (22). Тут есть, однако, одна тонкость, связанная с тем, что сейчас речь идет о больших напряжениях на стоке. В этом случае, как уже обсуждалось в связи с рис. 3, только половина состояний на вершине барьера заселена. Поэтому вместо (31) для заряда мы должны записать

$$Q = -q \frac{N_{2D}}{2} \mathfrak{F}_0(\eta_F) = -q \left( g_v \frac{m^* kT}{2\pi \hbar^2} \right) \mathfrak{F}_0(\eta_F). \quad (40)$$

Это вызвано тем, что электроны с положительными скоростями, инжектированные истоком, продолжают занимать соответствующие состояния на вершине барьера, тогда как аналогичные состояния на вершине барьера, соответствующие электронам с отрицательными скоростями, инжектированным стоком, остаются пустыми, поскольку эти электроны поступают со стока с низким уровнем Ферми, так что вероятность для электронов со стока иметь энергию выше, чем энергия на вершине барьера, пренебрежимо мала. Подставляя заряд по (40) вместо заряда в квадратных скобках выражения (39), окончательно получаем

$$I_{DSAT}^{ball} = W|Q| \langle\langle v_x^+ \rangle\rangle = W|Q| v_{inj}^{ball}, \quad (41)$$

где  $v_{inj}^{ball} = \langle\langle v_x^+ \rangle\rangle$  есть баллистическая скорость впрыскивания (22). Уравнение (41) идентично уравнению (37). В нем просто зависимость от заряда выписана в явном виде.

Итак, было показано, что баллистический ток в области насыщения (37) может быть записан традиционным образом (38), если скорость насыщения  $v_{sat}$ , лимитированную рассеянием, заменить на скорость впрыскивания  $v_{inj}$ , роль

которой играет баллистическая скорость впрыскивания  $v_{inj}^{ball} = \langle \langle v_x^+ \rangle \rangle$ , как в (41), которая есть ничто иное, как средняя тепловая скорость, с которой электроны впрыскиваются в канала проводимости истоком. Физика насыщения скорости в баллистических MOSFET обсуждалась ранее.

При совмещении модели виртуального истока с расчетными баллистическими характеристиками на рис. 5 параметры модели  $\mu_{app}$  и  $v_{inj}$  были подобраны так, чтобы наилучшим образом соответствовать расчетным данным. Вычислим оба параметра согласно (35) и (22) воспользовавшись данными для Si (100) с  $m^* = 0.19m_0$  и статистикой Максвелла – Больцмана. Получим значения

$$v_{inj}^{ball} = V_T = \sqrt{\frac{2kT}{\pi m^*}} = 1.2 \times 10^7 \text{ см/с},$$

$$\mu_B = \frac{v_{inj}^{ball} L}{2kT/q} = 692 \text{ см}^2/\text{В} \cdot \text{с},$$

которые довольно близки к подгоночным значениям  $v_{inj} = 1.24 \times 10^7 \text{ см/с}$  и  $\mu_{app} = 654 \text{ см}^2/\text{В} \cdot \text{с}$ .

### *Сравнение с экспериментальными результатами*

С целью проверки, насколько хорошо баллистическая модель MOSFET описывает реальные транзисторы, по этой модели были просчитаны следующие две ситуации: 30 нм ETSOI *n*-FET из [3] (рис. 6) и 30 нм III-V FET/HEMT из [5] (рис. 7). На каждом из этих рисунков приведена выходная характеристика, рассчитанная по баллистической модели в предположении статистики Максвелла – Больцмана, и экспериментальные данные, в которые вписаны результаты расчетов по модели виртуального истока.

Согласование экспериментальных данных с расчетными по модели виртуального истока позволило найти три важных параметра обоих транзисторов: 1) независимые от напряжения на затворе терминальные сопротивления; 2) кажущуюся подвижность; 3) скорость впрыскивания. Они собраны ниже в табл. 1. Там же приведены однонаправленные тепловые скорости  $v_T$ , вычисленные в предположении статистики Максвелла – Больцмана с  $m^* = 0.22m_0$  для Si и  $m^* = 0.016m_0$  для III-V HEMT, экспериментальные ограниченные диффузией подвижности  $\mu$  и баллистические подвижности  $\mu_B$ , вычисленные по (35) в предположении статистики Максвелла – Больцмана.

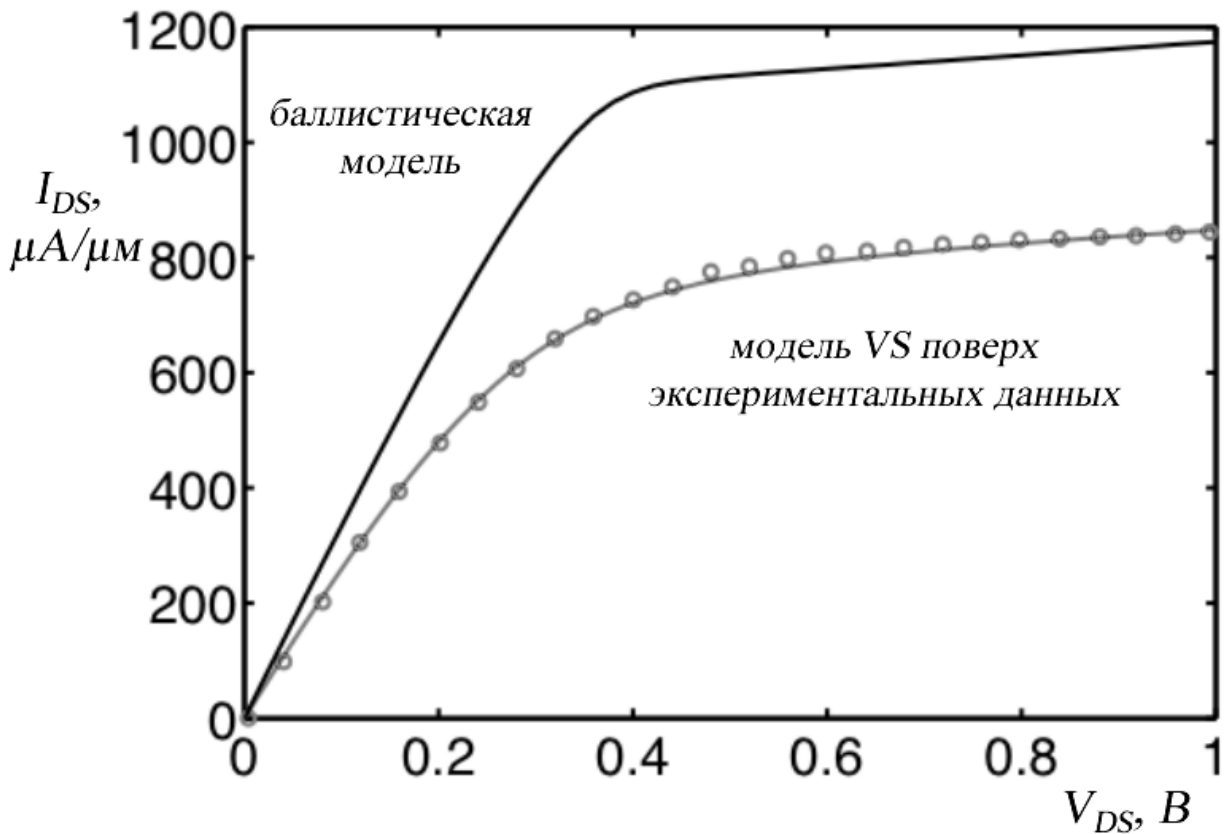


Рис. 6. Наверху: моделирование выходной характеристики баллистического ETSOI Si MOSFET: параметры, включая терминальные сопротивления, взяты из [3]; напряжение на затворе  $V_{GS} = 0.5V$ ; хотя это и  $n$ -MOSFET, пороговое напряжение меньше нуля, так что имеет место приличный ток при  $V_{GS} = 0V$ . Внизу: сплошной линией показана экспериментальная характеристика  $30\text{ нм}$  ETSOI  $n$ -FET [3] и вписанные в нее точки (кружочки) по модели VS [4].

Таблица 1  
Сравнение параметров Si ETSOI MOSFET и III-V HEMT

| СВ-во       | ETSOI Si FET       | III-V HEMT         | []                   |
|-------------|--------------------|--------------------|----------------------|
| $R_{SD}^*$  | 260                | 434                | $\Omega \cdot \mu m$ |
| $\mu_{app}$ | 220                | 1800               | $cm^2/B \cdot c$     |
| $v_{inj}$   | $0.82 \times 10^7$ | $3.85 \times 10^7$ | $cm/c$               |
| $v_T$       | $1.14 \times 10^7$ | $4.24 \times 10^7$ | $cm/c$               |
| $\mu$       | 350                | 12500              | $cm^2/B \cdot c$     |
| $\mu_B$     | 658                | 2446               | $cm^2/B \cdot c$     |

\*  $R_{SD} = R_S + R_D$



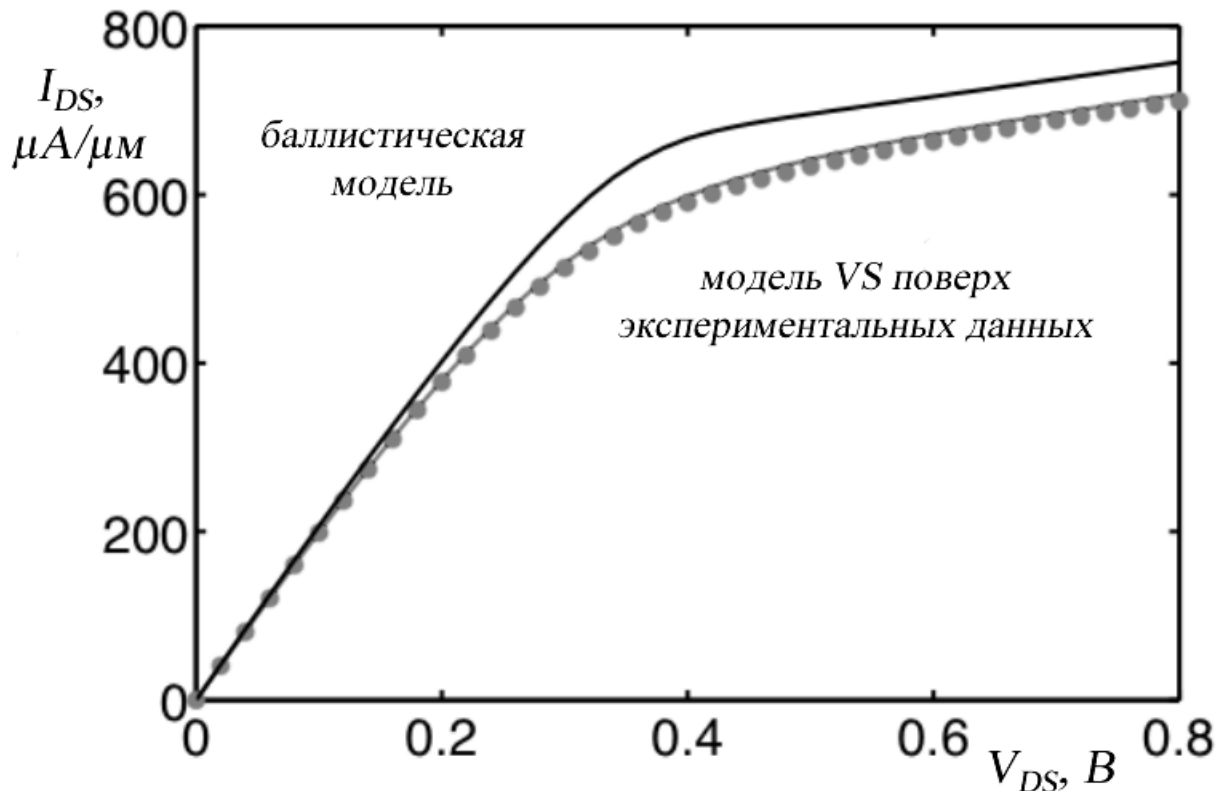


Рис. 7. Наверху: моделирование выходной характеристики баллистического III-V НЕМТ: параметры, включая терминальные сопротивления, взяты из [5]; напряжение на затворе  $V_{GS} = 0.5\text{ В}$ ; хотя это и  $n$ -НЕМТ, пороговое напряжение меньше нуля, так что имеет место приличный ток при  $V_{GS} = 0\text{ В}$ . Внизу: сплошной линией показана экспериментальная характеристика  $30\text{ нм } n$ -НЕМТ [5] и вписанные в нее точки (кружочки) по модели виртуального истока [4].

Кажущаяся подвижность  $\mu_{app}$  служила параметром подгонки модели виртуального истока к экспериментальным результатам. Видим, что в обоих случаях, ETSOI Si MOSFET и III-V FET,  $\mu_{app}$  меньше и подвижности  $\mu$ , ограниченной рассеянием, и баллистической подвижности  $\mu_B$ . Позже мы убедимся в том, что  $\mu_{app}$  не просто подгоночный параметр, а является хорошо определенным физическим параметром даже при наличии рассеяния.

Отношение измеренного значения  $I_{ON}^{meas}$  к вычисленному баллистическому значению  $I_{ON}^{ball}$  может служить мерой того, насколько близко к баллистическому пределу работает реальный транзистор.

Из графиков на рисунках 6 и 7 находим, что для ETSOI Si FET

$$B = \frac{I_{ON}^{meas}}{I_{ON}^{ball}} = 0.73,$$

а для III-V HEMT

$$B = \frac{I_{ON}^{meas}}{I_{ON}^{ball}} = 0.96.$$

Эти результаты свидетельствуют о том, что Si MOSFET работают весьма близко к баллистическому пределу, а для III-V HEMT характерен существенно баллистический режим. Обратим также внимание на то обстоятельство, что кажущаяся подвижность  $\mu_{app}$ , найденная из вписывания модели виртуального истока в экспериментальные результаты, сравнительно близка к традиционной подвижности  $\mu$  для Si FET, однако,  $\mu_{app} \ll \mu$  для III-V FET. Это также указывает на то, что Si FET работают ниже баллистического предела, а III-V FET – фактически на баллистическом пределе. Заметим также, что скорости впрыскивания  $v_{inj}$ , найденные из вписывания модели виртуального истока в экспериментальные результаты, меньше баллистических скоростей впрыскивания  $v_T$  и в тех и в других FET.

Еще одно замечание относительно использования статистики Максвелла – Больцмана в проведенном анализе данных. Выше порога более корректно использовать статистику Ферми – Дирака и учитывать такие усложнения как непараболичность подзон и вообще учет множественности подзон. Более тщательный анализ предполагает учет этих факторов, однако, сложившаяся практика привлечения статистики Максвелла – Больцмана к анализу экспериментальных данных, как правило, приводит к вполне физически правдоподобным результатам.

Итак, мы убедились, что результаты как в линейной области выходных характеристик, так и в области насыщения модели виртуального истока вполне совместимы с таковыми в баллистической модели. Мы теперь понимаем, почему традиционную подвижность для протяженных каналов проводимости нужно заменить на баллистическую подвижность, характерную для баллистического транспорта. Было также показано, что скорость насыщения в традиционной модели соответствует баллистической скорости впрыскивания в баллистической модели. Рисунки 5 – 7 показывают, что баллистическая модель предсказывает бóльшие токи по сравнению с экспериментальными данными, а

профиль баллистических характеристик  $I_{DS} \propto V_{DS}$  заметно отличается от экспериментальных характеристик: переход от линейной области к области насыщения происходит в меньшем диапазоне напряжений на стоке. Представляется, что профиль перехода от линейности к насыщению связан с зависимостью напряжения на стоке от рассеяния. Чтобы разобраться в этом, нам потребуется углубиться в теорию рассеяния электронов в полевых транзисторах. Лучшее понимание явлений рассеяния поможет нам объяснить почему скорость впрыскивания меньше баллистической скорости впрыскивания и как интерпретировать кажущуюся подвижность при наличии рассеяния.

Подведем итоги. В этой главе обсуждалась зависимость скорости электронов от напряжения на стоке в баллистических MOSFET, а также ее зависимость от напряжения на затворе и от инверсионного заряда. На первый взгляд кажется необычным, что в баллистических MOSFET скорость насыщается с ростом напряжения на стоке в отсутствие рассеяния электронов, однако, физика этого явления сейчас вполне понятна. В баллистических MOSFET скорость электронов насыщается не в стоковом конце канала проводимости, как в массивных транзисторах, где электрическое поле наибольшее и рассеяние интенсивное, а там, где заканчивается исток и начинается канал проводимости, т. е. на вершине барьера, где электрическое поле нулевое.

Мы также обсудили само по себе насыщение скорости, известное как баллистическая скорость впрыскивания. Именно эта скорость есть верхний предел скорости впрыскивания в реальных MOSFET. Если  $n_s \ll N_{2D}/2$ , то баллистическая скорость впрыскивания постоянная, однако, для  $n_s \gtrsim N_{2D}/2$  эта скорость растет с увеличением поверхностной плотности электронов. Прделаны простые расчеты баллистической скорости впрыскивания, которые могут послужить отправной точкой для более тщательных вычислений.

Показано как увязаны между собой баллистическая модель и модель виртуального истока. Простой заменой традиционной подвижности, лимитируемой рассеянием, в модели виртуального истока на баллистическую подвижность получаем правильный ход баллистического линейного тока. Заменой скорости насыщения  $v_{sat}$  в массивном канале проводимости на баллистическую скорость впрыскивания  $v_{inj}^{ball}$  получаем правильное значение баллистического тока  $I_{ON}$ . Также узнали, что баллистическая модель

предсказывает бóльшие токи по сравнению с экспериментальными данными. Это связано с эффектами рассеяния электронов, понимание которого исключительно важно для моделирования нанотранзисторов.

## Литература

1. J.-H. Rhew, Zhibin Ren, M. Lundstrom, *Solid-State Electron.*, **46**, 1899 – 1906 (2002).
2. M. Lundstrom, Zhibin Ren, *IEEE Trans. Electron Dev.*, **49**: 133 – 141 (2002).
3. A. Majumdar, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **61**: 351 – 358 (2014).
4. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
5. D. H. Kim, J. A. del Alamo, D. A. Antoniadis, B. Brar, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 861 – 864 (2009).

## Глава 7. Рассеяние электронов и модель прохождения MOSFET

### 1. Введение

При вычислении характеристик баллистических MOSFET мы начинали с уравнения (57/гл. 1), в котором предполагалось, что коэффициент прохождения  $T(E)$  равен единице. Рассеяние электронов на заряженных примесях, колебаниях решетки и тому подобные процессы уменьшают прохождение. Для вычисления характеристик транзисторов с учетом рассеяния в целом справедливо уравнение

$$I_{DS} = \frac{2q}{h} \int_{-\infty}^{+\infty} T(E)M(E)(f_S(E) - f_D(E))dE. \quad (1)$$

Траектория электрона при баллистическом переносе схематически показана на рис. 1 в сравнении с квазibalлистическим переносом, включающим множественные процессы рассеяния.

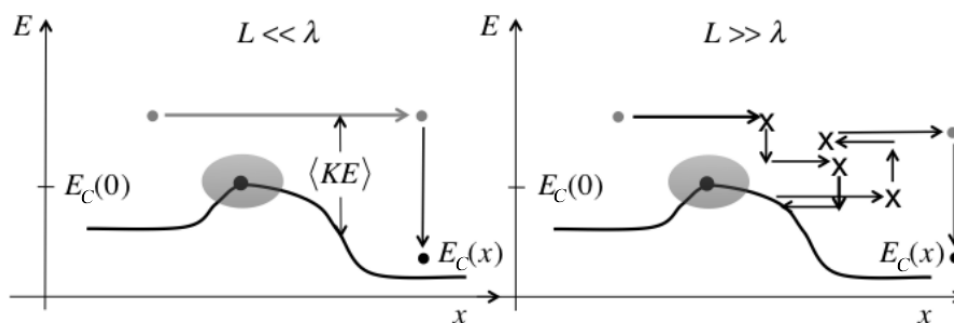


Рис. 1. Иллюстрация баллистического (слева) и квазibalлистического (справа) переноса электрона. В обоих случаях показаны траектории электрона, инжектированного истоком при определенной энергии  $E$ . Рассеяние – это стохастический процесс, так что траектория справа одна из множества возможных траекторий.

Как показано на левом рисунке 1, электроны впрыскиваются истоком, где они подвержены рассеянию, в канал проводимости, в котором не происходят процессы рассеяния, и входят в сток, где они опять подвержены рассеянию. Разница в электрохимических потенциалах на концах канала проводимости ускоряет электроны, они теряют кинетическую энергию  $KE$ , которая передается стоку. Это баллистический перенос.

Справа на рисунке 1 условно показана траектория переноса в присутствии актов рассеяния. Некоторые акты рассеяния упругие: изменяется направление движения, но не энергия. Другие акты неупругие: изменяются и направление и

энергия. Например, электроны могут приобрести энергию путем поглощения фононов (колебаний решетки) или могут передать энергию колебаниям решетки, генерируя фононы. В частности, для показанной справа траектории электрон, впрыснутый истоком, достигает стока, однако, рассеяние это стохастический процесс, и могут реализовываться траектории, когда инжектированный истоком электрон рассеивается назад и возвращается в исток. Коэффициент прохождения от истока до стока есть отношение потока электронов, впрыснутых истоком, к потоку электронов, покидающих канал проводимости через сток, с неизбежностью уменьшается из-за рассеяния.

Далее мы вначале качественно рассмотрим рассеяние электронов и введем понятие о коэффициенте прохождения, а затем построим модель прохождения MOSFET.

## 7.2. Рассеяние электронов и коэффициент прохождения

Наноразмерные MOSFET не являются ни чисто баллистическими ( $T(E)=1$ ), ни чисто диффузионными ( $T(E)\ll 1$ ): они работают в квазибаллистическом режиме с  $T(E) < 1$ . Наша задача – научиться пользоваться уравнением (1) в квазибаллистическом режиме. Более подробно с прохождением при рассеянии можно познакомиться в [1 – 4].

### *Характеристические времена и длины*

Удобный подход к пониманию рассеяния дают характеристические времена, такие как среднее время между столкновениями  $\tau$  или скорость рассеяния  $1/\tau$ , как вероятность акта рассеяния за единицу времени. Удобно ввести также характеристические длины, такие как средняя длина свободного пробега  $\Lambda$ , среднее расстояние между актами рассеяния, а  $1/\Lambda$  есть вероятность рассеяния на единицу длины. В общем случае эти характеристические времена и длины зависят от энергии электронов. Нас будут интересовать средние времена рассеяния и средние длины свободного пробега, когда усреднение проводится по физически существенному распределению энергий электронов.

Три важных характеристических времени проиллюстрированы на рис. 2. Рассмотрим поток электронов с квазиимпульсом  $\vec{p}(E) = p(E)\hat{x}$ , инжектированных в полупроводник в момент времени  $t=0$ . Будем считать, что энергия электронов  $E$  намного больше равновесной энергии  $3kT/2$ . По истечении времени  $\tau(E)$  пусть каждый электрон испытает в среднем хотя бы

один акт рассеяния. Величина  $\tau(E)$  есть среднее время рассеяния, а обратная величина  $1/\tau(E)$  – средняя скорость рассеяния. При этом предполагается, что все состояния, в которые рассеиваются электроны, свободны и отсутствуют электроны, привнесенные из других состояний. Более точно величину  $\tau(E)$  можно назвать средним временем рассеяния электронов наружу (out-scattering time).

Как показано на рис. 2, полезно определить и другие характеристические времена. Например, преимущественным механизмом рассеяния может быть упругое и анизотропное рассеяние, так что акт рассеяния не изменяет энергию и лишь незначительно отклоняет электрон. В этом случае по происшествии времени  $\tau(E)$  электроны все еще характеризуются импульсом преимущественно в направлении оси  $x$  и их энергия приблизительно та же, что и при впрыскивании. Позже, через время релаксации импульса  $\tau_m(E)$  импульсы электронов рандомизируются, релаксируют и суммарный импульс уже не направлен вдоль оси  $x$ , однако, средняя энергия все еще близка к энергии при инжектировании, если доминантный механизм рассеяния все еще остается упругим.

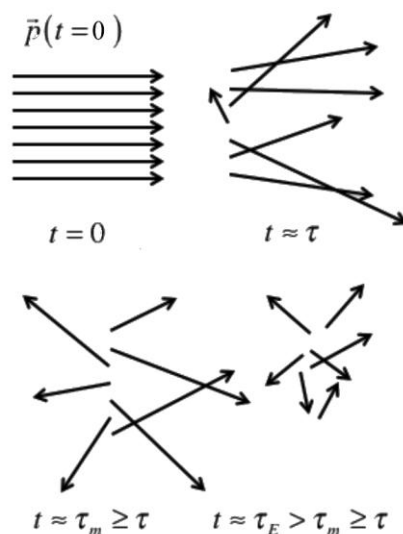


Рис. 2. Качественная визуализация различных характеристических времен рассеяния. В момент времени  $t = 0$  электроны с одинаковыми импульсами впрыскиваются в проводник в одном направлении. Электроны испытывают в среднем одно столкновение через время  $t \approx \tau(E)$ . Суммарный импульс первоначального ансамбля электронов зануляется через время  $t \approx \tau_m(E) \geq \tau(E)$ , а энергия принимает свое равновесное значение через время  $t \approx \tau_E(E) > \tau_m(E) \geq \tau(E)$ . Длина векторов на рисунках символически связана с энергией [1, 3].



Наконец, через несколько более длительное время, время релаксации энергии  $\tau_E(E)$ , впрыснутые электроны потеряют свою избыточную энергию и перейдут в равновесное состояние. Соотношение между этими характеристическими временами приблизительно такое:

$$\tau_E(E) \gg \tau_m(E) \approx \tau(E), \quad (2)$$

где отражено, что может потребоваться несколько актов неупругого рассеяния чтобы сбросить избыточную энергию. Если рассеяние изотропное, т. е. электроны равновероятно рассеиваются в любых направлениях, то  $\tau(E) = \tau_m(E)$ .

Можно определить также характеристические длины рассеяния, такие как средняя длина свободного пробега (mean-free-path/mfp)

$$\Lambda(E) = v(E) \tau(E), \quad (3)$$

иначе среднее расстояние между актами рассеяния, или mfp релаксации импульса и mfp релаксации энергии, если это необходимо.

### *Скорости рассеяния и энергия*

Фундаментальным понятием в теории рассеяния является матрица рассеяния  $S(\vec{p} \rightarrow \vec{p}')$ , переводящая систему частиц из первоначального состояния  $\vec{p}$  в некоторое конечное состояние  $\vec{p}'$ .

Скорость рассеяния, иначе вероятность рассеяния за единицу времени, есть просто единица, деленная на среднее время между соседними столкновениями, и получается путем суммирования по всем возможным конечным состояниям, а именно:

$$\frac{1}{\tau(\vec{p})} = \sum_{\vec{p}'} S(\vec{p} \rightarrow \vec{p}'). \quad (4)$$

В предположении, что электроны впрыскиваются в проводник с начальным импульсом, направленным вдоль оси  $x$ , для скорости импульсной релаксации, по аналогии с (4), имеем [1]:

$$\frac{1}{\tau_m(\vec{p})} = \sum_{\vec{p}'} S(\vec{p} \rightarrow \vec{p}') \frac{\Delta p_x}{p_x}. \quad (5)$$

Анизотропное рассеяние стремится отклонить электроны на небольшие углы, в результате чего время импульсной релаксации растет, а

соответствующая скорость падает. Аналогичным образом записывается скорость релаксации энергии через время соответствующей релаксации.

Время импульсной релаксации можно рассчитать из скорости перехода. Техника вычислений скорости перехода подробно изложена в [1]. Здесь мы обрисуем лишь ее основные моменты. Рассмотрим акт рассеяния подробнее (рис. 3).

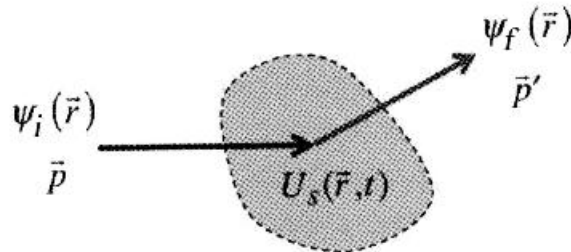


Рис. 3. Переход электрона из начального состояния в конечное в результате акта рассеяния на потенциале  $U_s(\vec{r}, t)$ .

Пусть электрон с импульсом  $\vec{p} = \hbar k$  и описываемый волновой функцией  $\psi_i(\vec{r})$  оказался в области действия потенциала рассеяния  $U_s(\vec{r}, t)$ , который может быть как статическим (рассеяние на заряженных примесях), так и динамическим (рассеяние на фононах). В результате акта рассеяния первоначальный импульс электрона становится иным  $\vec{p}'$ , изменяется и его волновая функция на  $\psi_f(\vec{r})$ . Нужно рассчитать вероятность (за единицу времени)  $S(\vec{p} \rightarrow \vec{p}')$  перехода электрона из первоначального состояния  $\vec{p}$  в конечное  $\vec{p}'$ .

В первом порядке теории возмущений

$$S(\vec{p} \rightarrow \vec{p}') = \frac{2\pi}{\hbar} |H_{p',p}|^2 \delta(E' - E - \Delta E), \quad (6)$$

где матричный элемент перехода

$$|H_{p',p}|^2 = \int_{-\infty}^{+\infty} \psi_f^*(\vec{r}) U_s(\vec{r}) \psi_i(\vec{r}) d\vec{r}.$$

Выражение (6) известно в квантовой теории как золотое правило Ферми – Дирака [5, 6]. Фигурирующая в (6)  $\delta$ -функция ответственна за сохранение энергии. В случае статического потенциала рассеяния (например, на заряженных примесях) энергия не релаксирует ( $\Delta E = 0$ ). Для периодического

потенциала с характерной частотой  $\omega$  (колебания решетки)  $\Delta E = \pm \hbar\omega$ , что соответствует излучению или поглощению фонона. В любом случае сначала нужно задаться потенциалом рассеяния, затем воспользоваться (6) и далее вычислить характерное время (4) или (5).

Согласно (4), скорость рассеяния зависит от числа конечных состояний при энергии  $E(\vec{p}')$ , доступных для электрона. Специфические механизмы рассеяния могут отбирать специальные конечные состояния, в чем мы убедимся ниже при обсуждении рассеяния на заряженных примесях, а в простейшем случае скорость рассеяния должна быть пропорциональной плотности конечных состояний. Для изотропного упругого рассеяния электронов в зоне проводимости

$$\frac{1}{\tau(E)} = \frac{1}{\tau_m(E)} \propto D(E - E_C), \quad (7)$$

а для изотропного неупругого рассеяния, когда электрон поглощает или излучает энергию  $\hbar\omega$ , например, при взаимодействии с фононом,

$$\frac{1}{\tau(E)} = \frac{1}{\tau_m(E - E_C)} \propto D(E \pm \hbar\omega - E_C). \quad (8)$$

Для простых зон параболического профиля известны аналитические выражения для времен рассеяния [1], а для более сложных зон не обойтись без численного суммирования по конечным состояниям.

В случае полупроводников времена рассеяния часто записывают в форме степенного закона

$$\tau_m(E) = \tau_{m0} \left( \frac{E - E_C}{kT} \right)^s, \quad (9)$$

где  $s$  – характерный показатель степени, соответствующий тому или иному механизму рассеяния. Например, рассеяние электронов на акустических фононах допустимо рассматривать как почти упругое и изотропное при комнатной температуре. Скорость рассеяния должна быть пропорциональной плотности состояний, которая для 3D электронов при параболической зонной структуре полупроводника пропорциональна  $(E - E_C)^{1/2}$ , так что для времени рассеяния можно ожидать пропорциональность величине  $(E - E_C)^{-1/2}$ . Характерный показатель степени для рассеяния на акустических фононах  $s = -1/2$ . Для 2D электронов плотность состояний не зависит от энергии, так что

характерный показатель степени  $s = 0$ . Для 1D электронов плотность состояний пропорциональна  $(E - E_c)^{-1/2}$  и характерный показатель степени в степенном законе (9)  $s = +1/2$ . Не всегда можно описать время рассеяния степенным законом, но если это возможно, тогда расчеты сильно упрощаются.

Когда при рассеянии учитываются электростатические взаимодействия, например, при рассеянии на заряженных примесях или при рассеянии на фононах в полярных средах, зависимость времени рассеяния от энергии иная. Как показано на рис. 4, расположенные случайным образом заряды вызывают флуктуации профиля дна зоны проводимости  $E_c(x)$ , приводящие к рассеянию электронов. Однако, высокоэнергетические электроны чувствуют флуктуирующий потенциал в меньшей степени чем электроны с более низкой энергией, так что в случае рассеяния на заряженных примесях или на полярных фононах можно ожидать, что скорость рассеяния  $1/\tau(E)$  уменьшится (время рассеяния  $\tau(E)$  увеличится) с ростом энергии электронов. Времени рассеяния для 3D электронов в степенном законе соответствует значение  $s = +3/2$  [1]. При рассеянии на неполярных фононах время рассеяния уменьшается с энергией электронов.

В отношении рассеяния на заряженных примесях важно отметить его анизотропию. Значительная доля электронов характеризуется большой энергией и почти не меняют свои траектории в поле заряженных примесей. В результате время релаксации импульса при рассеянии на заряженных примесях существенно больше среднего времени рассеяния:  $\tau_m(E) \gg \tau(E)$ .

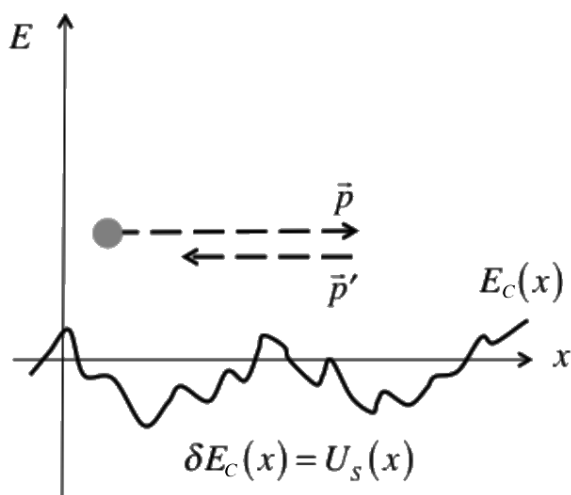


Рис. 4. Флуктуирующий потенциал, создаваемый случайно расположенными зарядами, иллюстрирует рассеяние на заряженных примесях. Электроны с высокой энергией чувствуют возмущающий потенциал в меньшей степени по сравнению с электронами с меньшей энергией и поэтому рассеиваются меньше.

Средняя длина свободного пробега также может быть записана в виде степенного закона. Из уравнения (3), помня что для параболических зон  $v(E) \propto (E - E_C)^{1/2}$  имеем

$$\Lambda(E) = v(E)\tau(E) \propto (E - E_C)^{1/2} \left( \frac{E - E_C}{kT} \right)^s = \Lambda_0 \left( \frac{E - E_C}{kT} \right)^r, \quad (10)$$

где характеристический показатель степени для средней длины свободного пробега  $r = s + 1/2$ . Для рассеяния на акустических фононах в 3D проводнике  $s = -1/2$ , так что  $r = 0$ : средняя длина свободного пробега не зависит от энергии. При рассеянии на акустических фононах в 2D проводнике  $s = 0$ , так что  $r = 1/2$ : средняя длина свободного пробега увеличивается с ростом энергии электронов.

### *Коэффициент прохождения*

Различие между прохождением  $T_{SD}(E)$  от истока к стоку и  $T_{DS}(E)$  от стока к истоку иллюстрируется на рис. 5. Величина  $T_{SD}(E)$  есть отношение стационарного потока электронов, выходящего из стока, к потоку, инжектированному истоком, а величина  $T_{DS}(E)$  есть отношение потока электронов, выходящего из истока, к потоку, инжектированному стоком. При нулевом или небольшом напряжении на стоке можно ожидать, что оба напряжения одинаковы:  $T_{SD}(E) \approx T_{DS}(E) = T(E)$ . Этот случай показан в верхней части рис. 5. В нижней части рисунка показана ситуация, когда напряжение на стоке велико. В этом случае совсем не очевидно, будут или не будут одинаковы  $T_{SD}(E)$  и  $T_{DS}(E)$ , но можно показать, что в случае упругого рассеяния оба прохождения одинаковы. Однако, при неупругом рассеянии они совершенно различны:  $T_{DS}(E) \ll T_{SD}(E)$ .

При моделировании тока в MOSFET тот факт, что  $T_{DS}(E) \ll T_{SD}(E)$  при большом напряжении на стоке, не играет большого значения, поскольку при большом  $V_{DS}$  поток электронов, инжектированных истоком, и так слишком незначителен. Таким образом, обычно предполагается, что достаточно вычислять лишь один коэффициент прохождения,  $T(E)$ , который определяет прохождение в любом из направлений. Ранее было показано, что коэффициент прохождения  $T(E)$  связан со средней длиной пробега между столкновениями назад  $\lambda(E)$  согласно (21/гл. 1):

$$T(E) = \lambda(E) / [\lambda(E) + L]. \quad (11)$$

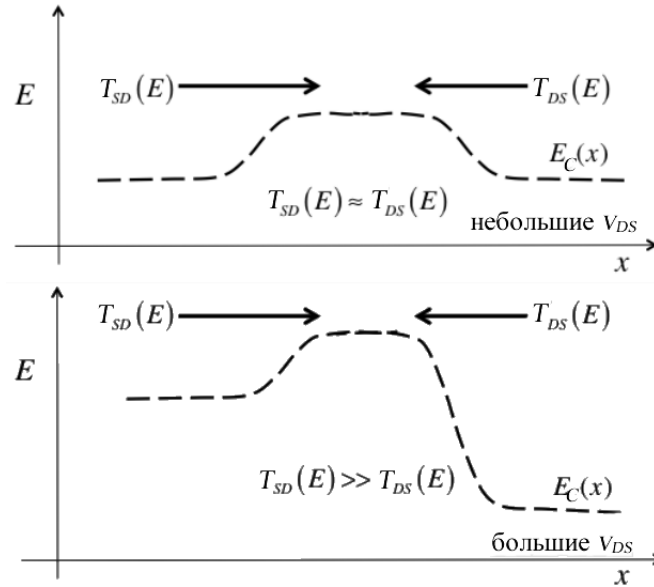


Рис. 5. Иллюстрация двух коэффициентов прохождения  $T_{SD}(E)$  и  $T_{DS}(E)$ . При впрыскивании потока электронов истоком прохождение  $T_{SD}(E)$  определяет долю электронов, покидающих сток, а при впрыскивании электронов стоком прохождение  $T_{DS}(E)$  определяет долю электронов, покидающих исток. Верхний рисунок соответствует нулевому или небольшому напряжению  $V_{DS}$ , а нижний рисунок соответствует большому напряжению на стоке.

Обращает на себя внимание, что в выражении для коэффициента прохождения используется средняя длина пробега  $\lambda(E)$ , а не  $L(E)$  по (3). Это уравнение (11) не сложно вывести [2, 3], как и не сложно убедиться в его справедливости.

Уравнение (11) пригодно для описания прохождения во всем интервале от баллистического предела до диффузионного предела. Когда длина проводника меньше средней длины свободного пробега, тогда

$$T(E) = \frac{\lambda(E)}{\lambda(E) + L} \rightarrow 1, \quad (L \ll \lambda(E)) \quad (12)$$

и перенос электронов в канале проводимости баллистический. Когда длина проводника велика по сравнению с длиной свободного пробега, тогда

$$T(E) = \frac{\lambda}{\lambda + L} \rightarrow \frac{\lambda}{L}. \quad (L \gg \lambda). \quad (13)$$

Уравнение (11) описывает прохождение электронов по проводнику в отсутствие электрического поля. Что происходит при приложении электрического поля к каналу проводимости иллюстрируется на рис. 6.

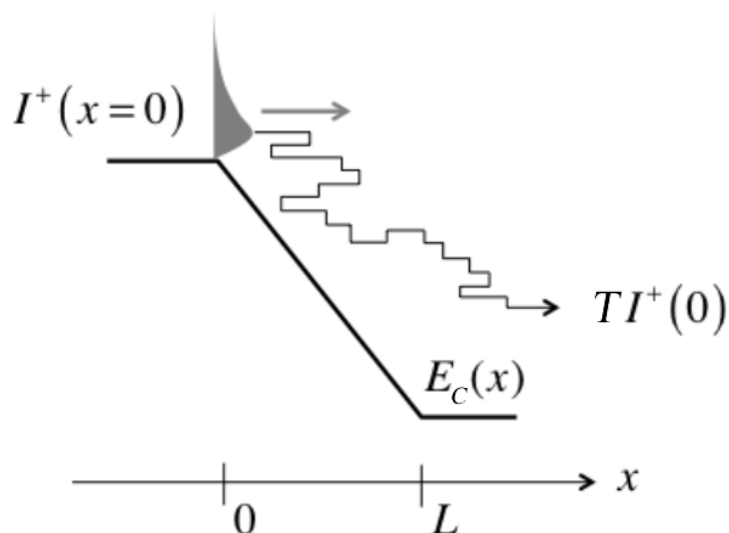


Рис. 6. Иллюстрация траектории электрона в коротком канале проводимости при высоком приложенном электрическом поле. Электроны впрыскиваются истоком из равновесного состояния и большая их часть достигает стока, даже если и происходят акты рассеяния [7].

На рис. 6 показан короткий канал проводимости, к которому приложено сильное электрическое поле. Равновесный поток электронов впрыскивается истоком. Электроны быстро наращивают кинетическую энергию и скорость их рассеяния увеличивается. Моделирование электронного транспорта через короткие каналы с большим перепадом напряжения, как в рассматриваемом случае, когда наблюдаются такие эффекты как резкие всплески скорости, одна из интригующих задач современной полуклассической теории транспорта [1]. Расчет средней скорости в зависимости от положения электрона в проводнике весьма непростая задача, однако, детальное моделирование ее с использованием понятия прохождения приводит к весьма понятным и простым результатам [7]. Обнаружено, что если инжектированные электроны проникают даже на короткое расстояние вглубь области с сильным электрическим полем, тогда даже если они и подвергаются актам рассеяния, то они все равно обречены достичь конца канала проводимости [7]. Даже если имеет место сильное рассеяние, коэффициент прохождения может быть почти равен единице, поскольку сильное электрическое поле подхватывает электроны и выносит их из канала проводимости в стоковый терминал.

В высококачественном MOSFET при большом напряжении на стоке электрическое поле вблизи вершины барьера у входа в канал проводимости слабое, а вблизи стока сильное. Для понимания того, что происходит в такой ситуации, рассмотрим модель на рис. 7.

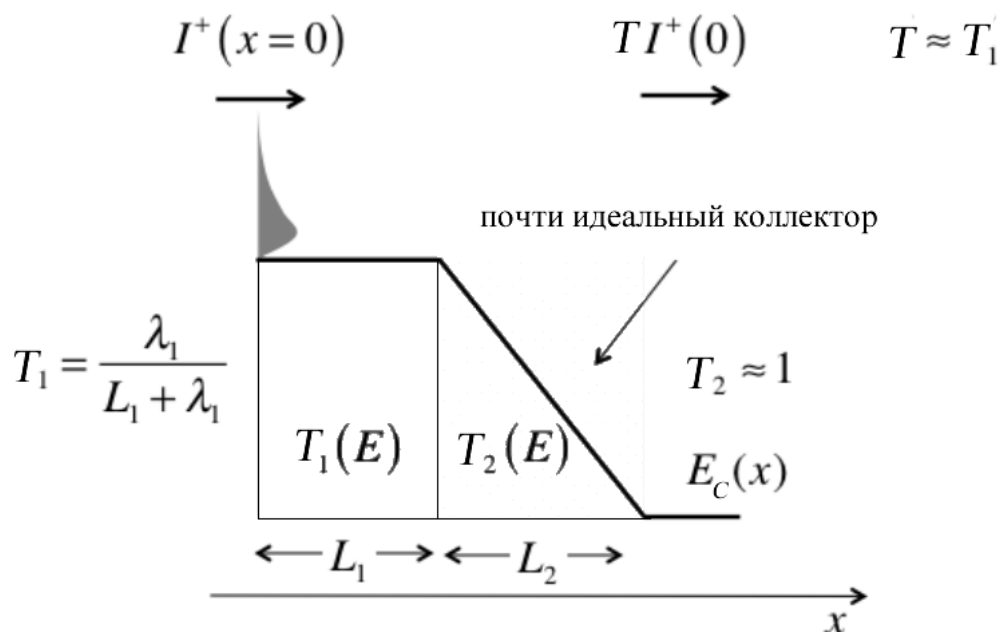


Рис. 7. Модельный профиль канала проводимости, иллюстрирующий прохождение электрона сначала через область с низким электрическим полем, а затем – с высоким электрическим полем.

В этой модели энергетический профиль канала проводимости состоит из двух участков: короткого участка длиной  $L_1$  с постоянным потенциалом и средней длиной свободного пробега  $\lambda_1$  и длинного участка длиной  $L_2$  с сильным электрическим полем. Прохождение через первый участок есть  $T_1 = \lambda_1 / (L_1 + \lambda_1)$ , а через второй участок –  $T_2 \approx 1$ . Суммарное прохождение, очевидно, будет  $T \approx \lambda_1 / (L_1 + \lambda_1)$ . Важным выводом из такого рассмотрения является следующий: прохождение через сложную структуру, состоящую из низкополевого участка и далее через участок с высоким электрическим полем, контролируется длиной низкополевого участка. На практике, когда электрическое поле меняется плавно вдоль канала проводимости, бывает трудно надежно определить длину низкополевого участка [8, 9], но это проведенное выше простое рассмотрение физически наглядно поясняет результаты математического моделирования [7].



Подведем итоги обсуждения понятия коэффициента прохождения:

1. Прохождение связано со средней длиной свободного пробега назад  $T = \lambda / (L + \lambda)$ .
2. При баллистическом транспорте  $T \rightarrow 1$ , если  $L \ll \lambda$ .
3. При диффузионном транспорте  $T \rightarrow \lambda / L \ll 1$ , если  $L \gg \lambda$ .
4. Область в канале проводимости с высоким электрическим полем является хорошим коллектором электронов,  $T \approx 1$ .
5. В канале проводимости, в котором электрическое поле меняется от слабого значения до сильного значения, как это имеет место в MOSFET при большом напряжении на стоке, прохождение контролируется областью с низким электрическим полем.

*Средняя длина свободного пробега назад*

Мы пользуемся двумя разными понятиями средней длины свободного пробега. Длина  $\Lambda$  (3) есть среднее расстояние между двумя актами рассеяния. Именно этот смысл большинство физиков вкладывает в понятие «средней длины свободного пробега». Обратная величина  $1/\Lambda$  есть вероятность рассеяния на единицу длины. Однако, в нашем контексте величина  $\lambda$ , как средняя длина свободного пробега назад, более существенна. Ее обратная величина  $1/\lambda$  есть вероятность на единицу длины того, что прямой (в положительном направлении) поток рассеется назад (в обратном направлении). Прохождение (11) определяется через понятие величины  $\lambda$ . Как связаны между собой понятия длин  $\lambda$  и  $\Lambda$ ?

Рисунок 8 иллюстрирует рассеяние в 1D канале проводимости, например, в NW MOSFET.

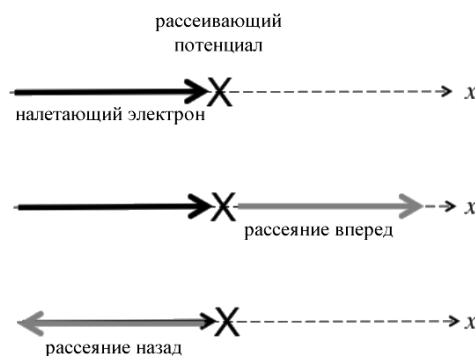


Рис. 8. Рассеяние вперед и рассеяние назад в канале 1D [2, 3].

Пусть рассеяние будет изотропным и среднее время между двумя актами рассеяния есть  $\tau$ . Если электрон, движущийся в прямом направлении, совершит акт рассеяния через время  $\tau$ , он с одинаковой вероятностью может рассеяться как вперед, так и назад. Только рассеяние назад, которое в среднем происходит через время  $2\tau$ , существенно для поддержания тока. В соответствии с этим очевидным утверждением, средняя длина свободного пробега назад в 1D канале есть

$$\lambda(E) = 2v(E)\tau_m = 2\Lambda, \quad (14)$$

где используется время релаксации импульса, поскольку предполагается изотропное рассеяние, для которого  $\tau_m = \tau$ .

В 2D и 3D каналах определение средней длины свободного пробега назад предполагает усреднение по углам [10], как это показано на рис. 9 для канала 2D.

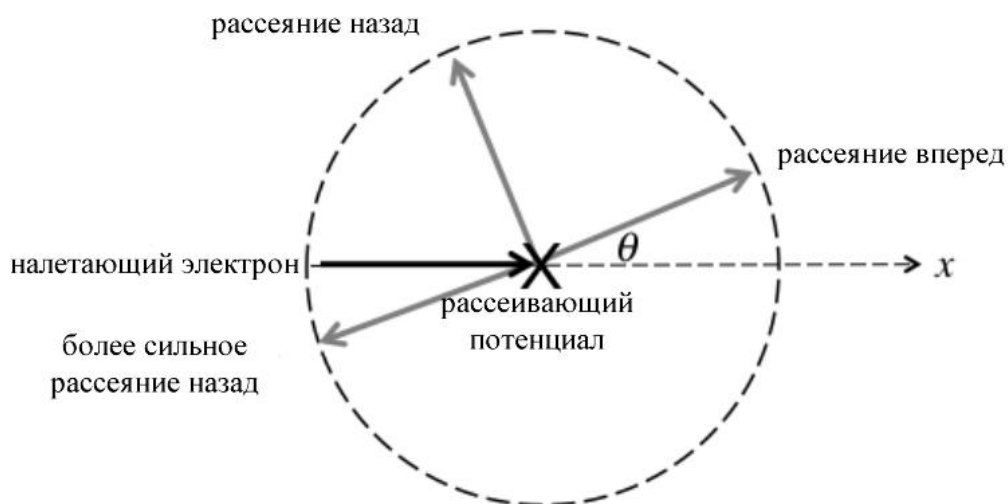


Рис. 9. Рассеяние вперед и рассеяние назад в канале 2D.

Для 2D канала проводимости окончательно имеем

$$\lambda(E) = \frac{\pi}{2}v(E)\tau_m = \frac{\pi}{2}\Lambda, \quad (15)$$

а для 3D канала  $\lambda(E) = 4\Lambda/3$  [2, 3, 10]. Для корректного расчета прохождения важно учитывать различие между двумя средними длинами  $\lambda$  и  $\Lambda$ .

Простое выражение для коэффициента прохождения (11) связывает между собой среднюю длину свободного пробега назад  $\lambda$  и длину низкополевой части канала проводимости, приходящейся на небольшую область вершины барьера.

Для вычисления прохождения нужно знать величину  $\lambda$ . Последнюю можно вычислить по (15) или определить экспериментально. Ситуация в классическом понимании выглядит как задача диффузии – частицы инжектируются левым терминалом, диффундируют через канал и покидают его через противоположный терминал. Выходящий поток электронов определяется коэффициентом прохождения, в классической постановке задачи все дело в коэффициенте диффузии. Тщательный анализ этой ситуации [2, 3] показывает, что имеет место следующее простое соотношение между коэффициентом диффузии и средней длиной свободного пробега назад  $\lambda$ , а именно:

$$D = \frac{v_T \langle \lambda \rangle}{2}, \quad (16)$$

где  $\langle \lambda \rangle$  – средняя длина свободного пробега, усредненная по энергии. Именно это удивительно простое соотношение открывает возможность определить среднюю длину свободного пробега экспериментально. Обратим только внимание на то, что выражение (16) пригодно лишь для невырожденной статистики электронов. Более общий случай статистики Ферми – Дирака применительно к этой задаче обсуждается в [2, 3].

Легче измерить подвижность, чем коэффициент диффузии, так что в литературе проще найти сведения о подвижности. Соотношение между ними дается уравнением Эйнштейна

$$\frac{D}{\mu} = \frac{kT}{q}, \quad (17)$$

которое, правда, справедливо вблизи равновесия, однако, и электроны в низкополевой части канала в области вершины барьера, которая, собственно, и контролирует прохождение, тоже находятся вблизи равновесия.

Таким образом, налицо довольно простой путь оценить среднюю длину свободного пробега назад через измерение подвижности электронов. Сначала определяется коэффициент диффузии из подвижности по (17), а затем – средняя длина свободного пробега по (16). Учет вырождения, что важно в области выше порога, обсуждается в [2, 3].

В качестве примера вычислим среднюю длину свободного пробега и коэффициент прохождения 22 нм MOSFET. Рассмотрим *n*-Si MOSFET с каналом длины  $L = 22 \text{ нм}$  при  $T = 300 \text{ K}$  в квазиклассическом режиме. Пусть

ориентация подложки (100) и заселена лишь нижняя подзона. Подвижность  $\mu = 250 \text{ см}^2/\text{В} \cdot \text{с}$ . Найдем  $\langle \lambda \rangle$  и  $T$ .

Вычисляем коэффициент диффузии из подвижности по (17):

$$D = \mu \frac{kT}{q} = 6.5 \text{ см}^2/\text{с}.$$

Возьмем  $m^* = 0.19m_0$ , так что  $v_T = 1.23 \times 10^7 \text{ см/с}$ , и определяем  $\langle \lambda \rangle$  по (16):

$$\langle \lambda \rangle = \frac{2D}{v_T} = \frac{2 \times 6.5}{1.2 \times 10^7} = 10.5 \text{ нм},$$

а далее определяем прохождение по (11):

$$T \approx \frac{\langle \lambda \rangle}{\langle \lambda \rangle + L} = \frac{10.5}{10.5 + 22} = 0.32.$$

Этот результат является приблизительным, поскольку величина  $\langle \lambda \rangle$  в этом выражении не усреднялась по энергии, как того требует формула (11). Тем не менее, полученный результат говорит о том, что такой MOSFET в линейной области работает на одну треть в баллистическом режиме. При высоком напряжении на стоке электроны обладают большей энергией, что ведет и к более интенсивному рассеянию. Однако, дальнейшее рассмотрение, как это не удивительно, покажет нам, что MOSFET при высоком напряжении на стоке работает ближе к баллистическому режиму, чем при низком напряжении.

Итак, существенные результаты таковы:

$$\begin{aligned} T(E) &= \frac{\lambda(E)}{\lambda(E) + L}, \\ \lambda(E) &= \frac{\pi}{2} v(E) \tau_m, \\ \langle \lambda \rangle &= \frac{2D}{v_T}, \end{aligned} \tag{18}$$

где под  $L$  нужно понимать низкополевую часть канала проводимости. Множитель  $\pi/2$  возникает из-за усреднения по углам в 2D канале. Последнее уравнение показывает простой путь оценить среднюю длину свободного пробега через коэффициент диффузии в предположении невырожденной статистики. Далее мы рассмотрим модель прохождения MOSFET и выясним, каким образом рассеяние назад влияет на работоспособность транзисторов.

### 7.3. Модель прохождения MOSFET

Перейдем к построению модели наноразмерных MOSFET с учетом рассеяния. Именно учет рассеяния делает моделирование транспортных явлений трудной задачей, а в случае MOSFET – особенно сложной [8, 11]. Тем не менее, мы убедимся в том, что основные принципы моделирования MOSFET с учетом рассеяния легко понять и использовать далее для интерпретации экспериментальных данных и результатов вычислений при моделировании.

Итак, для тока мы имеем общее выражение (2/гл. 1) с учетом всех возможных режимов протекания тока

$$I_{DS} = \frac{2q}{h} \int T(E)M(E)(f_S(E) - f_D(E))dE. \quad [A] \quad (19)$$

Если на сток подано большое напряжение, тогда  $f_S(E) \gg f_D(E)$  и ток насыщения

$$I_{DSAT} = \frac{2q}{h} \int T(E)M(E)f_S(E)dE. \quad (20)$$

В квазиравновесном режиме ток небольшой,  $f_S \approx f_D$ , и для тока и проводимости на линейном участке имеем (10/гл. 1):

$$I_{DLIN} = G_{ch}V_{DS}, \quad [A] \quad (21)$$

$$G_{ch} = \frac{2q^2}{h} \int T(E)M(E) \left( -\frac{\partial f_0}{\partial E} \right) dE. \quad [Cm]$$

По этим уравнениям можно вычислить линейный ток, величину  $I_{ON}$  и ток во всей области изменения  $V_{DS}$  от  $V_{DS}=0$  до  $V_{DS}=V_{DD}$ . С целью упрощения дальнейших построений будем предполагать, что средняя длина свободного пробега не зависит от энергии:

$$T(E) = \frac{\lambda(E)}{\lambda(E) + L} \rightarrow T = \frac{\lambda_0}{\lambda_0 + L}, \quad (22)$$

так что выражения для прохождения будем считать соответствующим образом усредненными по энергии.

## Резюме баллистической модели MOSFET

Выпишем окончательные результаты баллистической модели (69/гл. 1) следующим образом. Ток стока

$$I_{DS} = W |Q(V_{GS}, V_{DS})| F_{SAT} v_{inj}^{ball}. \quad (23)$$

Функция насыщения тока стока для обеих статистик дается выражениями:

$$F_{SAT} = \left[ \frac{1 - \mathfrak{F}_{1/2}(\eta_{FD}) / \mathfrak{F}_{1/2}(\eta_{FS})}{1 + \mathfrak{F}_0(\eta_{FD}) / \mathfrak{F}_0(\eta_{FS})} \right], \quad (\text{Ферми - Дирак / FD})$$

$$F_{SAT} = \left[ \frac{1 - e^{-qV_{DS}/kT}}{1 + e^{-qV_{DS}/kT}} \right], \quad (\text{Максвелл - Больцман / MB}) \quad (24)$$

$$\eta_{FS} = (E_{FS} - E_C(0))/kT, \quad \eta_{FD} = \eta_{FS} - qV_{DS}/kT.$$

Баллистическая скорость впрыскивания

$$v_{inj}^{ball} = v_T \frac{\mathfrak{F}_{1/2}(\eta_{FS})}{\mathfrak{F}_0(\eta_{FS})}, \quad (FD)$$

$$v_{inj}^{ball} = v_T, \quad (MB) \quad (25)$$

$$v_T = \sqrt{\frac{2kT}{\pi m^*}}.$$

Баллистический ток в линейной области

$$I_{DLIN} = W |Q(V_{GS}, V_{DS})| \left( \frac{v_{inj}^{ball}}{2kT/q} \right) \frac{\mathfrak{F}_{-1/2}(\eta_{FS})}{\mathfrak{F}_{1/2}(\eta_{FS})} V_{DS}, \quad (FD) \quad (26)$$

$$I_{DLIN} = W |Q(V_{GS}, V_{DS})| \left( \frac{v_T}{2kT/q} \right) V_{DS}, \quad (MB)$$

а для тока насыщения имеем:

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| v_{inj}^{ball}, \quad (FD) \quad (27)$$

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| v_T. \quad (MB)$$

Наконец, электронный заряд на вершине барьера

$$Q(V_{GS}, V_{DS}) = -q \frac{N_{2D}}{2} [\mathfrak{F}_0(\eta_{FS}) + \mathfrak{F}_0(\eta_{FD})], \quad (FD) \quad (28)$$

а для статистики Максвелла – Больцмана интегралы Ферми – Дирака просто сводятся к экспонентам.

Для учета рассеяния, казалось бы, нужно эти уравнения умножить только на прохождение  $T$ . Мы далее убедимся, что это действительно так для линейного тока, но не для тока насыщения и не для электронного заряда.

### *Линейная область*

Для вычисления линейного тока с учетом рассеяния начинаем с уравнения (21) для проводимости. Для числа мод пользуемся уравнением (60/гл. 1). При записи функции Ферми учитываем, что  $E_F \approx E_{FS} \approx E_{FD}$ . После вычисления интеграла получим уравнение, похожее на (47/гл. 1), в котором нужно лишь заменить прохождение в диффузионном пределе  $\lambda_0/L$  на прохождение  $T$ , как в исходном уравнении (21) для описания транспорта во всех режимах от баллистического до диффузионного, а именно:

$$I_{DLIN} = T \left[ W \frac{2q^2}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{2\pi \hbar} \right) \mathfrak{F}_{-1/2}(\eta_F) \right] V_{DS}, \quad (29)$$

которое есть ничто иное как уравнение баллистического тока в линейной области (61/гл. 1), умноженное на прохождение  $T$ .

Уравнение (29) справедливо для линейной области MOSFET, работающего во всех режимах от баллистического до диффузионного, однако, выглядит оно непривычно по сравнению с традиционным выражением (5/гл. 2)

$$I_{DLIN} = \frac{W}{L} \mu C_{ox} (V_{GS} - V_T) V_{DS}. \quad (30)$$

В полном соответствии этих уравнений друг другу мы убедимся позже.

### *Область насыщения*

Для вычисления тока в области насыщения начинаем с уравнения (20) и вычисляем интеграл таким же образом, как и в линейном случае выше. Окончательно получаем уравнение

$$I_{DSAT} = TW \frac{2q}{h} \left( \frac{g_v \sqrt{2m^* kT}}{\pi \hbar} \right) kT \frac{\sqrt{\pi}}{2} \mathfrak{F}_{1/2}(\eta_F), \quad (31)$$

которое является ничем иным как уравнением баллистического тока в области насыщения (64/гл. 1), умноженным на прохождение  $T$ . Уравнение (31) справедливо для области насыщения MOSFET, работающего во всех режимах от баллистического до диффузионного, однако, выглядит оно непривычно по сравнению с традиционным выражением (7/гл. 2) через скорость насыщения

$$I_{DSAT} = W C_{ox} (V_{GS} - V_T) v_{sat}. \quad (32)$$

Связь между этими уравнениями, полученными в разных моделях, обсудим позже.

### *От линейного тока до насыщения*

Для вычисления тока при любом значении напряжения на стоке с учетом рассеяния начинаем с уравнения (19) с вычислением интегралов таким же образом, как и прежде. Окончательно получаем

$$I_{DS} = TW \frac{q}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{\pi \hbar} \right) kT [\mathfrak{F}_{1/2}(\eta_{FS}) - \mathfrak{F}_{1/2}(\eta_{FD})], \quad (33)$$

$$\eta_{FS} = [E_{FS} - E_C(0)]/kT, \quad \eta_{FD} = \eta_{FS} - qV_{DS}/kT.$$

Уравнение (33) отличается от аналогичного результата в баллистическом пределе (66/гл. 1) лишь умножением на прохождение  $T$ . Оно переходит в уравнение (29) для малых  $V_{DS}$  и в (31) для больших  $V_{DS}$ , в чем не сложно убедиться. Мы видим, что ток стока с учетом рассеяния отличается от баллистического тока лишь умножением на коэффициент прохождения  $T$ . Если же ток выразить через электронный заряд, то ситуация окажется не такой же простой, в чем мы сейчас убедимся.

Мы до сих пор рассматривали 2D электроны. Не сложно вывести аналогичные формулы для 1D электронов в NW MOSFET.

### *Выражение для тока через электронный заряд*

Уравнение (33) в модели Ландауэра – Датта – Лундстрёма справедливо для произвольного напряжения на стоке, но в нем не видна зависимость от инверсионного электронного заряда  $Q$ . Однако, всегда предпочтительно выразить ток через заряд  $Q$ , связанный непосредственно с электростатикой структуры MOS. Для вычисления заряда  $Q$  нужно ввести в рассмотрение положительные потоки электронов, инжектированных истоком, которые



заселяют состояния  $+v_x$  на вершине барьера, и отрицательные потоки электронов, инжектированные стоком, которые заселяют там же состояния  $-v_x$ . В баллистических MOSFET результат нам известен в виде формул для тока и заряда (69/гл. 1). Этот результат изменится из-за необходимости учета рассеяния назад, поскольку нам нужно учесть все возможности заселения состояний на вершине барьера. Как показано на рис. 10, как и прежде, у нас есть баллистический поток, инжектированный истоком, заселяющий состояния на вершине барьера, но добавляется также и возвратный поток рассеяния назад в направлении истока, а величина баллистического потока, инжектированного стоком в направлении вершины барьера уменьшается на коэффициент прохождения. Таким образом, уравнение для заряда (29) нужно заменить на следующее уравнение

$$Q(V_{GS}, V_{DS}) = -q \frac{N_{2D}}{2} [\mathfrak{I}_0(\eta_{FS}) + (1-T)\mathfrak{I}_0(\eta_{FS}) + T\mathfrak{I}_0(\eta_{FD})]. \quad (34)$$

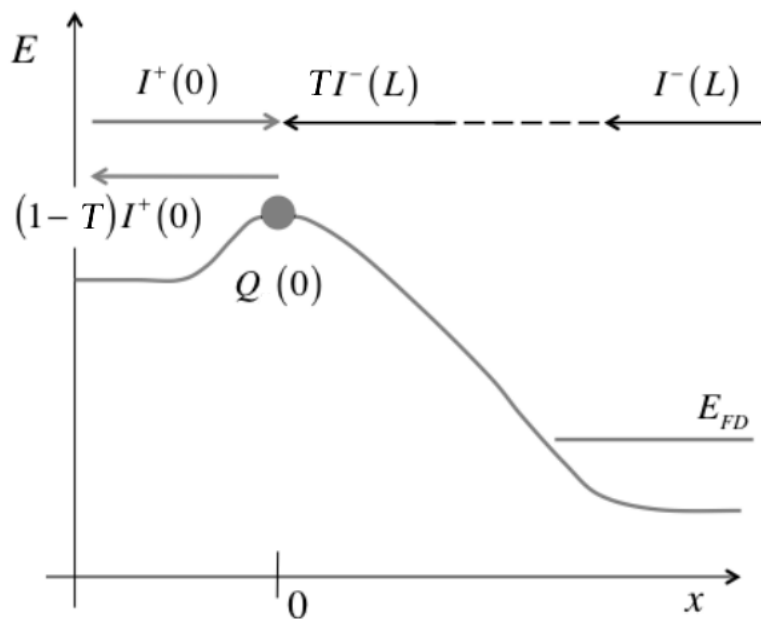


Рис. 10. Заселенность состояний на вершине барьера определяется тремя потоками электронов: со стороны истока, возвратным потоком рассеяния назад в направлении истока и потоком со стороны стока, уменьшенным на коэффициент прохождения.

Первое слагаемое в выражении для заряда на вершине барьера (34) есть баллистический поток, инжектированный истоком. Его величина зависит от положения электрохимического потенциала в истоке. Второе слагаемое есть вклад в заряд на вершине барьера от потока рассеяния назад в сторону истока. Поскольку этот возвратный поток берет свое начало на истоке, он тоже зависит

от положения электрохимического потенциала в истоке. Третье слагаемое обязано баллистическому потоку, берущему свое начало на стоке и потому зависящего от электрохимического потенциала на стоке и уменьшенному на коэффициент прохождения.

Уравнения (33) и (34) позволяют теперь выразить ток через электронный заряд. Для этого (33) умножим и разделим на заряд  $Q$ , так что

$$I_{DS} = TW \frac{|Q|}{|Q|} \frac{q}{h} \left( \frac{g_v \sqrt{2\pi m^* kT}}{\pi \hbar} \right) kT [\mathfrak{I}_{1/2}(\eta_{FS}) - \mathfrak{I}_{1/2}(\eta_{FD})], \quad (35)$$

а затем заряд в знаменателе заменим уравнением (34). После несложных алгебраических преобразований получаем следующие окончательные результаты с учетом рассеяния для тока, выраженного через электронный заряд, во всем диапазоне изменения напряжения на стоке:

$$I_{DS} = W |Q(V_{GS}, V_{DS})| v_{inj} \left[ \frac{1 - \frac{\mathfrak{I}_{1/2}(\eta_{FD})}{\mathfrak{I}_{1/2}(\eta_{FS})}}{1 + \frac{T}{2-T} \frac{\mathfrak{I}_0(\eta_{FD})}{\mathfrak{I}_0(\eta_{FS})}} \right],$$

$$Q = -q \frac{N_{2D}}{2} [\mathfrak{I}_0(\eta_{FS}) + (1-T)\mathfrak{I}_0(\eta_{FS}) + T\mathfrak{I}_0(\eta_{FD})],$$

$$v_{inj} = v_{inj}^{ball} \left( \frac{T}{2-T} \right), \quad (36)$$

$$v_{inj}^{ball} = \sqrt{\frac{2kT}{\pi m^*}} \frac{\mathfrak{I}_{1/2}(\eta_{FS})}{\mathfrak{I}_0(\eta_{FS})} = v_T \frac{\mathfrak{I}_{1/2}(\eta_{FS})}{\mathfrak{I}_0(\eta_{FS})},$$

$$\eta_{FD} = \eta_{FS} - qV_{DS}/kT.$$

Полученные уравнения с учетом рассеяния можно сравнить с уравнениями (69/гл. 1) для баллистического тока. Поскольку  $T < 1$ , скорость впрыскивания с учетом рассеяния назад  $v_{inj}$  меньше баллистической скорости впрыскивания,  $v_{inj} < v_{inj}^{ball}$ , что приводит к току меньшему по сравнению с баллистическим током.

Уравнения (36) можно упростить для небольших и больших напряжений на стоке таким же образом, как было получено выражение (70/гл. 1). Для линейного участка тока получаем

$$I_{DLIN} = W |Q(V_{GS})| T \frac{v_{inj}^{ball}}{2kT/q} \frac{\mathfrak{I}_{-1/2}(\eta_{FS})}{\mathfrak{I}_{1/2}(\eta_{FS})} V_{DS}, \quad (FD) \quad (37)$$

$$I_{DLIN} = W |Q(V_{GS})| T \frac{v_T}{2kT/q} V_{DS}, \quad (MB)$$

где обе скорости,  $v_{inj}^{ball}$  и  $v_T$ , нужно брать по (25). Аналогично получаем выражения для тока насыщения:

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| \left( \frac{T}{2-T} \right) v_{inj}^{ball}, \quad (FD)$$

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| \left( \frac{T}{2-T} \right) v_T. \quad (MB)$$
(38)

При использовании статистики Ферми – Дирака нужно знать положение уровня Ферми. Его значение находим по известному значению электронного заряда из второго уравнения в (36).

Эти результаты, полученные с учетом рассеяния, сравним с баллистическими результатами (26) и (27). Оказывается, что линейный ток с учетом рассеяния отличается от линейного баллистического тока лишь умножением на прохождение  $T$ , а ток насыщения с учетом рассеяния получается из баллистического тока насыщения умножением его на  $T/(2-T)$ . Это различие, связанное с балансом потоков (34), определяющих электронный заряд при учете рассеяния, мы обсудим ниже.

Уравнение (36) описывает цельную вольт-амперную характеристику MOSFET Ландауэра – Датта – Лундсрома (ЛДЛ) через электронный заряд на вершине барьера  $Q$  и коэффициент прохождения  $T$ . Однако, эти уравнения не так просто использовать на практике, поскольку коэффициент прохождения  $T$  зависит от напряжения на стоке  $V_{DS}$ , что мы обсудим ниже.

Выходную характеристику ЛДЛ (36) можно вычислить следующим образом. Сначала вычисляем  $Q(V_{GS}, V_{DS})$  из электростатики структуры MOS, используя, например, полуэмпирическое уравнение (46/гл. 4). Затем, задавшись значением  $Q$  во втором уравнении (36), вычисляем положение электрохимического потенциала истока  $\eta_{FS}$ . Это создает некоторые проблемы, потому что для этого нам нужно знать значение  $T$  и понять, как коэффициент прохождения  $T(V_{GS}, V_{DS})$  зависит от подачи напряжений. Далее вычисляем баллистическую скорость впрыскивания из четвертого уравнения (36), а затем скорость впрыскивания с учетом рассеяния из третьего уравнения (36). Наконец, вычисляем ток в точке характеристики  $(V_{GS}, V_{DS})$  из первого уравнения. Основная проблема вычислений по этой модели состоит в том, что до сих пор не существует достаточно хороших моделей для оценки  $T(V_{GS}, V_{DS})$ . Именно

поэтому такое широкое применение нашла полуэмпирическая модель виртуального истока.

На практике часто используется невырожденная максвелл-больцмановская форма этих уравнений. Это ведет к некоторой погрешности, особенно выше порога, однако, невырожденные уравнения намного проще, так что приходится выбирать между простотой вычислений и точностью. Отметим все же, что использование невырожденной статистики применительно к III-V FET может привести к слишком большим погрешностям.

### *Зависимость коэффициента прохождения от напряжения на стоке*

Несмотря на всю простоту предложенной выше модели MOSFET на практике возникают вопросы, связанные с тем, каким образом учесть зависимость коэффициента прохождения от напряжения на стоке. Эта зависимость прохождения от  $V_{DS}$  демонстрируется на рис. 11.

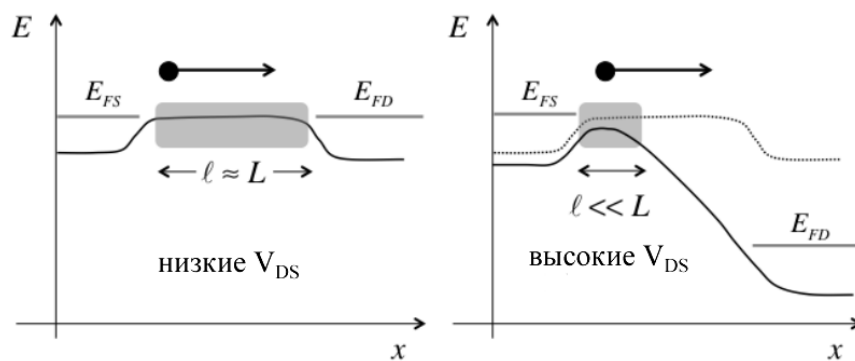


Рис. 11. Иллюстрация эффекта зависимости коэффициента прохождения от напряжения на стоке и почему этот эффект больше при бóльших  $V_{DS}$ , чем при меньших. Средняя длина свободного пробега в серой зоне предполагается  $\sim \lambda_0$  в обоих случаях.

При низком напряжении на стоке электрическое поле невелико вдоль всего канала проводимости. Как уже обсуждалось ранее, прохождение определяется длиной низкополевой части канала, так что при низком стоковом напряжении

$$T_{LIN} = \frac{\lambda_0}{\lambda_0 + L}. \quad (39)$$

При высоком напряжении на стоке в качественном транзисторе низкополевая область – небольшая, длиной  $l$  в самом начале канала проводимости. Высокополевая область канала работает почти как идеальный коллектор электронов с  $T \approx 1$ . Как уже обсуждалось ранее, прохождение канала

в этом случае определяется длиной низкополевого участка, так что при высоких напряжениях на стоке

$$T_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell}. \quad (40)$$

Таким образом приходим к выводу, что  $T_{SAT} > T_{LIN}$ , поскольку  $\ell \ll L$ . При высоком напряжении на стоке электроны в высокополевой области канала проводимости обладают бóльшим запасом энергии и рассеиваются более интенсивно, чем при низком  $V_{DS}$ . Тем не менее, коэффициент прохождения выше при высоком  $V_{DS}$  и транзистор демонстрирует ток, близкий к баллистическому пределу.

Расчет протяженности низкополевой области канала в зависимости от напряжений на затворе и стоке потребует, в принципе, самосогласованного решения задачи электростатики в присутствии потока электронов [8, 11]. Когда профиль канала  $E_c(x)$  известен, протяженность критической длины  $\ell$  может быть рассчитана [9, 12, 13]. Альтернативным подходом может быть использование эмпирической функции насыщения стока и скорости впрыскивания в модели виртуального истока.

Приведем численный пример обсуждаемых понятий на образце ETSOI MOSFET длиной  $L = 25 \text{ нм}$  [14]. Параметры транзистора при  $300 \text{ K}$  таковы [15]:

$$\begin{aligned} v_{inj} &= 0.82 \times 10^7 \text{ см/с}, \\ \lambda_0 &= 10.5 \text{ нм}. \end{aligned}$$

Воспользуемся также значением  $v_T = 1.2 \times 10^7 \text{ см/с}$ , согласно (15/гл. 6).

Для  $T_{LIN}$  по (39) имеем:

$$T_{LIN} = \frac{\lambda_0}{\lambda_0 + L} = \frac{10.5}{10.5 + 25} = 0.33.$$

Для вычисления  $T_{SAT}$  пользуемся третьим уравнением из (36):

$$T_{SAT} = \frac{2}{1 + v_T/v_{inj}} = 0.8.$$

Как и ожидалось, прохождение намного больше при высоком напряжении на стоке. Длину критической области  $\ell$  находим из уравнения (40):

$$\ell = \lambda_0 \left( \frac{1}{T_{SAT}} - 1 \right) = 2.6 \text{ нм} \ll L.$$

Таким образом, протяженность узкого места на вершине барьера, определяющего ток, составляет приблизительно 10% от длины канала проводимости.

Может показаться странным, что линейный ток с учетом рассеяния (37) отличается от баллистического тока множителем  $T$ , тогда как ток насыщения (38) – множителем  $T/(2-T)$ . Причиной этого в конечном итоге является электростатика MOS, а объяснение этому обстоятельству может дать рассмотрение ситуации при высоком напряжении на стоке с пренебрежимо малым впрыскиванием электронов на вершину барьера со стороны стока.

Рассмотрим сначала баллистическую ситуацию, показанную наверху рис. 12. Ток  $I_{ball}^+$  впрыскивается истоком. В этом случае заряд на вершине барьера создается лишь электронами, инжектированными истоком. Поскольку ток есть произведение заряда на скорость, то заряд на вершине барьера, в предположении статистики Максвелла – Больцмана,

$$Q|_{x=0} = -\frac{I_{ball}^+}{W v_T}. \quad (41)$$

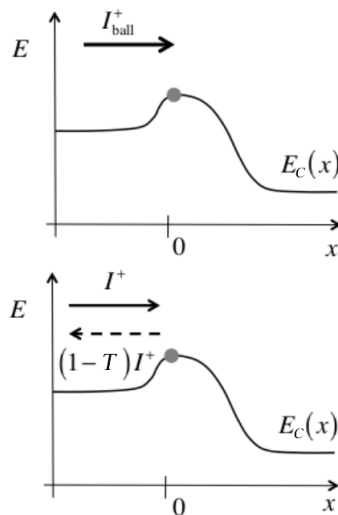


Рис. 12. Инжектированный и рассеянный назад поток электронов при высоком напряжении на стоке: баллистическая ситуация (наверху) и ситуация с рассеянием назад (внизу).

Теперь рассмотрим ситуацию с зарядом при наличии рассеяния. Как показано на рис. 12 внизу, заряд на вершине барьера формируется двумя

потоками электронов: потоком электронов со стороны истока с положительными скоростями и обратным потоком электронов с отрицательными скоростями, так что суммарный заряд

$$\underline{Q}|_{x=0} = -\frac{I^+ + (1-T_{SAT})I^+}{W v_T} = -\frac{(2-T_{SAT})I^+}{W v_T}. \quad (42)$$

В качественном транзисторе электронный заряд на вершине барьера  $\underline{Q}|_{x=0}$  определяется преимущественно электростатикой MOS и не должен зависеть от транспорта: заряд в баллистических условиях (41) должен быть таким же, как и при учете рассеяния (42). Уравнивая эти заряды, находим, что

$$I^+ = \frac{I_{ball}^+}{(2-T_{SAT})}. \quad (43)$$

При учете рассеяния  $T < 1$ , так что инжектированный меньший поток электронов создает на вершине тот же электронный заряд  $\underline{Q}|_{x=0}$ .

Ток стока больше инжектированного тока в  $T$  раз, так что для баллистического случая  $T = 1$

$$I_{DS}^{ball} = I^+ = I_{ball}^+, \quad (44)$$

а для тока  $I_{DS}$  в общем случае ( $T < 1$ ) с учетом (43) находим

$$I_{DS} = T_{SAT} I^+ = \frac{T_{SAT}}{(2-T_{SAT})} I_{ball}^+. \quad (45)$$

Требование решающей роли электростатики MOS приводит к тому, что ток насыщения в условиях рассеяния равен баллистическому току насыщения, помноженному на  $T/(2-T)$ .

Обсудим также какой смысл имеет понятие подвижности в наноразмерных MOSFET. Согласно уравнению (50/гл.1), подвижность пропорциональна средней длине свободного пробега. В транспортной теории понятие подвижности хорошо определено вблизи равновесия в массивном проводнике с длиной, во много раз превышающей среднюю длину свободного пробега [1]. В современных транзисторах длина канала проводимости сравнима со средней длиной свободного пробега и при высоком напряжении на стоке электроны далеки от состояния равновесия. И тем не менее, практика показывает, что квазиравновесная подвижность надежно описывает функционирование

наноразмерных транзисторов. Как объяснить значимость подвижности для наноразмерных MOSFET? Как уже обсуждалось в связи с рис. 11, почти равновесная средняя длина свободного пробега  $\lambda_0$  контролирует ток во всем диапазоне изменения напряжения на стоке. Пусть равновесный поток электронов впрыскивается истоком. При низком напряжении на стоке эти электроны остаются квазиравновесными во всем канале проводимости. При высоком напряжении на стоке электроны приобретают энергию, их скорость рассеяния увеличивается, а средняя длина свободного пробега уменьшается. Однако, как мы уже убедились, именно низкополевая область канала ответственна за прохождение. Электроны квазиравновесны в этой области канала проводимости, их средняя длина свободного пробега контролирует ток как при низких, так и при высоких напряжениях на стоке.

Экспериментально подтверждаемую корреляцию между функциональностью нанотранзисторов и подвижностью можно объяснить тем, что подвижность пропорциональна квазиравновесной средней длине свободного пробега, которая и контролирует ток нанотранзистора от низких до высоких значений напряжений на стоке. Конечно, это слишком очевидный аргумент. В короткоканальных транзисторах могут случиться перепады с механическим напряжением и допированием, так что электроны не находятся точно в равновесии. В очень коротких каналах электроны, входящие в канал со стороны истока, могут возбудить плазменные колебания вблизи истока, которые понизят среднюю длину свободного пробега [14, 16]. Этот эффект наблюдался экспериментально, однако, утверждение о том, что ток при высоком напряжении на стоке надежно коррелирует с квазиравновесной подвижностью, как представляется, вполне соответствует сложившимся представлениям о физике процессов в нанотранзисторах и приводит к достаточно точным результатам на практике.

Подведем итоги. Подробно рассмотрена качественная картина явлений рассеяния электронов в канале проводимости нанотранзисторов и были сформулированы важнейшие результаты в виде сводки формул (18), в которой ключевым является понятие о коэффициенте прохождения  $T(E)$ , связывающем среднюю длину свободного пробега назад  $\lambda(E)$  с длиной канала проводимости  $L$ . Затем обобщенная теория транспорта Ландауэра – Датта – Лундстрема позволила построить модель прохождения MOSFET с учетом рассеяния электронов. Ограничения, накладываемые электростатикой MOS, учтены в выражении Ландауэра для тока (19). Итогом явилась модель прохождения



MOSFET в виде сводки формул (36) для вычисления линейной области и области насыщения выходной характеристики нанотранзистора, которую полезно сравнить с баллистической моделью MOSFET (69/гл. 1). Если транзистор работает в подпороговом режиме, можно воспользоваться невырожденной статистикой. Однако, в режиме выше порога зона проводимости на вершине барьера близка или даже ниже уровня Ферми, так что нужно пользоваться статистикой Ферми – Дирака. Тем не менее, стало практикой в теории устройств со структурой MOS использовать невырожденную статистику Максвелла – Больцмана, поскольку использование ее сильно упрощает расчеты, ведет к более простым формулам и делает теорию более понятной. На практике также обычно некоторые параметры транзистора не известны с достаточной точностью, и в случае использования невырожденной статистики они могут рассматриваться в качестве параметров подгонки, обеспечивая вполне приемлемые результаты.

Полученные уравнения хорошо согласуются с физикой процессов в линейной области, как и в области насыщения, чего нельзя сказать о всей области изменения напряжения на стоке, поскольку у нас нет достаточно надежной модели для зависимости  $T(V_{DS})$ . Далее будет показано как можно объединить модель прохождения и модель виртуального истока и таким образом выйти на моделирование всей цельной выходной характеристики MOSFET.

## Литература

1. M. Lundstrom, *Fundamentals of Carrier Transport* (Cambridge, U.K.: Cambridge Univ. Press: 2000).
2. M. Lundstrom, C. Jeong, *Near-equilibrium transport. Fundamentals and Applications* (Singapore: World Scientific: 2013).
3. Ю. А. Кругляк, *Нанoeлектроника «снизу – вверх»* (Одесса: ТЭС: 2015).
4. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
5. P. A. M. Dirac, *Proc. Royal Soc. A*, **114**, 243 – 265 (1927).
6. E. Fermi, *Nuclear Physics* (Chicago: Univ. of Chicago Press: 1950).
7. P. J. Price, *Semiconductors and Semimetals*, **14**, 249 – 308 (1979).
8. P. Palestri, D. Esseni, S. Eminent, C. Fiegna, E. Sangiorgi, L. Selmi, *IEEE Trans. Electron Dev.*, **52**: 2727 – 2735 (2005).
9. R. Clerc, P. Palestri, L. Selmi, G. Ghibaudo, *J. Appl. Phys.*, **110**, 104502 (2011).
10. Changwook Jeong, Raseong Kim, M. Luisier, S. Datta, M. Lundstrom, *J. Appl. Phys.*, **107**, 023707 (2010).
11. M. V. Fischetti, T. P. O`Regan, N. Sudarshan, C. Sachs, S. Jin, J. Kim, Y. Zhang, *IEEE Trans. Electron Dev.*, **54**: 2116 – 2136 (2007).
12. G. Gildenblat, *J. Appl. Phys.*, **91**, 9883 – 9886 (2002).
13. S. Rakheja, M. Lundstrom, D. Antoniadis, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 35.1.1 – 35.1.4 (2014).
14. T. Uechi, T. Fukui, N. Sano, *Phys. Status Solidi C*, **5**, 102 – 106 (2008).
15. A. Majumdar, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **61**: 351 – 358 (2014).
16. M. V. Fischetti, S. E. Laux, *J. Appl. Phys.*, **89**, 1205 – 1231 (2001).



## Глава 8. Объединение модели прохождения и модели виртуального истока: модель MVS-прохождения

### 8.1. Введение

В заключение предстоит объединить модель прохождения и модель виртуального истока. Уравнения (36/гл.7) дают сводку формул модели прохождения для расчета выходных характеристик MOSFET. Ток стока в модели виртуального истока в нужной точке  $(V_{GS}, V_{DS})$  выходной характеристики вычисляется по уравнению (18/гл. 6), используя для вычисления заряда (19/гл. 6), а для вычисления скорости (24 – 26/гл. 6). Установим связь между этими двумя моделями.

Начинаем с уравнения (18/гл. 6) для тока в виде произведения заряда на скорость:

$$I_{DS} = W \left| \underline{Q(V_{GS}, V_{DS})} \right|_{x=0} \times \left| \underline{v_x(V_{GS}, V_{DS})} \right|_{x=0}. \quad (1)$$

Электронный заряд вычисляем из электростатики MOS. Далее нужно определить среднюю скорость электронов на вершине барьера. Она вычисляется по разному в модели прохождения и в модели виртуального истока.

#### *Резюме модели прохождения*

Пусть поведение электронов в канале проводимости подчиняется статистике Максвелла – Больцмана (*MB*). Ток дается уравнением (36/гл. 7). Заряд в нужной точке  $(V_{GS}, V_{DS})$  определяется электростатикой MOS. В случае статистики *MB* нет нужды знать значение энергии фермиевского уровня. Скорость впрыскивания

$$v_{inj} = v_T \left( \frac{T}{2-T} \right), \quad (2)$$

где баллистическая скорость впрыскивания в максвелл-больцмановском пределе  $v_T$  дается выражением (25/гл. 6):

$$v_T = \sqrt{\frac{2kT}{\pi m^*}}. \quad (3)$$

Средняя скорость электронов на вершине барьера в точке  $(V_{GS}, V_{DS})$

$$\left. \underline{v_x(V_{GS}, V_{DS})} \right|_{x=0} = F_{SAT}(V_{DS}) v_{inj}, \quad (4)$$

где

$$F_{SAT}(V_{DS}) = \left[ \frac{1 - e^{-qV_{DS}/kT}}{1 + \left(\frac{T}{2-T}\right) e^{-qV_{DS}/kT}} \right]. \quad (5)$$

Ток вычисляем окончательно по (1) с учетом терминальных сопротивлений.

Трудности в этом алгоритме вычислений связаны только лишь с расчетом  $T(V_{DS})$ . Для низких  $V_{DS}$  прохождение берется по (39/гл. 7):

$$T_{LIN} = \frac{\lambda_{LIN}}{\lambda_{LIN} + L}, \quad (6)$$

а для высоких  $V_{DS}$  – по (40/гл. 7):

$$T_{SAT} = \frac{L_{SAT}}{L_{SAT} + \ell}. \quad (7)$$

Как уже обсуждалось ранее,

$$\lambda_{LIN} \approx \lambda_{SAT} = \lambda_0. \quad (8)$$

Протяженность критической длины  $\ell$  оценить не так просто [1 – 3], а вот выражения для линейного тока и тока насыщения в MOSFET ЛДЛ легко увязать с уравнениями в модели VS. Согласно (37/гл.7) и (38/гл.7), эти выражения следующие:

$$I_{DLIN} = W |Q(V_{GS})| T_{LIN} \frac{v_T}{2kT/q} V_{DS}, \quad (9)$$

$$I_{DSAT} = W |Q(V_{GS})| v_{inj} = W |Q(V_{GS})| \left( \frac{T_{SAT}}{2 - T_{SAT}} \right) v_T.$$

Сейчас убедимся, что эти уравнения легко увязать с соотношениями, принятыми в традиционной модели MOSFET и в модели виртуального истока.

#### *Резюме модели виртуального истока*

Эта модель также начинается с уравнения (1), но средняя скорость на вершине барьера вычисляется согласно (24/гл. 6)

$$\langle v_x(V_{GS}, V_{DS}) \rangle_{x=0} = F_{SAT}(V_{DS}) v_{sat}, \quad (10)$$

где зависимость средней скорости от напряжения на стоке дается эмпирической функцией насыщения тока стока (25/гл. 6)

$$F_{SAT}(V_{DS}) = \frac{V_{DS}/V_{DSAT}}{\left[1 + (V_{DS}/V_{DSAT})^\beta\right]^{1/\beta}}, \quad (11)$$

где

$$V_{DSAT} = \frac{v_{sat}L}{\mu}. \quad (12)$$

Заряд вычисляется, скажем, по (19/гл. 6). Учесть нужно также терминальные сопротивления.

Для небольших напряжений на стоке

$$F_{SAT} \rightarrow V_{DS}/V_{DSAT}$$

и

$$\langle v_x(V_{GS}, V_{DS}) \rangle|_{x=0} \rightarrow \mu V_{DS}/L.$$

Линейный ток в модели VS

$$I_{DLIN} = \frac{W}{L} |Q(V_{GS})| \mu V_{DS} \quad (13)$$

совпадает с уравнением в традиционной модели MOSFET. Для больших значений  $V_{DS}$  уравнение (10) сводится к традиционному выражению для скорости насыщения

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| v_{sat}. \quad (14)$$

По существу, модель виртуального истока – это полуэмпирическая модель, позволяющая вписаться в экспериментально измеренную выходную характеристику. В случае короткоканальных MOSFET параметры  $\mu$  и  $v_{sat}$  нужно подогнать:

$$\mu \rightarrow \mu_{app}, \quad v_{sat} \rightarrow v_{inj}. \quad (15)$$

В главе 6 было показано, что в баллистическом пределе кажущаяся подвижность  $\mu_{app}$  и скорость впрыскивания  $v_{inj}$  имеют ясный физический смысл. Далее мы интерпретируем эти два параметра уже с учетом рассеяния.

## 8.2. Объединение модели прохождения и модели виртуального истока

Наша задача сейчас, объединяя обе модели, понять физический смысл кажущейся подвижности и скорости впрыскивания.

*Линейная область*

Используя выражение (39/гл.7) для прохождения  $T_{LIN} = \lambda_0/(\lambda_0 + L)$ , перепишем (9) следующим образом:

$$I_{DLIN} = \frac{W}{L} |Q(V_{GS})| (T_{LIN} L) \left( \frac{v_T}{2kT/q} \right) V_{DS} = \frac{W}{L} |Q| \left( \frac{1}{(1/\lambda_0) + (1/L)} \right) \left( \frac{v_T}{2kT/q} \right) V_{DS}. \quad (16)$$

Вспоминаем определение подвижности (50/гл. 5)

$$\mu = \frac{D}{kT/q} = \frac{v_T \lambda_0}{2kT/q} \quad (17)$$

и баллистической подвижности (54/гл. 5)

$$\mu_B = \frac{v_T L}{2kT/q}. \quad (18)$$

Используя их, переписываем (16) в виде

$$I_{DLIN} = \frac{W}{L} |Q| \left( \frac{1}{(1/\mu) + (1/\mu_B)} \right) V_{DS} = \frac{W}{L} |Q| \mu_{app} V_{DS}, \quad (19)$$

где кажущаяся подвижность определяется как

$$\frac{1}{\mu_{app}} = \frac{1}{\mu} + \frac{1}{\mu_B}. \quad (20)$$

Подобный рецепт нахождения результирующей величины как суммы двух независимых процессов применительно к электрическому сопротивлению известен как правило Матиссена [2].

Как уже обсуждалось в главе 6, баллистическая подвижность соответствует такой подвижности, когда средняя длина свободного пробега становится равной длине канала проводимости. Электроны рассеиваются интенсивно в контактах истока и стока, так что когда канал проводимости оказывается баллистическим, то расстояние между двумя последовательными актами рассеяния оказывается равным длине канала проводимости. Используя

баллистическую подвижность в (19), линейный ток в баллистическом MOSFET может быть записан в традиционной диффузионной форме (13).

Согласно (20), кажущаяся подвижность MOSFET всегда меньше наименьшей из двух подвижностей – баллистической  $\mu_B$  или диффузионной  $\mu$ . Для протяженного канала MOSFET  $\mu \ll \mu_B$ : кажущаяся подвижность ограничена диффузионной подвижностью. В случае нанотранзисторов  $\mu_B \ll \mu$ : кажущаяся подвижность определяется баллистической подвижностью. Обратим внимание на то, что традиционное выражение для линейного тока (13) может ошибочно предсказать ток выше баллистического предела, если канал проводимости достаточно короткий, но если традиционную подвижность заменить на кажущуюся подвижность, то это не случится.

В линейной области MOSFET ведет себя как резистор, контролируемый приложенным напряжением на затворе (рис. 1).

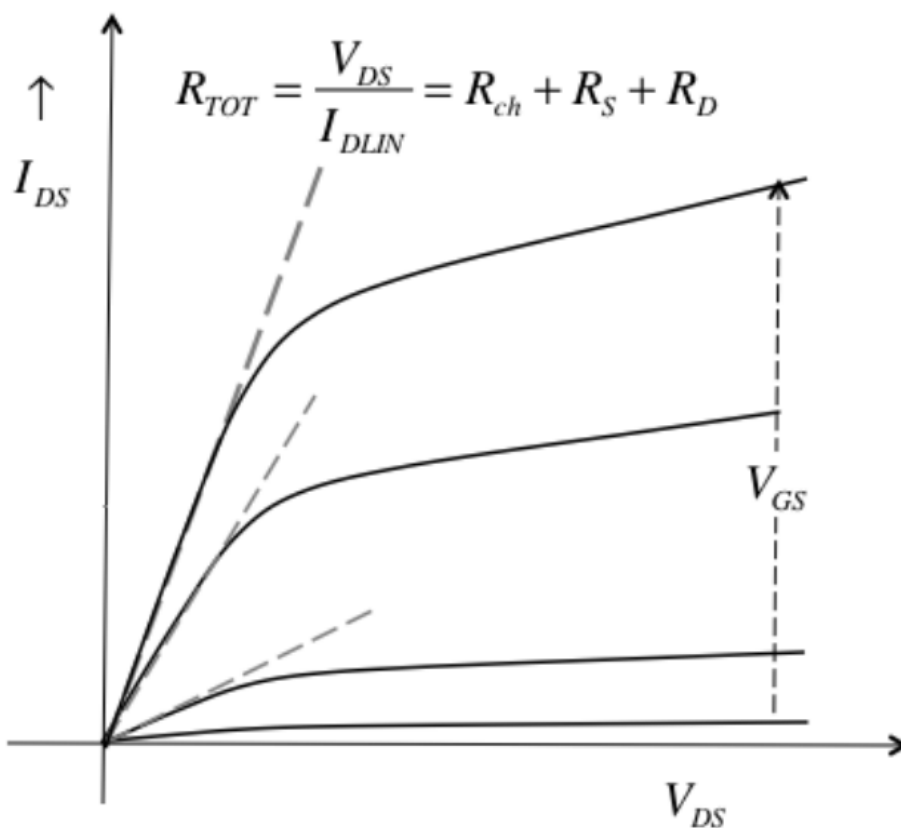


Рис. 1. Иллюстрация того, как линейный ток связан с сопротивлением канала и терминальными сопротивлениями. Для заданного напряжения на затворе сопротивление канала проводимости (21) обратно пропорционально кажущейся подвижности.



Из (19) следует, что сопротивление канала проводимости

$$R_{ch} = \frac{V_{DS}}{I_{DLIN}} = \frac{L}{W} \frac{1}{|Q| \mu_{app}}. \quad (21)$$

На практике обычно учитываются терминальные сопротивления, так что линейный ток

$$I_{DLIN} = \frac{V_{DS}}{R_{ch} + R_S + R_D} = \frac{V_{DS}}{R_{TOT}}. \quad (22)$$

Вписывая экспериментально измеренную входную характеристику в ее линейной области в модель виртуального истока, можно извлечь значения терминальных сопротивлений и кажущейся подвижности.

Итак, было показано, что линейный ток в модели прохождения (19) может быть записан в диффузионной форме (13), принятой в модели виртуального истока, если традиционную диффузионно контролируруемую подвижность  $\mu$  заменить на кажущуюся подвижность  $\mu_{app}$ .

#### *Область насыщения*

Уравнение (9) свидетельствует о том, что в области насыщения важную роль играет множитель  $T_{SAT} / (2 - T_{SAT})$ . Согласно (7) с учетом (8) этот множитель можно переписать в виде

$$\frac{T_{SAT}}{2 - T_{SAT}} = \frac{\lambda_0}{\lambda_0 + 2\ell}. \quad (23)$$

Тогда скорость впрыскивания (2)

$$v_{inj} = \frac{T_{SAT}}{2 - T_{SAT}} v_T = \frac{\lambda_0 v_T}{\lambda_0 + 2\ell} = \frac{1}{(1/v_T) + \ell / (\lambda_0 v_T / 2)}. \quad (24)$$

Вспоминая определение коэффициента диффузии (19/гл.5)  $D = v_T \lambda_0 / 2$ , перепишем (24) иначе:

$$v_{inj} = \left( \frac{1}{v_T} + \frac{1}{D/\ell} \right)^{-1} \quad (25)$$

или

$$\frac{1}{v_{inj}} = \frac{1}{v_T} + \frac{1}{D/\ell}. \quad (26)$$

Согласно (26), приходим к выводу, что скорость впрыскивания в MOSFET всегда меньше, чем наименьшая из двух величин – скорости баллистического впрыскивания и величины  $D/\ell$ , равной скорости, с которой электроны диффундируют через узкое место на вершине барьера длины  $\ell$ . Когда эта длина  $\ell$  достаточно протяженная или коэффициент диффузии  $D$  небольшой, то  $D/\ell \ll v_T$ , и скорость впрыскивания контролируется скоростью диффузии. Когда же длина  $\ell$  невелика или коэффициент диффузии  $D$  большой, то  $D/\ell \gg v_T$ , и скорость впрыскивания ограничена баллистической скоростью впрыскивания: скорость впрыскивания не может быть больше баллистической скорости впрыскивания, но может быть намного меньше.

Рисунок 2 служит иллюстрацией того, что происходит в наноразмерном MOSFET в режиме «ON». Электроны диффундируют через узкое место канала, но они не могут двигаться со скоростью быстрее, чем тепловая скорость, поскольку причиной диффузии является случайное тепловое движение. После диффузии через узкое место электроны попадают в высокополевую область канала проводимости, в которой электрическое поле подхватывает электроны и выносит их в сток. Узкое место канала аналогично базе в биполярном транзисторе, а высокополевая область аналогична коллектору.

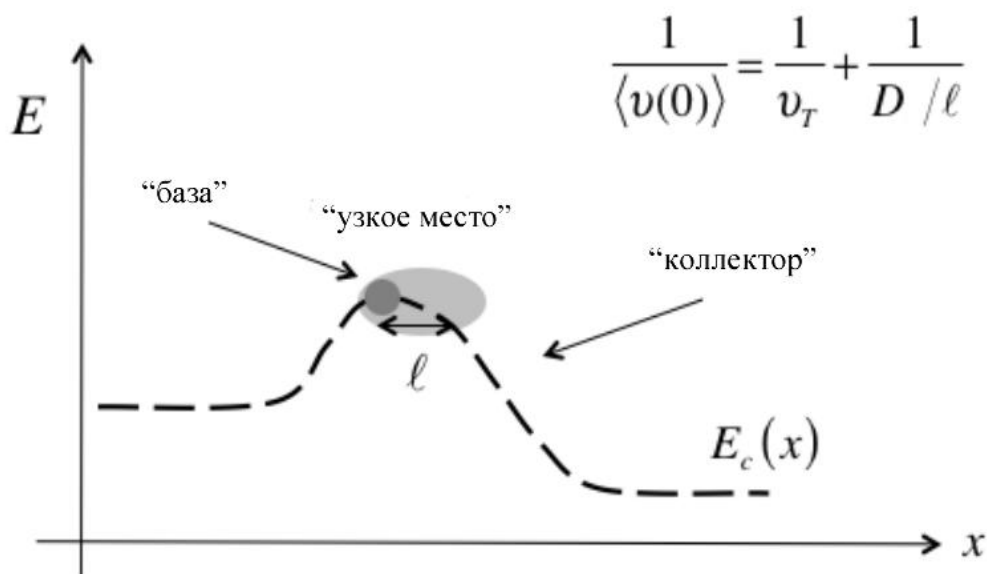


Рис. 2. Зонная диаграмма MOSFET в режиме «ON», демонстрирующая узкое место для тока с небольшим электрическим полем и высокополевую область канала проводимости. Это узкое место аналогично базе в биполярном транзисторе, а высокополевая область аналогична коллектору.

Таким образом, было показано, что ток насыщения (9) в модели прохождения может быть записан в традиционной форме (14) через скорость насыщения, принятой в модели VS, если традиционную диффузионно контролируемую скорость  $v_{sat}$  заменить на скорость впрыскивания  $v_{inj}$  (26). Предельным значением скорости впрыскивания является баллистическая скорость впрыскивания  $v_T$ .

Вписывая экспериментальные данные в модель виртуального истока, мы находим кажущуюся подвижность  $\mu_{app}$  и скорость впрыскивания  $v_{inj}$ . Если экспериментальные данные относятся к длинноканальному транзистору, то определим и стандартную диффузионную подвижность  $\mu$ . Позже мы покажем, что традиционную подвижность  $\mu$  можно определить и в случае коротких каналов в нанотранзисторах. Предполагая, что нам известны эти три параметра, покажем как вычислить коэффициент прохождения в линейной области и в области насыщения.

Уравнение (9) дает линейный ток через  $T_{LIN}$ , а (19) – через  $\mu_{app}$ . Уравнивая эти два выражения, получаем

$$T_{LIN} = \frac{\mu_{app}}{L} \left( \frac{v_T}{2kT/q} \right)^{-1} = \frac{\mu_{app}}{\mu_B},$$

а используя далее определение кажущейся подвижности по (20), окончательно находим, что

$$T_{LIN} = \frac{\mu_B \mu}{\mu_B + \mu} \times \frac{1}{\mu_B} = \frac{\mu}{\mu_B + \mu}. \quad (27)$$

Для вычисления  $T_{SAT}$  воспользуемся выражением (2)

$$v_{inj} = v_T \left( \frac{T_{SAT}}{2 - T_{SAT}} \right),$$

откуда для  $T_{SAT}$  имеем

$$T_{SAT} = \frac{2}{1 + v_T/v_{inj}}. \quad (28)$$

Скорость впрыскивания  $v_{inj}$  определяется путем вписывания опытных данных в модель VS, однако, баллистическую скорость впрыскивания  $v_T$  определить труднее. Ее можно извлечь из анализа экспериментальных входных

характеристик [5], вместе с тем ее часто вычисляют пользуясь известными эффективными массами и гипотезой о числе занятых подзон.

Оценим  $\mu_B$ ,  $\mu_{app}$  и  $T_{LIN}$  для линейной области  $n$ -MOSFET с длиной канала  $L = 22 \text{ нм}$  при  $T = 300 \text{ К}$ . Пусть в подложке Si (100) заселена только одна нижняя подзона, а подвижность  $\mu = 250 \text{ см}^2/\text{В} \cdot \text{с}$ .

В этих условиях, согласно (15/гл.6),  $v_T = 1.2 \times 10^7 \text{ см/с}$ . Баллистическую подвижность находим по (18)

$$\mu_B = \frac{v_T L}{2kT/q} = \frac{(1.2 \times 10^7) \times (22 \times 10^{-7})}{2 \times 0.026} = 508 \text{ см}^2/\text{В} \cdot \text{с}. \quad (29)$$

Поскольку  $\mu_B$  сопоставима с  $\mu$ , этот транзистор работает в квазибаллистическом режиме.

Кажущуюся подвижность находим из (20), а именно:

$$\mu_{app} = \frac{\mu \mu_B}{\mu + \mu_B} = \frac{250 \times 508}{250 + 508} = 191 \text{ см}^2/\text{В} \cdot \text{с}.$$

Как и ожидалось, кажущаяся подвижность меньше наименьшей из двух подвижностей – баллистической и диффузионной. И, наконец, находим коэффициент прохождения в линейной области по (27):

$$T_{LIN} = \frac{\mu}{\mu_B + \mu} = \frac{250}{508 + 250} = 0.33. \quad (30)$$

Мы убедились, что как для линейной области выходной характеристики, так и для области насыщения всегда можно установить соответствие между результатами, полученными в модели виртуального истока и в модели прохождения. Теперь нам ясно, почему традиционную диффузионную подвижность, применимую для длинноканальных транзисторов, нужно заменять на кажущуюся подвижность, охватывающую весь квазибаллистический транспорт. Также было показано, что скорости насыщения в традиционной модели MOSFET соответствует скорость впрыскивания в модели прохождения. Эта модель предоставляет физически ясное понимание изменения как линейного тока, так и тока насыщения в нанотранзисторах, однако, полуэмпирическая модель виртуального истока лучше приспособлена для вписывания экспериментальных данных  $I_{DS} \propto V_{DS}$  во всем диапазоне изменения напряжения на стоке. Это различие между моделями

не носит фундаментальный характер, оно возникает только из-за трудностей вычисления зависимости  $T(V_{DS})$ .

Мы до сих пор встречались с тремя разными подвижностями: 1) традиционная диффузионная подвижность  $\mu$ , 2) баллистическая подвижность  $\mu_b$  и 3) кажущаяся подвижность  $\mu_{app}$ . В теории MOSFET нередко используется т. наз. эффективная подвижность  $\mu_{eff}$  [6]. Термин «эффективная подвижность» – неудачный, он подразумевает то обстоятельство, что электроны вблизи границы раздела Si/SiO<sub>2</sub> должны иметь меньшую подвижность по сравнению с электронами в глубине канала проводимости, в результате дополнительного рассеяния на шероховатостях поверхности раздела. Эффективная подвижность – это подвижность электронов в канале проводимости, усредненная по глубине. Для Si MOSFET  $\mu_{eff}$  намного меньше диффузионной подвижности  $\mu$  в массивной кремниевой подложке. Для транзисторов III-V НЕМТ высокая подвижность сохраняется благодаря технологической возможности создания атомарно регулярной границы раздела. В современных нанотранзисторах, однако, квантовые ограничения столь сильны, что все электроны в канале проводимости испытывают рассеяние на шероховатостях поверхности раздела. Говорить об усреднении подвижности по глубине применительно к нанотранзисторам не приходится. Обе подвижности стоит уравнивать,  $\mu_{eff} = \mu$ , и рассматривать их как диффузионную подвижность.

### **8.3. Транспорт в нанотранзисторах с позиции модели MVS/прохождения**

О физике процессов в нанотранзисторах можно многое выяснить из анализа вольт-амперных характеристик наноразмерных MOSFET. Разработано немало подходов для такого анализа [7 – 12]. Перспективность объединенной модели VS/прохождения хорошо зарекомендовала себя при изучении транспортных явлений в нанотранзисторах [13 – 16]. Продемонстрируем это [17, 18] путем анализа экспериментальных результатов на примере как ETSOI MOSFET [19, 20], так и III-V НЕМТ [16, 21].

#### *Обзор модели MVS/прохождения*

Прежде чем переходить к анализу экспериментальных данных приведем основные положения последней версии этой объединенной модели, получившей название модели MVS/прохождения, поскольку она была сформулирована в Массачусетском технологическом институте [22].

Модель MVS описывает ток стока в виде произведения электронного заряда на скорость электронов [22 – 24], а именно:

$$I_{DS} = W \left| \underline{Q(V_{Gi}, V_{Di})} \right|_{x=0} \times F_{SAT}(V_{Di}) v_{inj}, \quad (31)$$

где  $F_{SAT}(V_{Di}) v_{inj}$  – это скорость электронов на виртуальном истоке. Напряжения  $V_{Gi}$  и  $V_{Di}$  – это собственные (intrinsic) напряжения на затворе и на стоке. Электронный заряд взят по абсолютной величине, поскольку инверсионный заряд отрицательный в  $n$ -MOSFET.

В модели MVS заряд на виртуальном истоке  $Q(V_{Gi}, V_{Di})$  вычисляется из полуэмпирического выражения, подобного (45/гл. 4) [22], а именно:

$$\left| \underline{Q(V_{Gi}, V_{Di})} \right|_{x=0} = m C_G^{inv} \left( \frac{kT}{q} \right) \ln \left\{ 1 + \exp \left[ \frac{q(V_{Gi} - V_T - \alpha(kT/q) F_f)}{mkT} \right] \right\}. \quad (32)$$

В этом выражении используется инверсионная переходная функция (inversion transition function) [22]

$$F_f = \frac{1}{1 + \exp \left\{ \frac{V_{Gi} - [V_T - \alpha(kT/q)/2]}{\alpha kT/q} \right\}}, \quad (33)$$

которая ответственна за увеличение порогового напряжения на величину  $\alpha(kT/q)/2$  по мере того как MOSFET переходит из подпорогового режима в режим сильной инверсии:  $F_f \rightarrow 1$  в подпороговом режиме и  $F_f \rightarrow 0$  в режиме сильной инверсии. Эмпирический параметр  $\alpha$  обычно выбирают равным 3.5 [17, 22]. В уравнении (32) пороговое напряжение зависит от напряжения на стоке согласно (47/гл. 4), а именно:

$$V_T = V_{T0} - \delta V_{Di}, \quad (34)$$

где  $V_{T0}$  – это пороговое напряжение в режиме сильной инверсии при  $V_{DD} = V_{Di} = 0$ , а  $\delta$  – это параметр *DIBL*. Подпороговый параметр наклона в (32)

$$m = m_0 + m' V_{Di}, \quad (35)$$

где  $m_0$  – это подпороговый параметр при  $V_{DD} = V_{Di} = 0$ , а  $m' = dm/dV_{Di}$  описывает изменение параметра  $m$  с изменением напряжения на стоке.

В модели MVS используется эмпирическая функция насыщения тока стока (45/гл. 4) [22]

$$F_{SAT}(V_{Di}) = \frac{V_{Di}N_{DSATs}}{\left[1 + (V_{Di}N_{DSATs})^\beta\right]^{1/\beta}} \quad (36)$$

с (56/гл. 4)

$$V_{DSATs} = \frac{v_{inj}L_{eff}}{\mu_{app}}, \quad (37)$$

где  $L_{eff}$  – это эффективная длина канала проводимости по [25]. Обращаем внимание, что в индекс  $SAT$  в  $V_{DSATs}$  добавлена буква  $s$ , чтобы учесть тот факт, что  $F_{SAT}$  описывает насыщение тока стока в режиме сильной инверсии. В условиях подпорогового режима  $V_{DSAT} = kT/q$  [25]. Модель MVS рассматривает этот переход от  $V_{DSAT}$  в подпороговом режиме в режим сильной инверсии эвристически с использованием инверсионной переходной функции (33) [22], а именно:

$$V_{DSAT} = V_{DSATs}(1 - F_f) + (kT/q)F_f. \quad (38)$$

Собственные терминальные напряжения  $V_{Di}, V_{Gi}$  связаны с внешними терминальными напряжениями  $V_{DS}, V_{GS}$  следующими соотношениями:

$$\begin{aligned} V_{Di} &= V_{DS} - I_{DS}R_{SD0}, \\ V_{Gi} &= V_{GS} - I_{DS}R_{SD0}/2, \end{aligned} \quad (39)$$

где суммарное терминальное сопротивление  $R_{SD0} = R_{S0} + R_{D0}$  есть сумма сопротивления истокового терминала  $R_{S0}$  и сопротивления стокового терминала  $R_{D0}$ , которые предполагаются одинаковыми и независимыми от напряжений на затворе и на стоке.

Модель MVS может быть вписана в экспериментально измеренные передаточные характеристики  $I_{DS} \propto V_{GS}$  и выходные характеристики  $I_{DS} \propto V_{DS}$ , что позволяет определить несколько важных параметров MOSFET. Дальнейший анализ сосредоточим на линейной области при низких  $V_{DS}$  и на области насыщения при высоких  $V_{DS}$ .

При небольших напряжениях на стоке

$$F_{SAT} \rightarrow \frac{V_{DS}}{V_{DSATs}}, \quad \frac{v_x(V_{GS}, V_{DS})|_{x=0}}{\mu_{app}V_{DS}} \rightarrow \frac{\mu_{app}V_{DS}}{L_{eff}}.$$

Уравнение для тока (31) в линейной области становится таким:

$$I_{DLIN} = \frac{W}{L_{eff}} |Q(V_{GS})| \mu_{app} V_{DS} = \frac{V_{DS}}{R_{ch}}, \quad (40)$$

где  $R_{ch}$  – сопротивление канала проводимости.

Для больших напряжений на стоке  $F_{SAT} \rightarrow 1$  и уравнение (31) сводится к традиционному выражению для тока насыщения

$$I_{DSAT} = W |Q(V_{GS}, V_{DS})| v_T, \quad (41)$$

где

$$v_T = \sqrt{\frac{2kT}{\pi m^*}} = v_{inj}^{ball} \quad (42)$$

есть баллистическая скорость впрыскивания в случае использования статистики Максвелла – Больцмана. Обратим внимание на то, что баллистическую скорость впрыскивания на практике вычислить трудно. Физические напряжения и квантовые ограничения в нанотранзисторах могут сказаться на величине  $m^*$ , к тому же уравнение (42) предполагает, что только одна нижняя подзона заселена, что не всегда верно.

Кажущаяся подвижность (20) в модели MVS зависит от  $L_{eff}$  и дается выражением

$$\frac{1}{\mu_{app}(L_{eff})} = \frac{1}{\mu} + \frac{1}{\mu_B(L_{eff})}, \quad (43)$$

где диффузионная подвижность (17)

$$\mu = \frac{D}{kT/q} = \frac{v_T \lambda_0}{2kT/q}, \quad (44)$$

а баллистическая подвижность (18)

$$\mu_B(L_{eff}) = \frac{v_T L_{eff}}{2kT/q}. \quad (45)$$

Скорость впрыскивания при высоком напряжении на стоке определяется из (26), а именно:

$$\frac{1}{v_{inj}} = \frac{1}{v_T} + \frac{1}{D/\ell}, \quad (46)$$



где  $\ell \ll L_{eff}$  и

$$D = \frac{v_T \lambda_0}{2}. \quad (47)$$

Мы предположили, что в линейной области средняя длина свободного пробега  $\lambda_{LIN}$  совпадает со средней длиной свободного пробега в области насыщения  $\lambda_{SAT}$ . Хотя равенство  $\lambda_{LIN} = \lambda_{SAT} = \lambda_0$ , строго говоря, не совсем корректно, однако, физически оно вполне приемлемо [26] и подтверждается в экспериментах [17]. Наконец, полезно вспомнить как параметры в модели MVS связаны с коэффициентом прохождения. Из уравнения (27) для линейной области имеем

$$T_{LIN} = \frac{\lambda_0}{\lambda_0 + L_{eff}} = \frac{\mu_{app}}{\mu_B} = \frac{\mu}{\mu_B + \mu}, \quad (48)$$

а из (28) для области насыщения

$$T_{SAT} = \frac{\lambda_0}{\lambda_0 + \ell} = \frac{2}{1 + v_T/v_{inj}}. \quad (49)$$

Измеряемая скорость впрыскивания связана с коэффициентом прохождения как

$$v_{inj} = v_T \left( \frac{T_{SAT}}{2 - T_{SAT}} \right). \quad (50)$$

Мы привели основные результаты новейшей версии модели MVS/прохождения, в основе которой лежит объединенная модель прохождения и модель виртуального истока. Как только экспериментально измеренные вольт-амперные характеристики вписаны в модель MVS, результатами измерений будем считать суммарное терминальное сопротивление  $R_{SD0}$ , кажущуюся подвижность  $\mu_{app}$  и скорость впрыскивания  $v_{inj}$ . Мы также увидим, что из экспериментальных данных можно извлечь сведения о баллистической скорости впрыскивания, диффузионной подвижности, средней длине свободного пробега, критической длине и о коэффициентах прохождения в линейной области и в области насыщения.

## 8.4. ETSOI MOSFET и III-V HEMT

Проанализируем Si MOSFET, имеющий простую и тщательно детализированную физическую структуру, что весьма благоприятствует поставленной задаче. Как показано на рис. 3, это Si SOI (silicon-on-insulator) структура с исключительно тонким слоем SOI толщиной  $T_{SOI} = 6.1 \pm 0.4 \text{ нм}$  [17].

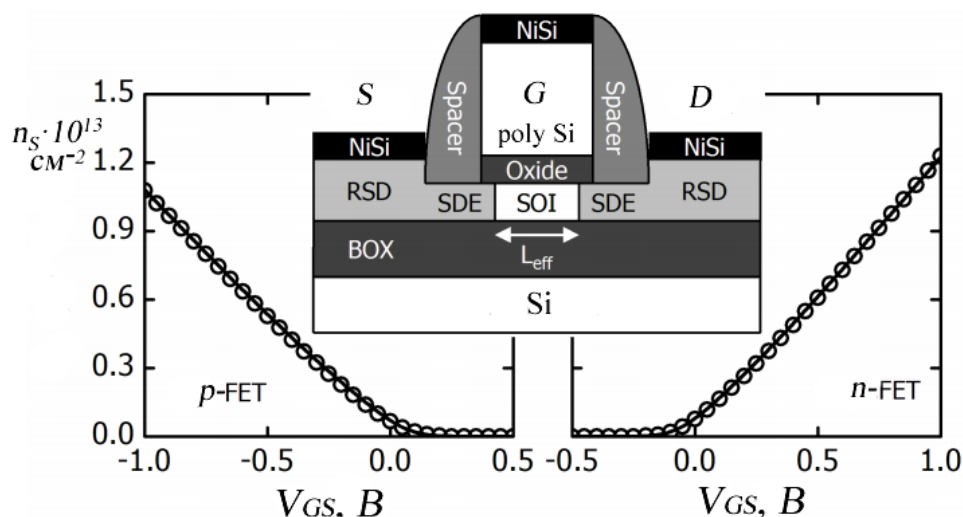


Рис. 3. Сечение ETSOI MOSFET и зависимость  $n_s \propto V_{GS}$ , полученная вписыванием экспериментальных данных (кружки) в модель VS (линии). Показаны SD область расширения (SD extension region/SDE), расширенный SD (raised SD/RSD), слой углубленного оксида (buried oxide/BOX) толщиной  $145 \text{ нм}$  и длина канала проводимости  $L_{eff}$  [17].

Плоскость канала – (100), направление транспорта электронов –  $\langle 110 \rangle$ . Электрод затвора – poly Si и оксид затвора – SiON с толщиной эквивалентной емкости  $CET = 1.1 \text{ нм}$ . Емкость затвора в режиме сильной инверсии (21/гл. 3)  $C_G^{inv}$  определена из вольт-емкостных (C-V) измерений на длинноканальном транзисторе [17]. Для рассматриваемого конкретного n-FET  $C_G^{inv} = 1.98 \text{ мФ} \cdot \text{см}^{-2}$  [17]. Измеренная квазиравновесная подвижность для длинноканального транзистора оказалась равной  $350 \text{ см}^2/\text{В} \cdot \text{с}$ , что соответствует средней длине свободного пробега  $15.8 \text{ нм}$ .

В этом транзисторе использовались прокладки для снятия механического напряжения (neutral stress liners), так что кремниевый канал номинально не напряжен, что упрощает вычисление  $v_T$ . Предполагая, что  $m^* = 0.22m_0$ , находим  $v_T = 1.14 \text{ см}/\text{с}$ . Технологический процесс создания SD областей расширения привел к очень резким p-n-переходам с низким значением суммарного

терминального сопротивления [19]. Физическая длина электрода затвора определялась путем  $C$ - $V$ -измерений [18]. Детальное моделирование процесса изготовления транзистора показало, что имеется перекрытие в 1–2 нм между электродом затвора и  $S$ - и  $D$ -расширениями для  $n$ -MOSFET и  $p$ -MOSFET, соответственно, так что  $L_{eff} = L_G - 2$  нм для  $n$ -FET и  $L_{eff} = L_G - 4$  нм для  $p$ -FET, где  $L_G$  – это физическая длина электрода затвора. Перечисленные длины были подтверждены путем тщательного анализа 2D электростатики [19, 20].

Транзистор с высокой подвижностью электронов НЕМТ это полевой транзистор, в котором для создания проводимости канала используется контакт двух полупроводниковых материалов III-V групп с различной шириной запрещенной зоны (вместо легированной области как у обычных MOSFET): широкая запрещенная зона работает как изолятор, а узкая запрещенная зона работает как канал проводимости. Моделировался III-V НЕМТ (рис. 4), подробно описанный в [16, 21]. Транзистор собран на подложке InP. Буфер наращивался на подложке, далее шли слои 2 нм  $In_{0.53}Ga_{0.47}As$ , 5 нм InAs и 3 нм  $In_{0.53}Ga_{0.47}As$ . Решетки слоя 2 нм  $In_{0.53}Ga_{0.47}As$  и подложки InP совместимы, однако, имеет место несоответствие между решетками  $In_{0.53}Ga_{0.47}As$  и InAs, так что слой InAs – псевдоморфный, напряжен, но он все же достаточно тонкий и не создает дефекты. Поверх этого 10 нм канала создан барьерный слой  $In_{0.52}Al_{0.48}As$ , играющий роль изолятора. Т-образная структура затвора понижает сопротивление затвора, что важно для радиочастотных приложений. Сильно допированные подконтактные  $SD$  слои (показаны на рисунке белым цветом) обеспечивают низкие терминальные сопротивления.

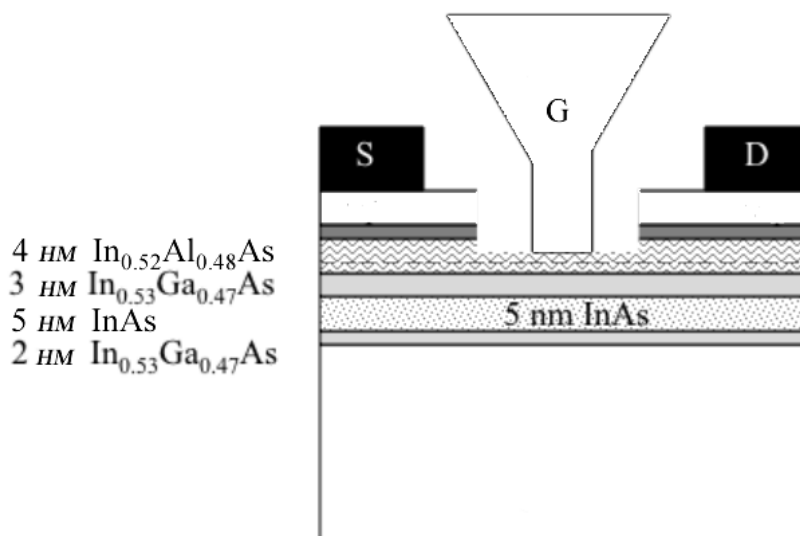


Рис. 4. III-V НЕМТ [21], взятый для анализа в рамках модели MVS/прохождения.

Измеренная подвижность длинноканального НЕМТ оказалась равной  $12500 \text{ см}^2/\text{В} \cdot \text{с}$ , что дает среднюю длину свободного пробега  $153 \text{ нм}$  [18]. Эффективная масса  $m^* = 0.022m_0$ , откуда  $v_T = 3.62 \times 10^7 \text{ см/с}$  [18]. Слой  $4 \text{ нм}$   $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$  поверх канала проводимости показывает емкость  $C_G^{inv} = 1.08 \text{ мФ} \cdot \text{см}^{-2}$  [18].

#### Вписывание модели MVS/прохождения в экспериментальные данные

Анализ экспериментальных характеристик обычно проводится как для выходных, так и для передаточных характеристик. Предполагается, что физические и эффективные длины затвора измерены независимо друг от друга, так же как и емкость затвора в режиме сильной инверсии. Параметр  $\alpha$ , который контролирует переход от слабой до сильной инверсии, взят равным 3.5 [17, 22]. Параметр  $\beta$  в  $F_{SAT}$  (36) подобран таким, чтобы вписаться в ток насыщения. Обычно он оказывается в узкой области  $\beta \approx 1.6 - 2.0$  [22]. Чтобы вписаться в экспериментальные данные, показанные на рис. 5, в качестве подгоночных использовались четыре параметра.

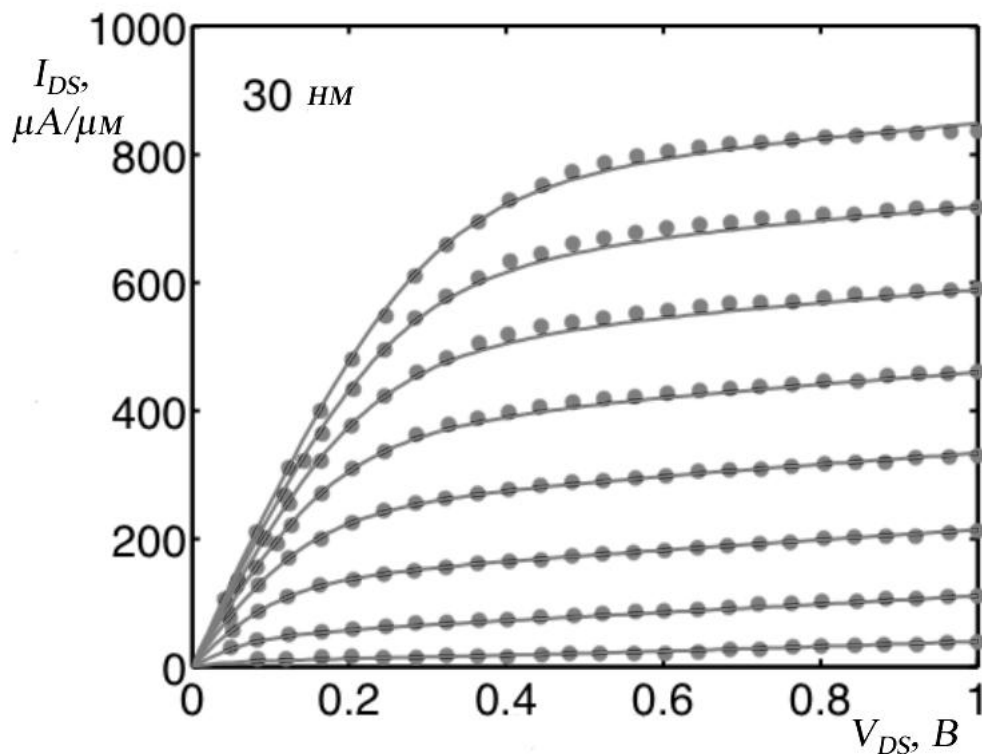


Рис. 5. Экспериментальные характеристики ETSOI MOSFET с  $L_{eff} = 30 \text{ нм}$  (точки) и вписанные в них уравнения модели MVS/прохождения (линии). Данные приведены для восьми значений  $V_{GS} = -0.2(0.1)0.5 \text{ В}$  [27].

Пороговое напряжение  $V_{T0}$  подбиралось таким образом, чтобы получить измеренный ток  $I_{OFF}$  при низком напряжении на стоке  $V_{DS}$ . Значение параметра  $\delta$  *DIBL* подбиралось таким, чтобы получить измеренное значение *DIBL*. Подпороговый параметр наклона  $m_0$  и параметр пробоя  $m'$  выбраны таким образом, чтобы вписаться в подпороговый наклон при низких и при высоких  $V_{DS}$ . Значение кажущейся подвижности  $\mu_{app}$  выбрано так, чтобы соответствовать наклону линейного участка характеристики  $I_{DS} \propto V_{DS}$ . Скорость впрыскивания  $v_{inj}$  соответствует измеренному току насыщения. Терминальное сопротивление  $R_{SD0}$  влияет как на линейную область, так и на область насыщения. Поскольку это влияние сказывается по разному на этих участках характеристик, то оказывается возможным независимо друг от друга определить  $\mu_{app}$  и  $R_{SD0}$ .

Результатом вписывания модели MVS/прохождения в экспериментально измеренные данные являются значения параметров  $R_{SD0}$ ,  $\mu_{app}$  и  $v_{inj}$ . Если транзистор собран правильно и качественно, то процедура вписывания проходит безупречно. В дополнение к трем названным параметрам удается получить скорость баллистического впрыскивания  $v_T$ , диффузионную подвижность  $\mu$ , среднюю длину свободного пробега  $\lambda_0$ , критическую длину  $\ell$ , а также коэффициенты прохождения  $T_{LIN}$  и  $T_{SAT}$ .

### *Результаты анализа экспериментальных данных ETSOI MOSFET и III-V HEMT в модели MVS/прохождения*

Согласование экспериментальных данных по ETSOI MOSFET с моделью MVS/прохождения показано на рис. 5. Результаты следующие:

$$\begin{aligned} R_{SD0} &= R_{S0} + R_{D0} = 130 \Omega \cdot \mu\text{м}, \\ \mu_{app} &= 220 \text{ см}^2/\text{В} \cdot \text{с}, \\ v_{inj} &= 0.82 \times 10^7 \text{ см/с}. \end{aligned}$$

Вычислим коэффициенты прохождения для линейной области и области насыщения. Для вычисления  $T_{LIN}$  по (48) нужно знать баллистическую подвижность. Находим ее согласно (45):

$$\mu_B = \frac{v_T L_{eff}}{2kT/q} = \frac{(1.12 \times 10^7)(30 \times 10^{-7})}{2 \times 0.026} = 658 \text{ см}^2/\text{В} \cdot \text{с}.$$

Тогда

$$T_{LIN} = \frac{\mu_{app}}{\mu_B} = \frac{220}{646} = 0.34.$$

Коэффициент прохождения в области насыщения оцениваем по (49):

$$T_{SAT} = \frac{2}{1 + v_T/v_{inj}} = \frac{2}{1 + 1.12/0.82} = 0.85.$$

Отношение тока  $I_{ON}$  к его баллистическому значению

$$B_{SAT} = \frac{I_{ON}}{I_{ON}^{ball}} = \frac{T_{SAT}}{2 - T_{SAT}} = 0.72.$$

Эти результаты типичны для Si MOSFET. Они показывают, что данный транзистор работает прилично ниже баллистического предела в линейной области, однако, весьма близко к баллистическому пределу в области насыщения.

Для III-V НЕМТ (рис. 6) подгоночные параметры следующие:

$$R_{SD0} = R_{S0} + R_{D0} = 400 \Omega \cdot \mu\text{м},$$

$$\mu_{app} = 1800 \text{ см}^2/\text{В} \cdot \text{с},$$

$$v_{inj} = 3.5 \times 10^7 \text{ см/с}.$$

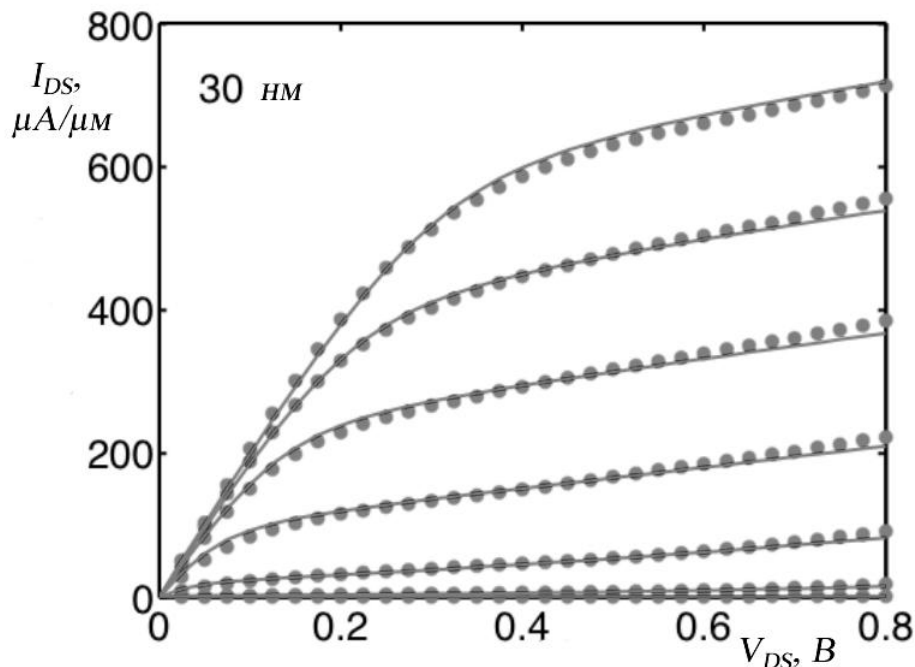


Рис. 6. Экспериментальные характеристики III-V НЕМТ с  $L_{eff} = 30 \text{ нм}$  [21] (точки) и вписанные в них уравнения модели MVS/прохождения (линии) [27].

Для баллистической подвижности, коэффициентов прохождения и значения  $V_{SAT}$  находим:

$$\begin{aligned}\mu_B &= 2088 \text{ см}^2/\text{В} \cdot \text{с}, \\ T_{LIN} &= 0.86, \\ T_{SAT} &= 0.98, \\ V_{SAT} &= 0.96.\end{aligned}$$

Эти результаты типичны для III-V НЕМТ. Они показывают, что этот транзистор в линейной области работает довольно близко к баллистическому пределу, а в области насыщения режим его работы существенно баллистический. Этот результат можно было бы предвидеть по двум причинам. Во-первых, средняя длина свободного пробега, найденная из диффузионной подвижности, оказалась равной  $153 \text{ нм}$ , в несколько раз большей, чем длина канала проводимости. Во-вторых, баллистическая подвижность оказалась меньше диффузионной подвижности.

Хотя этот транзистор и работает довольно близко к баллистическому пределу, если судить по току  $I_{ON}$ , однако, важно помнить, что «вблизи баллистического предела» означает только, что критический участок канала проводимости короткий по сравнению со средней длиной свободного пробега.

Обсуждаемые результаты помогают нам понять работу транзистора через такие показатели, как коэффициент прохождения и отношение к  $I_{ON}^{ball}$ . Как сейчас будет показано, более тщательный анализ линейной области и области насыщения позволяет дополнительно получить полезные параметры. Наконец, отметим, что в этих расчетах есть некоторые неопределенности. Выбор в расчетах значений эффективных электронных масс зависит от степени механической напряженности в кремниевой подложке, которая может увеличить или уменьшить  $m^*$ , и от степени непараболичности зоны проводимости, также влияющей на значение электронной массы, особенно в условиях квантовых ограничений. Высокие подзоны также могут быть заселены, так что невырожденная статистика может оказаться неприемлемой, особенно для III-V FET. Например, как будет показано далее, значение  $v_T$  можно получить непосредственно из анализа характеристик в области насыщения.

## Анализ линейной области

Анализ линейной области FET позволяет обнаружить наличие баллистической компоненты в сопротивлении канала проводимости, а также открывает возможность измерения диффузионной подвижности  $\mu$ . Процедура вписывания модели MVS/прохождения в экспериментальные данные позволяет извлечь физически существенное значение кажущейся подвижности. Из уравнений (43 – 45) находим, что

$$\frac{1}{\mu_{app}} = \frac{1}{\mu} + \left(\frac{\lambda_0}{\mu}\right) \frac{1}{L_{eff}}, \quad (51)$$

где второе слагаемое есть обратное значение баллистической подвижности. Ожидается, что зависимость  $\mu_{app}^{-1} \propto L_{eff}^{-1}$  должна быть прямой линией, пересечение которой с осью ординат даст обратное значение диффузионной подвижности, а угол наклона даст отношение средней длины свободного пробега к диффузионной подвижности. Если зависимость  $\mu_{app}^{-1} \propto L_{eff}^{-1}$  окажется прямой линией с физически приемлемым углом наклона, то зависимость кажущейся подвижности от длины канала должна быть обязана в первую очередь зависимости баллистической подвижности (45) от длины канала. Именно эта концепция подтверждается экспериментальными данными (рис. 7). Из этого графика находим, что  $\mu = 12195 \text{ см}^2/\text{В} \cdot \text{с}$ , а  $\lambda_0 = 171 \text{ нм}$ . Эти значения очень близки к значениям, ожидаемым для длинноканальных FET [18].

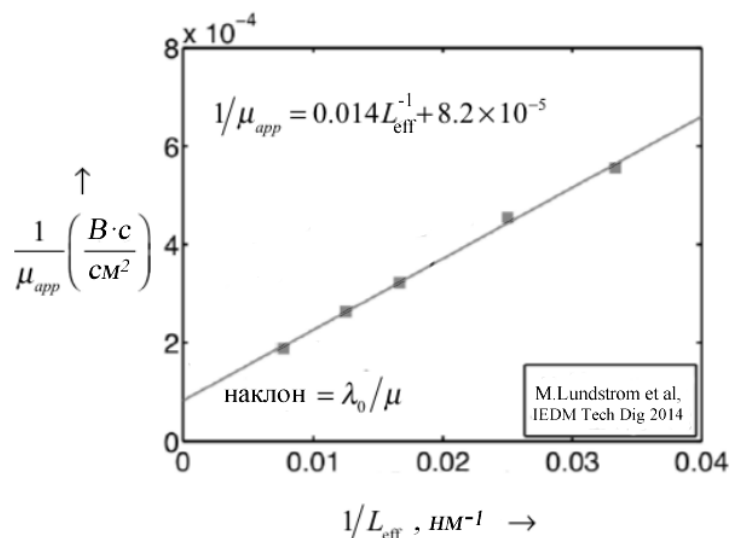


Рис. 7. Зависимость  $1/\mu_{app} \propto 1/L_{eff}$  для III-V НЕМТ: из пересечения с осью ординат находим диффузионную подвижность, а по углу наклона определяем квазиравновесную среднюю длину свободного пробега [18].



Зависимость  $1/\mu_{app} \propto 1/L_{eff}$  не будет прямой линией, если угол наклона, определяемый отношением  $\lambda_0/\mu$ , будет переменной величиной, т. е. когда средняя длина свободного пробега и/или диффузионная подвижность будут зависеть от длины канала. Например, для зависимости  $1/\lambda_0 \propto 1/L_{eff}$  из тех же уравнений (43) – (45) находим, что

$$\frac{1}{\lambda_0(L_{eff})} = \frac{v_T}{2(kT/q)} \frac{1}{\mu_{app}} - \frac{1}{L_{eff}}, \quad (52)$$

откуда может быть определена зависимость  $\lambda_0 \propto L_{eff}$ . Эта зависимость была построена для ETSOI MOSFET (рис. 8).

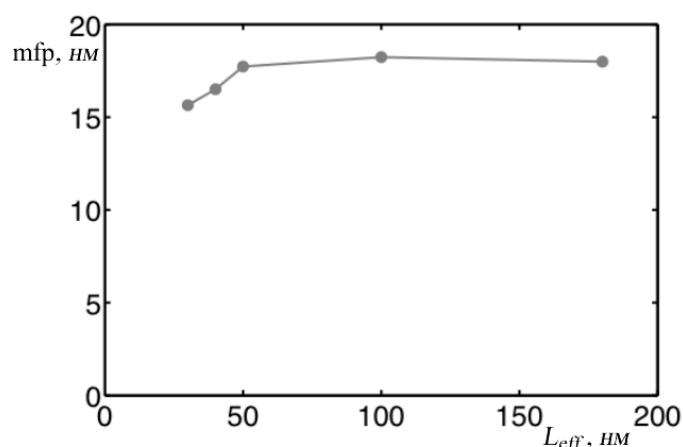


Рис. 8. Зависимость средней длины свободного пробега (mfp)  $\lambda_0$  от длины канала проводимости ETSOI MOSFET при низком  $V_{DS}$  [27]. Альтернативный подход, в котором предполагается  $\lambda_0 = \text{Const}$ , а прослеживается зависимость  $\mu \propto L_{eff}$  путем учета рассеяния назад, см. [28, 29].

Обращает на себя внимание уменьшение  $\lambda_0$  с укорочением длины канала. Этот эффект возникает, возможно, из-за технологических особенностей изготовления транзистора, однако, он также был и предсказан как следствие далекодействующих кулоновских осцилляций [30, 31].

Полезно отметить, что в ситуации, когда диффузионная подвижность не зависит от длины канала, как на рис. 7, тогда для определения обоих параметров – и диффузионной подвижности и средней длины свободного пробега не требуется знания величины  $v_T$ . Если же подвижность изменяется с длиной канала, то зависимость  $\lambda_0 \propto L_{eff}$  можно найти из (52), но потребуется знание тепловой скорости  $v_T$ . Это может вызвать затруднения без надежных

сведений об эффективных массах электронов и заселенности подзон. Однако, далее будет показано, что значение  $v_T$  можно получить из анализа зависимости скорости впрыскивания от длины канала проводимости.

### Анализ области насыщения

Скорость впрыскивания уменьшается с увеличением длины канала проводимости. Из (49) следует, что

$$v_{inj} = v_T \frac{\lambda_0}{\lambda_0 + 2\ell}, \quad (53)$$

что можно переписать как

$$\frac{1}{v_{inj}} = \frac{1}{v_T} + \frac{2\ell}{\lambda_0 v_T}. \quad (54)$$

Правдоподобно предположить, что критическая длина  $\ell$  пропорциональна длине канала  $L_{eff}$ . Доказать строго это предположение затруднительно, однако, тщательный анализ экспериментальных данных указывает на то, что это допущение вполне приемлемо на практике [17]. Предположив, что

$$\ell = \xi L_{eff}, \quad (55)$$

перепишем (54) в виде

$$\frac{1}{v_{inj}} = \frac{1}{v_T} + \frac{2\xi}{\lambda_0 v_T} L_{eff}. \quad (56)$$

Зависимость  $1/v_{inj} \propto L_{eff}$  ожидается линейной. Пересечение с осью ординат позволит определить значение  $v_T$ , а наклон даст значение коэффициента  $\xi$ , а стало быть и критической длины  $\ell$ , согласно (55). Результаты такого анализа для III-V НЕМТ [18] показаны на рис. 9.

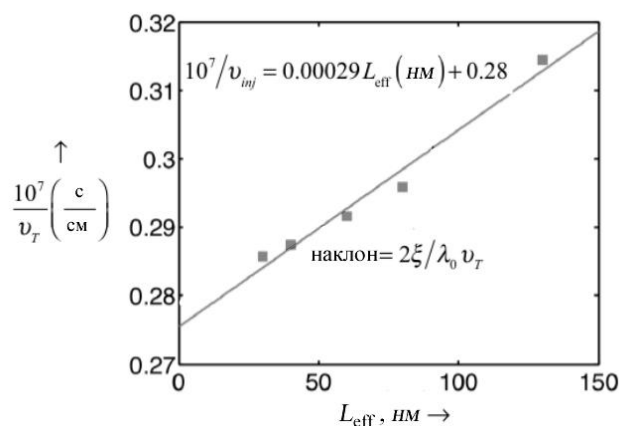


Рис. 9. Определение тепловой скорости  $v_T$  из анализа экспериментальных данных для III-V НЕМТ [18].

Из этого графика получаем  $v_T = 3.57 \times 10^7 \text{ см/с}$  и  $\xi = 0.09$ . Это значение тепловой скорости очень близко к значению, ожидаемому по известной эффективной массе, а критическая длина действительно есть лишь небольшая часть длины канала проводимости.

#### *Анализ перехода линейной области к насыщению*

Одна из проблем при моделировании наноразмерных MOSFET заключается в том, что нет аналитических выражений для зависимости  $T \propto V_{DS}$ . Уравнения (48) и (49) дают выражения для прохождения в пределе низких и высоких значений напряжений на стоке. Если бы мы имели аналитическую модель для  $T \propto V_{DS}$ , не было бы нужды в эмпирической функции насыщения тока стока  $F_{SAT}$  (36).

Из экспериментальных характеристик качественно собранных MOSFET можно извлечь «экспериментальные» зависимости  $T \propto V_{DS}$ . Процедура следующая. Сначала вписываем модель MVS/прохождения в экспериментальные характеристики. Далее генерируем собственные характеристики транзистора, положив  $R_{S0} = R_{D0} = 0$  в модели MVS и строим результирующие характеристики. Затем используем уравнение (36/гл.7) в невырожденном пределе, а именно:

$$I_{DS} = W |Q(V_{GS}, V_{DS})| v_T \left( \frac{T}{2-T} \right) \left[ \frac{1 - e^{-qV_{DS}/kT}}{1 + \frac{T}{2-T} e^{-qV_{DS}/kT}} \right]. \quad (57)$$

Инверсионный заряд  $Q(V_{GS}, V_{DS})$  берем по (32) с параметрами, которые уже были определены при вписывании модели MVS/прохождения в экспериментальные характеристики. Предполагая, что баллистическая скорость впрыскивания  $v_T$  известна, для любой точки  $(V_{Gi}, V_{Di})$  вписываем (57) в ранее вычисленные собственные характеристики транзистора и тем самым получаем численные значения  $T(V_{Gi}, V_{Di})$ . Полученный таким образом график зависимости  $T(V_{Gi}, V_{Di}) \propto V_{Di}$  при  $V_{Gi} = V_{DD}$  показан на рис. 10 для двух каналов разной длины. Как и ожидалось, коэффициент прохождения увеличивается с ростом напряжения на стоке и прохождение меньше для более длинного канала проводимости.

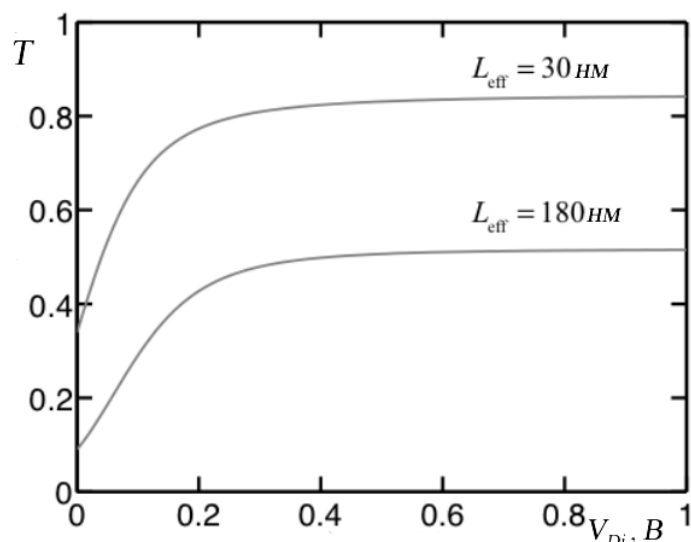


Рис. 10. Зависимость «экспериментального» коэффициента прохождения  $T$  от напряжения на стоке для ET-SOI MOSFET с каналами разной длины [27].

Из графика  $T(V_{Gi}, V_{Di}) \propto V_{Di}$  на рис. 10 можно определить критическую длину  $L_C(V_{Di})$ . Запишем коэффициент прохождения в виде

$$T(V_{Di}) = \frac{\lambda_0}{\lambda_0 + L_C(V_{Di})}. \quad (58)$$

Положив  $\underline{L_C}|_{V_{Di}=0} = L_{eff}$  и воспользовавшись графиками на рис. 10, получаются зависимости  $L_C(V_{Di})$ , показанные на рис. 11.

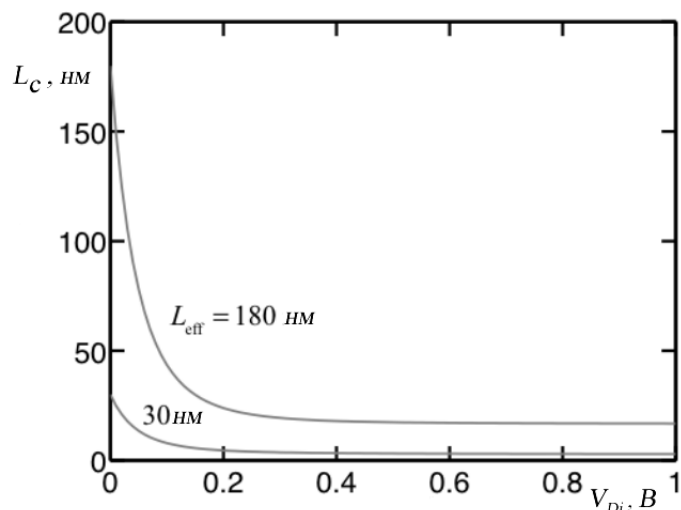


Рис. 11. Зависимость «экспериментальной» критической длины с учетом рассеяния назад от напряжения на стоке для ET-SOI MOSFET с каналами разной длины [27].

Как и ожидалось,  $L_C = \ell \ll L_{eff}$  по мере того, как  $V_{Di} \rightarrow V_{DD}$ . Рисунки 10 и 11 подтверждают ожидаемое поведение коэффициента прохождения и критической длины с изменением напряжения на стоке. Из значения  $T|_{V_{Di}=0}$  на рис. 10 и  $L_C|_{V_{Di}=0} = L_{eff}$ , согласно (58), находим следующие значения длины свободного пробега для каналов проводимости разной длины:

$$\begin{aligned} \lambda_0|_{L_{eff}=30\text{ нм}} &= 15.4 \text{ нм}, \\ \lambda_0|_{L_{eff}=180\text{ нм}} &= 17.8 \text{ нм}. \end{aligned}$$

Модель виртуального истока дает полуэмпирическое описание вольт-амперных характеристик полевых нанотранзисторов. Подбирая значения лишь нескольких параметров модели, удается с высокой точностью воспроизвести экспериментальные характеристики. Обобщенная транспортная модель Ландауэра – Датта – Лундстрёма позволяет параметрам модели придать физический смысл. Приведенные выше примеры иллюстрируют возможности модели MVS/прохождения извлечь из экспериментальных характеристик физически значимую информацию о транспортных процессах в нанотранзисторах. К такой информации относятся сведения о баллистической скорости впрыскивания  $v_T$ , средней длине свободного пробега с учетом рассеяния назад  $\lambda_0$ , диффузионной подвижности  $\mu$ , критической длине канала проводимости  $\ell$  и др. Позже мы обсудим ограничения и неопределенности, присущие этой модели.

Подведем итоги. Показано, что модель прохождения может быть тесно увязана с моделью виртуального истока. Простой заменой диффузионной подвижности  $\mu$  в модели виртуального истока на кажущуюся подвижность  $\mu_{app}$  для линейного тока получаем корректные результаты от баллистического до диффузионного предела. Заменой лимитируемой рассеянием скорости насыщения  $v_{sat}$  на скорость впрыскивания  $v_{inj}$  получаем корректную величину тока  $I_{ON}$ . Сравнение экспериментально измеренных характеристик показывает, что нанотранзисторы на кремниевой подложке работают весьма далеко от баллистического предела, тогда как наноразмерные III-V FET работают очень близко к баллистическому пределу.

К модели прохождения есть две серьезные претензии. Одна из них связана с трудностью вычисления зависимости  $I_{DS} \propto V_{DS}$  из-за проблем расчета  $T(V_{DS})$ .

А другая вызвана трудностями предсказания тока  $I_{ON}$ , что в свою очередь связано с трудностью вычисления критической длины  $\ell$  при высоком напряжении на стоке, в результате чего трудно предсказать величину  $T_{SAT}$ . Из-за этих ограничений модели прохождения и виртуального истока объединяют таким образом, что параметры модели прохождения берутся из вписывания экспериментальных результатов в модель виртуального истока, а физический смысл параметров берется из модели прохождения.

Показано каким образом можно анализировать характеристики нанотранзисторов на основе модели MVS/прохождения. Разнообразие типов транзисторов порождает новые частные задачи, однако, методология анализа экспериментальных данных принципиально не меняется. Подчеркивалось, что применение модели MVS/прохождения обосновано, если транзистор собран качественно. Для таких транзисторов модель позволяет надежно получить физически значимые параметры.

## Литература

1. P. Palestri, D. Esseni, S. Eminent, C. Fiegna, E. Sangiorgi, L. Selmi, *IEEE Trans. Electron Dev.*, **52**: 2727 – 2735 (2005).
2. P. Palestri, R. Clerc, D. Esseni, L. Lucci, L. Selmi, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 945 – 948 (2006).
3. R. Clerc, P. Palestri, L. Selmi, G. Ghibaudo, *J. Appl. Phys.*, **110**, 104502 (2011).
4. M. Lundstrom, *Fundamentals of Carrier Transport* (Cambridge, U.K.: Cambridge Univ. Press: 2000).
5. Y. Tsididis, C. McAndrew, *Operation and Modeling of the MOS Transistor* (New York: Oxford Univ. Press: 2011).
6. K. Y. Lim, X. Zhou, *Solid State Electron.*, **45**, 193 – 197 (2001).
7. M. J. Chen, H. T. Huang, K. C. Huang, P. N. Chen, C. S. Chang, C. H. Diaz, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 39 – 42 (2002).
8. V. Barral, T. Poiroux, M. Vinet, J. Widiez, B. Previtali, P. Grosgeorges, G. Le Carval, S. Barraud, J. L. Autran, D. Munteanu, S. Deleonibus, *Solid State Electron.*, **51**, 537 – 542 (2007).
9. M. Zilli, P. Palestri, D. Esseni, L. Selmi, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 105 – 108 (2007).
10. R. Wang, H. Liu, R. Huang, J. Zhuge, L. Zhang, D. W. Kim, X. Zhang, D. Park, Y. Wang, *IEEE Trans. Electron Dev.*, **55**: 2960 – 2967 (2008).

11. V. Barral, T. Poiroux, J. Saint-Martin, D. Munteanu, J. L. Autran, S. Deleonibus, *IEEE Trans. Electron Dev.*, **56**: 408 – 419 (2009).
12. V. Barral, T. Poiroux, D. Munteanu, J. L. Autran, S. Deleonibus, *IEEE Trans. Electron Dev.*, **56**: 420 – 430 (2009).
13. A. Khakifirooz, D. A. Antoniadis, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 667 – 670 (2006).
14. A. Khakifirooz, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **55**: 1391 – 1400 (2008).
15. A. Khakifirooz, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **55**: 1401 – 1408 (2008).
16. D. H. Kim, J. A. del Alamo, D. A. Antoniadis, B. Brar, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 861 – 864 (2009).
17. A. Majumdar, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **61**: 351 – 358 (2014).
18. S. Rakheja, M. Lundstrom, D. Antoniadis, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 35.1.1 – 35.1.4 (2014).
19. A. Majumdar, Z. Ren, S. J. Koester, W. Haensch, *IEEE Trans. Electron Dev.*, **56**: 2270 – 2276 (2009).
20. A. Majumdar, X. Wang, A. Kumar, J. R. Holt, D. Dobuzinsky, R. Venigalla, C. Ouyang, S. J. Koester, W. Haensch, *IEEE Electron Dev. Lett.*, **30**, 413 – 415 (2009).
21. D. H. Kim, J. A. del Alamo, *IEEE Electron Dev. Lett.*, **29**, 830 – 833 (2008).
22. A. Khakifirooz, O. M. Nayfeh, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **56**: 1674 – 1680 (2009).
23. Shaloo Rakheja, D. Antoniadis, *MVS 1.0.1 Nanotransistor Model (Silicon)*: [www.nanohub.org/resources/19684](http://www.nanohub.org/resources/19684).
24. Shaloo Rakheja, D. Antoniadis, *MVS Nanotransistor Model (Silicon) 1.1.1*: [www.nanohub.org/publications/15/4](http://www.nanohub.org/publications/15/4).
25. Y. Taur, T. Ning, *Fundamentals of Modern VLSI Devices* (New York: Oxford Univ. Press: 2013).
26. M. S. Lundstrom, *IEEE Electron Dev. Lett.*, **18**, 361 – 363 (1997).
27. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
28. D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **63**: 2650 – 2656 (2016).
29. K. Natori, H. Iwai, K. Kakushima, *J. Appl. Phys.*, **118**, 234502 (2015).
30. M. V. Fischetti, S. E. Laux, *J. Appl. Phys.*, **89**, 1205 – 1231 (2001).
31. T. Uechi, T. Fukui, N. Sano, *Phys. Status Solidi C*, **5**, 102 – 106 (2008).

## Глава 9. Фундаментальные пределы и ограничения

### 9.1. Введение

По мере того, как размеры нанотранзисторов становятся все меньше и меньше, все более актуальным представляется вопрос «Каковы фундаментальные физические пределы MOSFET?». Можно ли пользоваться классическим описанием транспорта электронов вблизи физических пределов? Даже если классические модели транспорта приемлемы, насколько обоснованы допущения, принятые в модели виртуального истока/прохождения, как лучшей из известных ныне моделей для анализа нанотранзисторов?

### 9.2. Фундаментальные пределы

Уже простейшая барьерная модель MOSFET (рис. 1) позволяет наметить некоторые фундаментальные пределы работы нанотранзисторов в роли переключателей тока. Аналогичная модель применима и к биполярному транзистору. Излагаемый далее эвристический подход приводит к тем же выводам, что и результаты, полученные ранее в работе [1].

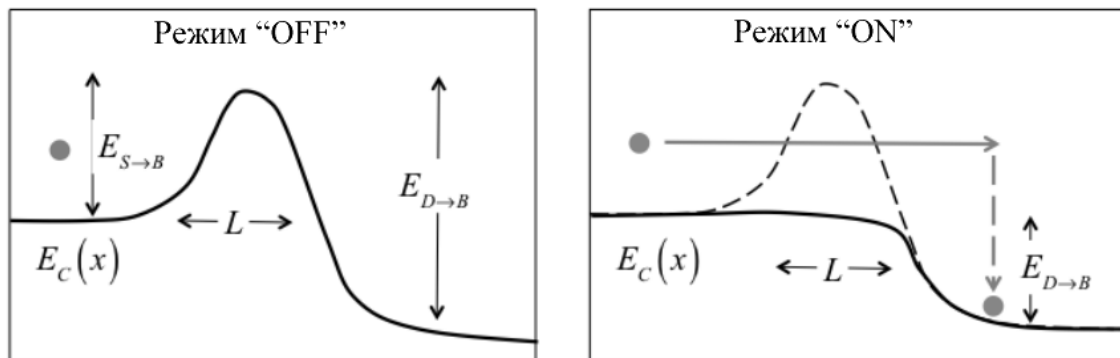


Рис. 1. Простая модель MOSFET как переключателя тока в режимах «OFF» и «ON». Показаны энергетические барьеры для электронов, движущихся от истока к барьеру  $E_{S \rightarrow B}$  и от стока к барьеру  $E_{D \rightarrow B}$ .

На рисунке слева показана ситуация в режиме «OFF». Высокий энергетический барьер не позволяет электронам из истока попасть в сток (предполагается высокое напряжение на стоке). Справа показана ситуация в режиме «ON». Большое напряжение, поданное на затвор, практически зануляет барьер и электроны из истока по каналу проводимости поступают в сток. Предполагаем, что в канале имеет место баллистический транспорт, так что



электроны отдают свою энергию стоку и релаксируют в результате интенсивных неупругих взаимодействий в терминале стока.

Как показано на рис. 2, эта простая модель позволяет установить минимальную энергию, необходимую для переключения из одного режима в другой. Большое напряжение на затворе в режиме «ON» нивелирует барьер между истоком и каналом, но барьер  $E_{D \rightarrow B}$  между стоком и вершиной барьера остается, поскольку на сток подано положительное напряжение. После того как электроны термически релаксировали в стоке, диссипировали свою кинетическую энергию в результате рассеяния, остается некоторая вероятность  $P$  того, что электроны все же преодолеют барьер  $E_{D \rightarrow B}$  и вернуться в исток: в этом случае переключения не произойдет. Требуя, чтобы эта вероятность была меньше  $1/2$ , имеем

$$P = e^{-E_{D \rightarrow B}/kT} < \frac{1}{2}, \quad (1)$$

откуда находим минимальную энергию переключения

$$E_{min} \equiv \underline{E_S}|_{min} = kT \ln 2, \quad (2)$$

что составляет  $0.017 \text{ эВ}$  при комнатной температуре. Эти простые соображения носят эвристический характер. Фундаментальные рассуждения Ландауэра [2 – 4] и дальнейший тщательный анализ [5] приводят к тем же самым результатам о величине минимальной энергии переключения.

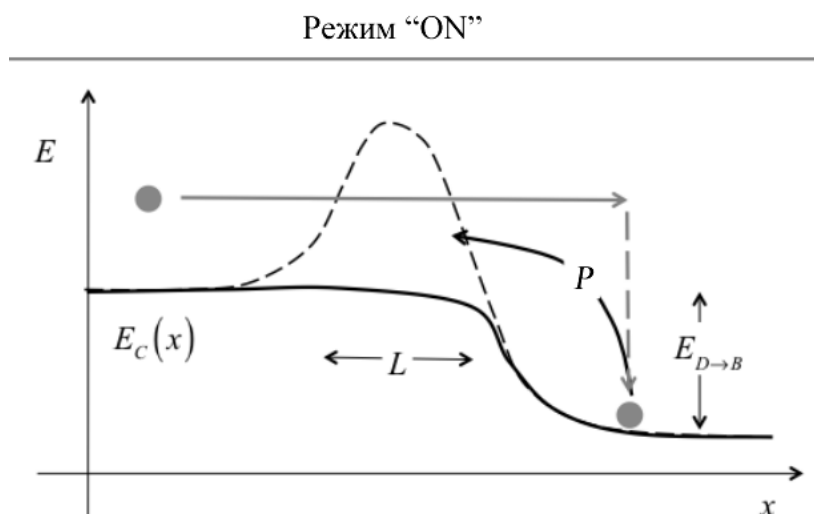


Рис. 2. К иллюстрации события переключения из одного режима в другой. Вероятность переключения есть  $1 - P$ , где  $P$  есть вероятность термоионной реэмиссии электрона из стока в исток.

Фундаментальный закон Ландауэра об энергии, необходимой для стирания одного бита информации  $E_{\text{bit}} \geq kT \ln 2$  [2, 6] был открыт в 1961 году, а его первое экспериментальное подтверждение поступило в 2012 году [7, 8]. Вместе с тем, в виду грядущей эры квантовых вычислений и квантовых компьютеров уместно было поставить вопрос не отличаются ли фундаментально между собой классические и квантовые биты. Согласно новейшим экспериментам [9, 10], ответ отрицательный. Экспериментальное подтверждение фундаментальности принципа Ландауэра получено в условиях полностью квантованной системы, когда квантуются не только биты, но и тепловой резервуар, с которым биты обмениваются энергией.

Вернемся к нашей задаче и оценим возможную минимальную длину канала проводимости MOSFET. Как показано на рис. 3, когда транзистор находится в режиме «OFF», барьер достаточно высок и достаточно широк, чтобы не пропустить электроны от истока к стоку. Для оценки нужной высоты барьера обратимся к явлению термоионной эмиссии. Высота барьера в режиме «OFF» должна быть по крайней мере не меньше  $E_{\text{min}}$ , что служит гарантией того, что электроны, если и преодолели барьер, то с вероятностью, меньшей 1/2. Минимальная ширина барьера (длина канала) определяется квантовомеханическим туннелированием через барьер. Вероятность того, что электрон из истока туннелирует через барьер стандартно оценивается в приближении ВКБ [1].

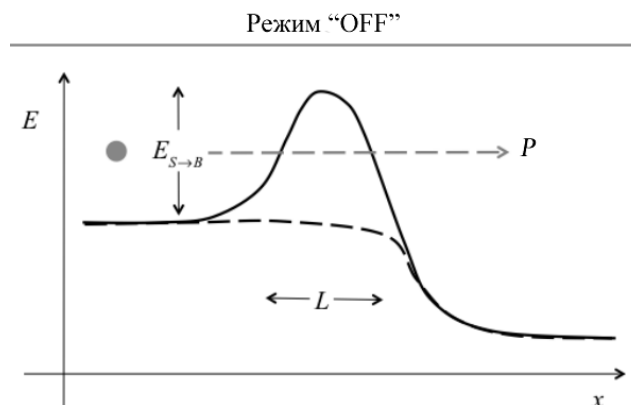


Рис. 3. К иллюстрации того, что в режиме «OFF»  $P$  есть вероятность квантовомеханического просачивания электрона через барьер ширины  $L$ .

Из требования, чтобы в режиме «OFF» вероятность туннелирования была меньше 1/2, следует, что

$$P = e^{-2\sqrt{2m^*E_{S \to B}}L/\hbar} < \frac{1}{2}, \quad (3)$$

откуда следует, что

$$L > \frac{\hbar}{\sqrt{2m^* E_{S \rightarrow B}}} . \quad (4)$$

Подставляя  $E_{S \rightarrow B} = \underline{E_S}|_{min}$ , находим, что минимальная длина канала проводимости

$$L_{min} = \frac{\hbar}{\sqrt{2m^* \underline{E_S}|_{min}}} . \quad (5)$$

Можем также оценить время переключения транзистора из одного режима в другой. В режиме «ON» электроны просто движутся по каналу с баллистической скоростью. Минимальное время пролета через канал есть

$$\tau_{min} = \frac{L_{min}}{v_T} . \quad (6)$$

Подставляя  $L_{min}$  по (5), тепловую скорость  $v_T$  по (3/гл.6) и пренебрегая множителем порядка единицы, находим, что

$$\tau_{min} = \frac{\hbar}{\underline{E_S}|_{min}} . \quad (7)$$

Оценив все три предельных показателя при комнатной температуре  $T = 300 \text{ K}$  и положив  $m^* = m_0$ , получим:

$$\begin{aligned} \underline{E_S}|_{min} &= kT \ln 2 = 0.017 \text{ эВ}, \\ L_{min} &= \frac{\hbar}{\sqrt{2m^* \underline{E_S}|_{min}}} = 1.5 \text{ нм}, \\ \tau_{min} &= \frac{\hbar}{\underline{E_S}|_{min}} = 40 \text{ фс}. \end{aligned} \quad (8)$$

Эти же значения получаются [1] из принципа неопределенности Гейзенберга  $\Delta x \Delta p_x \geq \hbar/2$ ,  $\Delta E \Delta t \geq \hbar/2$ . Дополним также перечень фундаментальных пределов (8) еще двумя ограничениями [1]: максимальная плотность упаковки транзисторов  $n_{max}$  и диссипация мощности  $\wp$  на единицу поверхности в этой предельной Si технологии, а именно:

$$\begin{aligned} n_{max} &= \frac{1}{L_{min}^2} = 4.7 \times 10^{13} \text{ см}^{-2}, \\ \wp &= \frac{n_{max}}{\tau_{min}} \underline{E_S}|_{min} = 3.7 \times 10^6 \text{ Вт} \cdot \text{см}^{-2}. \end{aligned} \quad (9)$$

Минимальная энергия переключения одиночного транзистора по оценке (8) намного ниже энергии переключения типичной интегральной схемы CMOS. Последняя может быть оценена из  $E_s = C_s V_{DD}$ , где  $C_s$  – это среднее значение переключаемой емкости. Для типичной ситуации  $C_s \approx 1 \phi\Phi$  и  $V_{DD} \approx 1 B$ , что приводит к оценке на 4 – 5 порядков больше фундаментального предела. Это сильное расхождение связано с тем, что типичная емкость интегральной схемы намного больше собственной емкости затвора одиночного транзистора. Увеличение емкости реальных схем связано с различными паразитными эффектами и не только и уменьшить емкость в нынешней технологической генерации практически не представляется возможным.

Предельные значения длины канала проводимости и диссипации мощности были оценены из предположения, что отношение «ON»/«OFF» равно двойке. В реальных интегральных схемах это отношение достигает  $10^4$  (см. ниже рис. 5), а это значит, что длина канала и энергия переключения всегда будут значительно больше фундаментальных пределов. И тем не менее, длина канала проводимости на сегодняшний день в транзисторах составляет около 20 нм и даже меньше, т. е. того же порядка величины, что и предельное значение в 1.5 нм. Реальное время переключения транзистора сегодня также того же порядка величины, что и предельное фундаментальное значение. Однако, особенности нынешних интегральных схем таковы, что характерные для них времена переключения всегда будут на порядки выше, чем для отдельного транзистора.

В связи с проведенными выше рассуждениями невольно возникает вопрос «Не существует ли иной переключатель тока, фундаментально лучший чем MOSFET?». Однако, тот факт, что предельные ограничения  $E_s|_{min}$  и  $\tau_{min}$  следуют из фундаментального принципа неопределенности Гейзенберга, говорит о том, что нельзя ожидать появления более эффективного бинарного переключателя тока.

### **9.3. Квантовый транспорт в MOSFET с длиной канала меньше 10 нм**

Практические пределы уменьшения длины канала проводимости транзисторов могут быть определены путем математического моделирования. Результаты моделирования квантового транспорта в Si NW MOSFET [11, 12] представлены ниже. Электростатический контроль в таком транзисторе превосходный, поскольку речь идет о MOSFET [11], в котором роль 1D канала проводимости выполняет углеродная нанотрубка (Carbon NanoTube/CNT), так

что пределы масштабирования определяются квантовомеханическим туннелированием электронов от истока через барьер в режиме OFF. Рисунки 4 демонстрируют ток в каналах CNT с длиной от 12 до 5 нм при большом напряжении на стоке в режиме OFF.

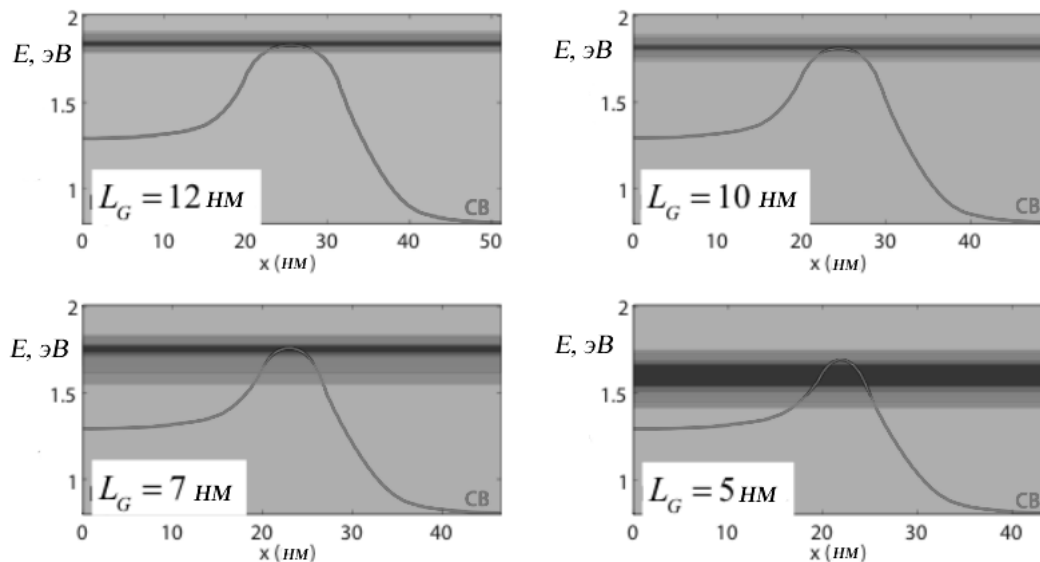


Рис. 4. Ток в каналах CNT разной длины при большом напряжении на стоке в режиме OFF, вычисленный квантовомеханическим моделированием [11 – 13].

При длине канала  $L_G = 12 \text{ нм}$  ток протечки  $I_{OFF}$  идет практически полностью выше барьера. Такой транзистор работает в привычном классическом режиме, контролируемом барьером. Когда длина канала уменьшается до 10 нм, небольшая доля электронов уже туннелирует через барьер: при  $L_G = 10 \text{ нм}$  все еще в основном сохраняется привычный режим, контролируемый барьером. При  $L_G = 7 \text{ нм}$  уже существенная доля тока  $I_{OFF}$  туннелирует через барьер. Функциональность транзистора нарушается: подпороговый разброс  $SS$  уже значительно превышает фундаментальный предел в 60 мВ/декада (рис. 5).

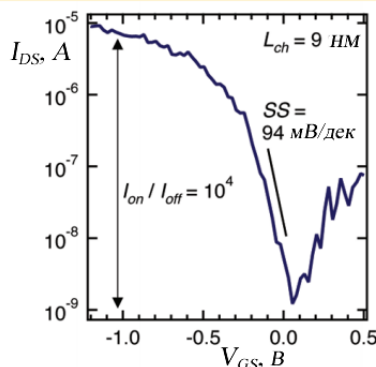


Рис. 5. Подпороговый разброс в Si NW MOSFET с  $L_G = 9 \text{ нм}$  [11].

Наконец, при  $L_G = 5 \text{ нм}$  уже бóльшая часть тока  $I_{OFF}$  обязана туннелированию через барьер. При такой длине канала проводимости невозможно управлять током контролируя высоту барьера, поскольку барьер стал проницаемым для электронов.

Представленные выше результаты моделирования свидетельствуют о том, что классическая модель прохождения MOSFET, подробно изложенная ранее, вполне оправдана применительно к Si транзисторам с длиной канала проводимости до 10 нм и даже еще несколько меньше. Дальнейшее укорочение канала с одновременным подавлением туннелирования требует бóльших значений эффективной массы электронов [14]. Масштабирование до 5 нм уже ставит серьезные проблемы как прикладного характера (возрастающая роль паразитных сопротивлений и емкостей при очень коротких каналах проводимости), так и фундаментального характера – туннелирование через барьер [15]. Численное моделирование транзисторов с напряженными подложками и специальным выбором их ориентации относительно направления канала проводимости показывает, что, по-видимому, окажется возможным реализовать приемлемый режим функционирования MOSFET с длиной канала проводимости даже меньше 5 нм [15]. Однако, нужно все-таки констатировать, что прикладные и фундаментальные пределы скейлинга MOSFET уже достигнуты.

#### **9.4. Упрощающие допущения в модели прохождения**

Опираясь на обобщенную транспортную модель Ландауэра – Датта – Лундстрема, в предыдущих главах подробно изложена модель прохождения для описания вольт-амперных характеристик наноразмерных транзисторов. Первоначальный подход Ландауэра, лежащий в основе обобщенной транспортной модели ЛДЛ, при некоторых условиях может быть непосредственно сформулирован исходя из квантовомеханического рассмотрения диссипативного квантового транспорта [16]. Модель Ландауэра также следует непосредственно из транспортного уравнения Больцмана при некоторых упрощающих допущениях [17, 18]. Как было обосновано в предыдущем разделе, транзисторы с длиной канала проводимости выше 10 нм вполне могут быть описаны классически. Соответственно этому, мы далее уделим больше внимания тем допущениям, которые лежат в основе полуклассического подхода Ландауэра, обстоятельно изложенного в [19 – 22].

Во-первых, обратим внимание на то, что модель прохождения в соответствии с подходом Ландауэра выражает терминальный ток в виде линейной комбинации фермиевских функций контактов. Этот результат следует из транспортного уравнения Больцмана (ТУБ), только если в матрице плотности отбросить нелинейные члены, обязанные принципу запрета Паули. Оправдать неучет таких членов можно в двух случаях: 1) когда имеет место упругое рассеяние и 2) в ситуации, когда для носителей тока справедлива невырожденная статистика. Мы, как правило, считали, что реализуется именно такая ситуация, однако, для описания тока в режиме «ON» может потребоваться статистика Ферми – Дирака. В этом случае оправдать подход Ландауэра при наличии лишь упругого рассеяния можно только весьма приближенно. Если ни одно из этих двух условий не выполняется, то уравнение Ландауэра для тока не следует из ТУБ: оно все еще может работать, но каждый случай применения уравнения Ландауэра потребует тщательного анализа конкретной ситуации.

Во-вторых, далее обратим внимание на то, что модель Ландауэра предполагает идеальные контакты. Это означает, что электроны, поступающие из канала проводимости в контакты, полностью ими поглощаются: отражения электронов назад в канал не происходит. Как только электроны попали в контакт, они подвергаются интенсивному рассеянию и немедленно приходят в равновесное состояние. Более того, контакты считаются бесконечно емкими источниками электронов, имея в виду под этим, что контакты могут обеспечить любой ток, не испытывая при этом истощения. Реальные контакты на практике далеки от идеала.

В-третьих, нужно учесть ограничения, связанные с понятием прохождения: независимость средней длины свободного пробега от напряжения, или иначе ее квази-равновесность, а также зависимость критической длины канала от напряжения. Нет уверенности в том, что эти ограничения в полной мере приемлемы для нанотранзисторов, рассеяние в которых может носить весьма сложный характер.

В-четвертых, нужно стремиться к самосогласованному учету электростатики. Мы не рассматривали пространственное распределение потенциала по всему каналу проводимости, а ограничились лишь вершиной барьера и ввели параметр *DIBL* для учета эффектов 2D электростатики. Такой подход нуждается в обосновании и дальнейшем развитии.

В-пятых, уже упоминалось о роли статистики Ферми – Дирака, хотя мы пользовались в основном невырожденной статистикой Максвелла – Больцмана.

Статистика Ферми – Дирака усложняет применение модели прохождения, однако, она может быть необходима для таких нанотранзисторов как III-V FET [23, 24].

Наконец, мы предполагали, что инверсионный заряд контролируется только электростатикой и транспорт на него не влияет. Это несомненно в общем случае не так, и может оказаться важным для III-V FET [23, 24].

Можно обратить внимание еще на следующее. Мы всегда предполагали простую изотропную зонную структуру, однако, учет непараболичности зоны проводимости может оказаться важным так же, как и множественность долин зоны проводимости кремния и искривление валентных зон.

Далее мы более подробно рассмотрим перечисленные допущения и ограничения, и начнем с вывода формулы Ландауэра для тока из транспортного уравнения Больцмана.

### Вывод формулы Ландауэра для тока

Наша конечная цель не столько получить уравнение Ландауэра для тока (2/гл.5), сколько понять допущения, лежащие в основе вывода этого важнейшего уравнения в теории электрических явлений. Рассмотрим полупроводниковый проводник (рис. 6) с двумя массивными контактами 1 и 2 в равновесном состоянии, через которые инжектируются потоки электронов  $F_1(E)$  и  $F_2(E)$ . Левый контакт впрыскивает поток электронов  $F_1(E)$ , доля которого  $T(E)F_1(E)$  переходит в правый контакт, куда возвращается также часть  $(1-T(E))F_2(E)$  потока электронов  $F_2(E)$ , впрыснутых правым контактом и рассеянным назад. Ситуация с потоком  $F_2(E)$  полностью симметрична. Рассеяние предполагается упругим, так что коэффициенты прохождения слева направо и справа налево совпадают.

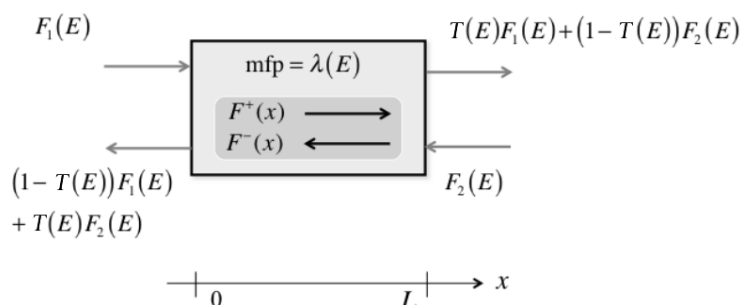


Рис. 6. Полупроводниковый проводник с идеальными контактами, находящимися в равновесном состоянии. Входящие потоки электронов  $F^+(x)$  и  $F^-(x)$  направлены как в сторону  $+x$ , так и  $-x$ .



Внутри проводника реализуются положительно направленные потоки  $F^+(x)$  и отрицательно направленные  $F^-(x)$ . Положительно направленный поток истощается в результате рассеяния назад в отрицательно направленный поток и возрастает за счет рассеяния назад отрицательно направленного потока. Отрицательно направленный поток ведет себя аналогичным образом. Соответственно этому можно написать, что

$$\begin{aligned}\frac{dF^+(x)}{dx} &= -\frac{F^+}{\lambda} + \frac{F^-}{\lambda}, \\ \frac{dF^-(x)}{dx} &= -\frac{F^-}{\lambda} + \frac{F^+}{\lambda},\end{aligned}\tag{10}$$

где предполагается, что оба потока происходят в канале с одной и той же энергией (упругое рассеяние). Знаки в обоих уравнениях одинаковые, поскольку поток  $F^-$  считается положительным, если движется в сторону  $-x$ . Уравнения (10) это простые транспортные уравнения Больцмана в стационарном состоянии динамического равновесия, в которых пространство скоростей включает только два противоположные друг другу направления. Величина  $\lambda$  в (10) это средняя длина свободного рассеяния назад. Величина  $dx/\lambda$  есть вероятность на единицу длины того, что положительный (отрицательный) поток при рассеивании назад превратится в отрицательный (положительный) поток. Как подробно показано в [6, 20], решение системы уравнений (10) при заданных граничных условиях непосредственно приводит к выражению для коэффициента прохождения (21/гл. 5), а именно:

$$T(E) = \frac{\lambda(E)}{\lambda(E) + L}.\tag{11}$$

До сих пор рассматривалась ситуация, когда проводник не подвержен действию внешнего электрического поля. Прохождение еще можно вычислить, если к проводнику приложено лишь небольшое напряжение [13], но если электрическое поле велико, то решение такой транспортной задачи сильно усложняется, поскольку вся система вышла из равновесия и предположение о независимости друг от друга мод проводимости нарушилось: коэффициенты прохождения слева направо и справа налево теперь различные, прохождение в одном направлении теперь стремится к единице, а в противоположном – к нулю.

Вернемся к рис. 6. Суммарный поток на правом контакте при  $x = L$  есть

$$F(E) = TF_1(E) + (1-T)F_2(E) - F_2(E) = T(F_1(E) - F_2(E)) \quad (12)$$

и такой же поток имеет место на левом контакте при  $x = 0$ .

Ток на левом контакте в интервале энергий от  $E$  до  $E + dE$  есть

$$I_1(E)dE = qF_1(E)dE = qv_x^+ \frac{D(E)}{2} f_1(E)dE, \quad (13)$$

где  $v_x^+$  есть скорость электронов в направлении  $+x$ ,  $D(E)$  есть плотность состояний, а множитель  $1/2$  связан с тем, что скорости электронов только у половины состояний направлены в сторону  $+x$ ,  $f_1(E)$  есть функция Ферми левого контакта. Аналогично записывается ток, инжектируемый правым контактом в предположении, что  $|v_x^-(E)| = |v_x^+(E)|$ , а именно:

$$I_2(E)dE = qv_x^+ \frac{D(E)}{2} f_2(E)dE. \quad (14)$$

Остается ввести безразмерное число мод проводимости при энергии  $E$ :

$$M(E) \equiv \frac{h}{4} v_x^+ D(E). \quad (15)$$

Теперь суммарный ток при энергии  $E$

$$I(E) = I_1(E) - I_2(E) = \frac{2q}{h} T(E) M(E) (f_1(E) - f_2(E)). \quad (16)$$

Полный ток получается интегрирование по всему спектру энергий:

$$I = \int I(E) dE. \quad (17)$$

Окончательно получаем искомое уравнение Ландауэра для тока (2/гл. 5):

$$I = \frac{2q}{h} \int_{-\infty}^{+\infty} T(E) M(E) (f_1(E) - f_2(E)) dE. \quad (18)$$

Приведенные простые рассуждения достаточны, чтобы продемонстрировать происхождение уравнения Ландауэра для тока. Более глубокое обоснование этого уравнения можно найти в [22].

### *Неидеальные контакты*

Контакты ограничивают работоспособность и функциональные возможности транзисторов. Дело не только в терминальных сопротивлениях, о которых всегда нужно помнить, сказаться могут и другие эффекты. Эти эффекты характерны не только для Si MOSFET, но могут создать проблемы для III-V HEMT и для GaN FET [23 – 25].

Ток на вершине барьера можно записать в виде

$$I_D = qn_s(0)\langle v_x(0) \rangle.$$

Величина зарядовой плотности на вершине барьера контролируется электростатикой затвора, однако, если исток недостаточно допирован, он может не обеспечить нужный заряд на вершине барьера. Исток истощается, возникает сильное электрическое поле и функционирование транзистора нарушается. Этот эффект назвали истощением истока (source exhaustion) [26]. Может случиться и другое явление. Канал проводимости обычно уже истока, и электронам может оказаться затруднительным попасть из истока в канал. Этот эффект назвали голоданием истока (source starvation) и он может быть важным для III-V FET [27]. Парадоксально, но это тот случай, когда рассеяние может улучшить функционирование транзистора. Квантовомеханическое моделирование MOSFET с терминальными контактами показало, что функционирование транзисторов при наличии рассеяния оказывается более эффективным по сравнению с баллистическим пробросом, поскольку рассеяние помогает «слить» электроны в канал [28].

Эффекты неидеальности истока моделировались с учетом зависимости терминальных сопротивлений от напряжения на затворе. Такое моделирование можно провести эмпирически [23], а можно более обосновано физически, вставив беззатворный FET в областях истока и стока, граничащих с каналом проводимости [25]. Мы далее еще обсудим ситуацию, когда электроны рассеиваются назад непосредственно от истока и стока. Этот эффект, наблюдавшийся в некоторых транзисторах, может быть также отнесен к контактными эффектам [29].

### *Критическая длина при рассеянии назад*

Вычисление прохождения в бесполево́м проводнике сравнительно простая задача, и как видно из (11), результат очевиден. Однако, в канале MOSFET может возникнуть сильное и быстро меняющееся в пространстве электрическое

поле. В этом случае расчет прохождения нуждается в более тщательном анализе так называемых нелокальных транспортных эффектов, например, таких как всплески скорости (velocity overshoots) [30] и результат может быть записан в виде

$$T(E) = \frac{\lambda_0(E)}{\lambda_0(E) + L_c}, \quad (19)$$

где  $\lambda_0(E)$  – это квазиравновесная средняя длина свободного пробега, а  $L_c \ll L$  – это критическая длина рассеяния назад.

Как уже обсуждалось ранее, подобное уравнение физически вполне приемлемо, а именно, коэффициенты прохождения, извлекаемые из вписывания экспериментальных характеристик в модель MVS/прохождения, ведут себя в соответствии с (19):  $L_c \rightarrow L$  при низком напряжении на стоке и  $L_c \rightarrow \ell \ll L$  при высоком напряжении. Уравнение (19) можно вывести в предположении квазиравновесного транспорта (низкое напряжение на стоке), однако, в режиме высокого напряжения на стоке транспорт в значительной части канала проводимости далек от равновесного, в результате чего использование в (19) квазиравновесной средней длины свободного пробега, казалось бы, не может быть оправдано, поскольку рассеяние, вызывающее возврат электронов в исток, имеет место очень близко к вершине барьера, еще до того как электроны были достаточно разогреты полем стока. Кажется удивительным, что это простое уравнение способно описать столь сложную ситуацию, однако, моделирование такой транспортной задачи вдали от равновесия методом Монте-Карло свидетельствует о том, что уравнение (19) фактически вполне пригодно [31].

Ток  $I_{ON}$  в MOSFET пропорционален скорости впрыскивания (24/гл. 8),

$$v_{inj} = \frac{T_{SAT}}{2 - T_{SAT}} v_T = \frac{\lambda_0 v_T}{\lambda_0 + 2\ell},$$

которая определяется коэффициентом прохождения в области насыщения или, что эквивалентно, критической длиной  $\ell$  при высоком напряжении на стоке. Вписывание экспериментальных характеристик в модель MVS/прохождения позволяет вычислить скорость впрыскивания и критическую длину [32]. Для предсказания же тока  $I_{ON}$  нужно иметь значение критической длины  $\ell$ .

Длина  $\ell$  приблизительно есть то расстояние, на котором потенциал увеличивается на величину  $kT/q$  по сравнению с его значением на вершине

барьера [31], но это лишь грубая оценка. В предположении невырожденной статистики и квазиравновесных условий можно получить выражение для  $\ell$  через потенциал в канале  $V(x)$  [33]. Можно вывести также аналитическое выражение, которое не подразумевает квазиравновесности [34].

Известны результаты моделирования рассеяния назад в наноразмерных MOSFET методом Монте-Карло с учетом нелокальных транспортных эффектов и самосогласования с уравнением Пуассона [35 – 38]. Результаты из [33] подтверждают, что именно рассеяние вблизи вершины барьера возвращает электроны назад в исток, однако, критическая длина оказалась несколько больше расстояния, на котором потенциал увеличивается на  $kT/q$ . Критическая длина зависит также от профиля потенциала, зависящего в свою очередь от самосогласованности электростатики, так что, например, баллистическое моделирование профиля потенциала не способно предсказать критическую длину. В работе [35] пришли к выводу, что при высоком напряжении на стоке  $\ell \ll L$ , однако, достаточно точное вычисление величины  $\ell$  требует при моделировании учета самосогласования и правдоподобного учета рассеяния.

*Зависимость средней длины свободного пробега и подвижности от длины канала проводимости*

Модель прохождения MOSFET базируется на понятии баллистической скорости впрыскивания, которая зависит от зонной структуры и от средней длины свободного пробега назад  $\lambda$ , которая в свою очередь зависит от зонной структуры и от физики рассеяния, а также от распределения электронов в импульсном пространстве. При большом напряжении на стоке скорость рассеяния электронов и величина  $\lambda$  сильно меняются вдоль канала проводимости по мере того как электроны, приближаясь к стоку, приобретают все большую энергию. Ключевое предположение в этой модели заключается в том, что подходящей средней длиной свободного пробега при вычислении прохождения является ее квазиравновесное значение ( $\lambda \approx \lambda_0$ ), поскольку рассеяние, контролирующее прохождение, проявляется очень близко к истоку прежде чем электроны получат шанс приобрести значительную энергию.

Ранее, объединяя модель прохождения и модель виртуального истока, была определена величина (17/гл. 8), имеющая размерность подвижности,

$$\mu = \frac{v_T \lambda_0}{2kT/q}. \quad (20)$$

Строго говоря, понятие подвижности определено только вблизи равновесия и для массивного проводника [30], и тем не менее модель прохождения удобно формулировать в традиционной форме и выражать среднюю длину свободного пробега через подвижность. Если скорость инжектированных электронов есть  $v_T$  и если квазиравновесная длина свободного пробега на вершине барьера  $\lambda_0$  та же самая, что и в длинноканальном транзисторе, тогда и подвижность по (20) есть та же самая подвижность, что и измеряемая в MOSFET с длинным каналом: мы часто использовали длинноканальную диффузионную подвижность для оценки квазиравновесной длины свободного пробега  $\lambda_0$  в наноразмерных FET.

Строя модель VS/прохождения, мы видели, что ток стока в линейной области пропорционален кажущейся подвижности, определяемой, согласно (20/гл. 8), из

$$\frac{1}{\mu_{app}} = \frac{1}{\mu} + \frac{1}{\mu_B} \quad (21)$$

и зависящей не только от диффузионной подвижности, но и от баллистической подвижности (18/гл. 8)

$$\mu_B = \frac{v_T L}{2kT/q}. \quad (22)$$

Мы видим, что кажущаяся подвижность, которую легко извлечь анализируя выходные характеристики, уменьшается для короткоканальных FET, поскольку баллистическая подвижность падает с уменьшением длины канала.

Для некоторых типов транзисторов зависимость кажущейся подвижности от длины определяется исключительно баллистической подвижностью (рис. 7/гл. 8), а для других типов оказывается, что  $\lambda_0$  уменьшается с сокращением длины канала (рис. 8/гл. 8). Причина уменьшения  $\lambda_0$  для короткоканальных FET еще до конца не понята. Некоторые наблюдения указывают на то, что причиной этого могут быть заряженные дефекты, введенные неумышленно при технологическом процессе [39]. Возможной причиной считают также дальнедействующее кулоновское рассеяние. В этом случае электроны в канале взаимодействуют на расстоянии с собственными электронами истока и стока и возбуждают плазменные колебания [40]. Такой дополнительный механизм рассеяния становится более эффективным с уменьшением длины канала. Использование металлических затворов вместо поликремниевых (poly-Si) могло бы экранировать дальнедействующие

кулоновские взаимодействия, но эти взаимодействия фундаментальны и должны бы присутствовать во всех типах FET. Ясное понимание причин уменьшения  $\lambda_0$  с укорочением канала – то ли причины носят фундаментальный характер, то ли все дело в используемой технологии и стало быть ситуация поправима, представляется особенно важным, когда длина канала становится меньше 20 нм.

Важно отдавать себе отчет в том, что все три типа подвижности в (21) в принципе зависят от длины канала:

$$\frac{1}{\mu_{app}(L)} = \frac{1}{\mu(L)} + \frac{1}{\mu_B(L)}.$$

Уравнения (21) и (22) можно скомбинировать и для зависимости  $\mu_{app}(L)$  получить следующее выражение:

$$\mu_{app}(L) = \frac{\mu}{1 + \alpha K_B \frac{\mu}{L}}, \quad K_B = \frac{2kT/q}{v_T},$$

где  $\alpha$  – численный коэффициент. Мы до сих пор пользовались только значением  $\alpha=1$ . Однако, уже ясно, что  $\alpha=1$  приемлемо не для всех типов транзисторов [29]. Для некоторых из них зависимость  $\mu_{app} \propto L$  требует  $\alpha < 1$ . Например, при анализе FinFET оказалось, что  $\alpha \approx 0.4$  [29]. Такое поведение объяснили тем, что электроны рассеиваются назад не только в канале проводимости, но и на границе перехода от канала к стоку [41]. В этой модели

$$\alpha = 1 - R_D,$$

где  $R_D$  есть коэффициент рассеяния назад на стоке. Объяснение  $\alpha \approx 0.4$  в FinFET требует большого значения этого коэффициента, примерно равного 0.6. Возможно, важны детали устройства границы раздела канал/сток, объясняющие почему у одних транзисторов  $\alpha \approx 1$ , а у других  $\alpha < 1$ . Нужны дополнительные исследования.

### *Роль рассеяния*

Утверждалось, что рассеяние в глубине канала, далеко от вершины барьера, не может играть существенной роли, поскольку электроны не смогут преодолеть барьер и вернуться в исток. Рассеяние, конечно, замедляет электроны, но поскольку имеется установившийся поток электронов из истока,

то внутри канала плотность электронов возрастает. Это изменяет электростатический потенциал везде, включая и область, охватывающую вершину барьера.

Обратимся к показательным результатам расчетов методом Монте-Карло DG SOI MOSFET с длиной канала 25 нм [42]. Рисунок 7 демонстрирует важную роль рассеяния в наноразмерном транзисторе, учет которого уменьшает ток в два раза.

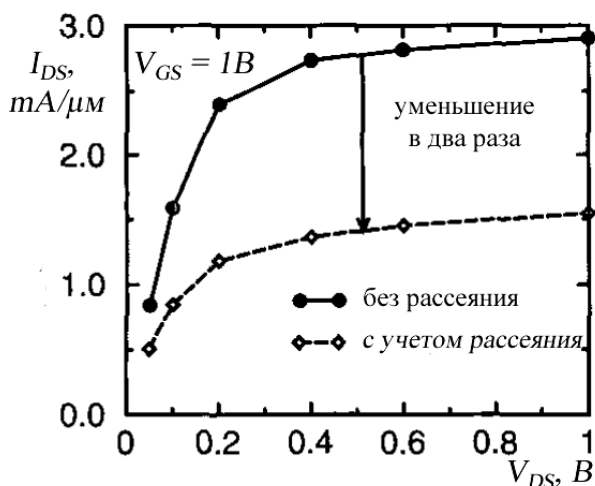


Рис. 7. Выходные характеристики DG SOI MOSFET с длиной канала 25 нм [42].

Поведение тока от истока к стоку и встречного тока показано на рис. 8 с указанием положения вершины барьера в канале. Ток стока  $I_{DS} = I^+ - I^-$ , конечно, не зависит от координаты в канале, его максимумы приходятся, естественно, на контакты истока и стока.

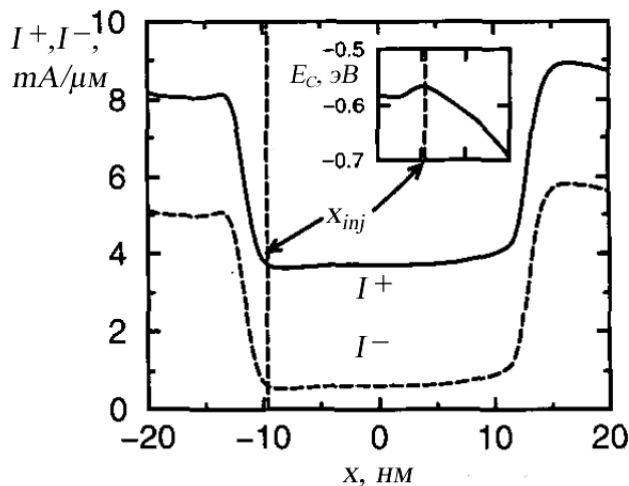


Рис. 8. Токи от истока к стоку  $I^+$  и от стока к истоку  $I^-$  в DG SOI MOSFET с длиной канала 25 нм при  $V_{GS} = V_{DS} = 1B$ . Во врезке показано положение вершины барьера [42].



Влияние рассеяния на плотность электронного заряда и профиль потенциала показано на рис. 9. Как и ожидалось, рассеяние увеличивает плотность инверсионного заряда в канале, что в свою очередь приводит к подъему и расширению зоны проводимости (рис. 9/справа). В результате критическая длина рассеяния назад увеличивается, прохождение и ток уменьшаются.

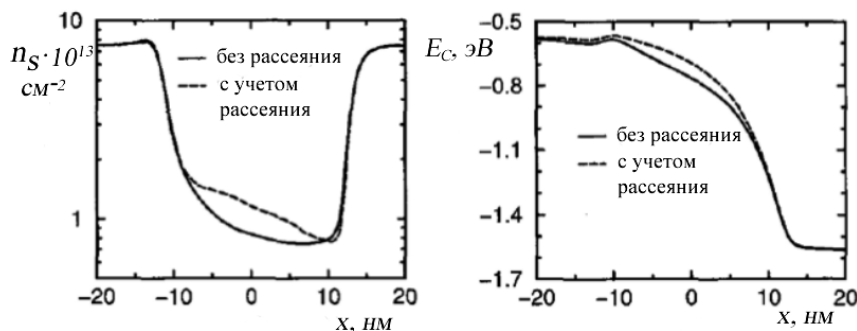


Рис. 9. Демонстрация влияния учета рассеяния в DG SOI MOSFET с длиной канала 25 нм при  $V_{GS} = V_{DS} = 1 \text{ В}$  на плотность инверсионного заряда вдоль канала проводимости (слева) и на профиль потенциальной энергии (справа) [42].

В итоге, однозначно приходим к выводу, что рассеяние в глубине канала влияет на величину тока [31, 35]. Однако, для качественно собранного транзистора влияние рассеяния невелико, поскольку в таких транзисторах потенциал на вершине барьера и вблизи нее контролируется преимущественно напряжением на затворе, а не напряжением на стоке или потенциалом в глубине канала проводимости. Об этом свидетельствует также тот факт, что в «хороших» транзисторах эффекты 2D электростатики (сдвиг подпороговой характеристики и выходная проводимость в режиме насыщения) в подпороговом режиме (инверсионный заряд в канале небольшой) и в надпороговом (большой заряд в канале) могут быть учтены одним и тем же параметром *DIBL*.

#### *Вырождение носителей тока*

При построении модели прохождения предполагалась максвелл-большцмановская статистика для электронов. Это предположение как будто бы неплохо работает для Si MOSFET [32] и более или менее приемлемо для III-V FET. Однако, для полупроводников с высокой подвижностью электронов учет их вырождения оказывается более существенным [43] и переход к статистике Ферми – Дирака позволяет построить более точные модели НЕМТ

[23, 24]. Как следует из (36/гл. 7), если учитывать статистику Ферми – Дирака, скорость баллистического впрыскивания увеличивается с ростом заряда  $|Q|$ . Соотношение между подвижностью и средней длиной свободного пробега при переходе к статистике Ферми – Дирака также изменяется от (50/гл. 5),

$$\mu = \frac{v_T \lambda_0}{2kT/q},$$

к [6, 20, 44]

$$\mu = \frac{v_T \langle \lambda \rangle}{2kT/q} \times \frac{\mathfrak{F}_{-1/2}(\eta_F)}{\mathfrak{F}_0(\eta_F)}.$$

Напомним, что  $v_T$  – это невырожденная однонаправленная тепловая скорость.

Емкость затвора в режиме сильной инверсии понижается при переходе к статистике Ферми – Дирака из-за уменьшения квантовой емкости. Вырождение аккуратно учтено в расширенной версии модели виртуального истока [25]. Полезные выражения при переходе к статистике Ферми – Дирака можно найти в [44].

#### *Плотность заряда и транспорт*

Ток стока пропорционален произведению заряда на скорость. В модели MVS/прохождения заряд на вершине барьера определяется электростатикой MOS посредством полуэмпирического выражения (32/гл. 8), зависящего только от напряжений на затворе и на стоке. Скорость впрыскивания, согласно (50/гл. 8), зависит от коэффициента прохождения. Однако, отделение заряда от скорости является лишь допущением, поскольку они связаны между собой.

Как показано на рис. 10/гл. 7, заряд на вершине барьера формируется положительными и отрицательными потоками и, согласно (36/гл. 7), зависит от прохождения следующим образом:

$$Q = -q \frac{N_{2D}}{2} [\mathfrak{F}_0(\eta_{FS}) + (1-T)\mathfrak{F}_0(\eta_{FS}) + T\mathfrak{F}_0(\eta_{FD})]. \quad (23)$$

В диффузионном пределе ( $T \ll 1$ ) состояния с положительными и отрицательными скоростями на вершине барьера заселены при любых напряжениях на стоке, однако, в баллистическом пределе ( $T \rightarrow 1$ ) при высоком напряжении на стоке заселены только состояния с положительными скоростями. Значение коэффициента прохождения определяет положение

уровня Ферми (через  $\eta_{FS}$ ), что в свою очередь определяет скорость баллистического впрыскивания  $v_{inj}^{ball}$ .

В модели MVS/прохождения мы используем уравнение (32/гл. 8) для определения заряда  $Q$  по значениям напряжений на затворе и на стоке, а затем пользуемся уравнением (23) для определения  $\eta_{FS}$ , после чего определяется  $v_{inj}^{ball}$ . Однако, в принципе сама величина заряда  $Q$  зависит от коэффициента прохождения  $T$ . Как обсуждалось ранее, емкость затвора в режиме инверсии определяется емкостью оксидного слоя и емкостью полупроводниковой подложки, согласно (99/гл. 3). В случае ETSOI емкость подложки есть просто квантовая емкость  $C_Q$  (103/гл. 3). В диффузионном пределе ( $T \ll 1$ ) состояния с положительными и отрицательными скоростями на вершине барьера заселены, и квантовая емкость в вырожденном пределе пропорциональна плотности состояний согласно (103/гл. 3). В баллистическом пределе ( $T \rightarrow 1$ ) заселены только состояния с положительными скоростями, так что квантовая емкость в вырожденном пределе пропорциональна половине плотности состояний. Это различие может быть важным для III-V FET, поскольку использование малых эффективных масс электронов приводит к малым значениям  $C_Q$ , что существенно понижает емкость затвора. Поскольку III-V FET работают вблизи баллистического предела, уже малые значения  $C_Q$ , уменьшающиеся еще в два раза при большом напряжении на стоке, могут быть важным фактором, определяющим функционирование III-V FET и должны быть учтены [23, 24].

Свою основную задачу мы видели в том, чтобы обратить внимание на существенные физические аспекты функционирования полевых нанотранзисторов, как это следует из экспериментов и численного моделирования. Именно эти физические аспекты полезны и достаточны для понимания и интерпретации результатов моделирования и экспериментальных данных при построении полуэмпирических моделей полевых транзисторов такого типа как модель MVS/прохождения. Это возможно только в том случае, если не упущено что-либо физически существенное. В отношении модели прохождения известны результаты численного моделирования, подтверждающие эту модель [35 – 38]. В частности, подтверждаются наши представления о том, что рассеяние, которое ограничивает ток  $I_{ON}$ , имеет место в небольшой области вблизи виртуального истока, а также о том, что вполне достижимы токи, близкие по величине к  $I_{ON}^{ball}$ , даже при наличии весьма

интенсивного рассеяния, лишь бы оно не происходило в узкой критической области [35].

Продемонстрируем отдельные важные результаты моделирования [34, 36] на рисунках 10 и 11. Отметим, что моделирование выполнялось самосогласованно, учитывалось уравнение Пуассона и квантовые ограничения, тщательно моделировались процессы рассеяния. Рисунок 10 показывает как заселяются  $k$ -состояния в 2D канале MOSFET с  $L = 25 \text{ нм}$  в режиме «ON» при высоких значениях напряжения как на стоке, так и на затворе.

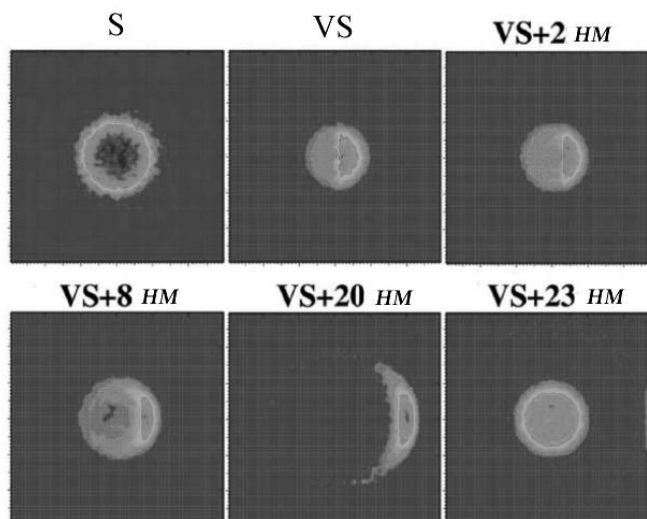


Рис. 10. Заселенность  $k$ -состояний в 2D канале MOSFET с  $L = 25 \text{ нм}$  в состоянии «ON» показана в шести локациях канала – истока, виртуального истока и далее вплоть до стока [34].

На истоке  $S$  имеет место симметричное квазиравновесное распределение заселенных  $k$ -состояний: исток является хорошим ландауэровским контактом. На вершине барьера  $VS$  распределение заселенных состояний существенно асимметрично: практически заселена половина состояний с положительными скоростями, а в результате рассеяния назад заселены лишь некоторые состояния с отрицательными скоростями. По мере вхождения вглубь канала протяженность заселенных состояний растет с увеличением электрического поля, распределение состояний становится все более и более асимметричным. Наконец, на стоке мы снова наблюдаем симметричное распределение заселенных состояний.

На рисунке 11 показана заселенность  $k$ -состояний в зависимости от скорости вдоль направления 2D канала MOSFET с  $L = 14 \text{ нм}$  в режиме «ON» при высоких значениях напряжения как на стоке, так и на затворе [36]. Показаны две ситуации – с учетом рассеяния (пунктир) и без рассеяния (сплошные).

Опять, на истоке наблюдается равновесное распределение, а на вершине барьера имеет место высоко асимметричное распределение (приблизительно половинное максвелловское). В баллистическом пределе, как и ранее на рис. 3/гл. 6, не наблюдаются электроны с отрицательными скоростями на вершине барьера, однако, как только «включается» рассеяние, появляется небольшая заселенность электронами с отрицательными скоростями.

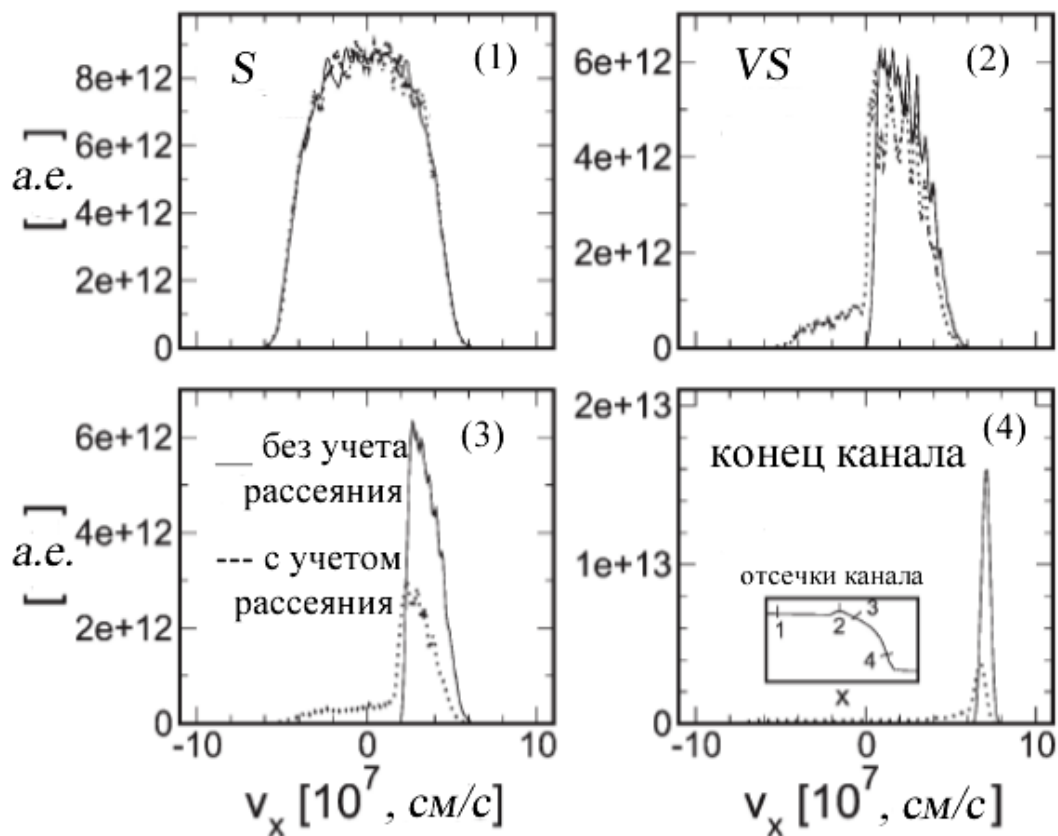


Рис. 11. Заселенность  $k$ -состояний в 2D канале MOSFET с  $L = 14$  нм в состоянии «ON» показана в четырех локациях канала – истока (1), виртуального истока (2), сразу после него (3) и в конце канала (4), где показаны также все четыре отсечки канала. Заселенность показана как функция скорости вдоль направления канала. Сплошные линии соответствуют отсутствию рассеяния, а пунктирные – наличию рассеяния везде по каналу. Заметны шумы, вызванные стохастическими процессами по ходу решения транспортного уравнения Больцмана [36].

По мере углубления в канал (3) вплоть до стока (4) по ходу увеличения электрического поля формируется все более симметричный баллистический пик. Аналогичная картина наблюдается при моделировании нанотранзисторов в чисто баллистическом режиме [45]. Можно надежно утверждать, что

монте-карловское моделирование транзисторов [35 – 38, 45] подтверждает физику процессов, положенных в основу модели MVS/прохождения.

Вместе с тем, другие не менее тщательные работы по численному моделированию процессов в нанотранзисторах ставят новые проблемы [46, 47]. Они связаны с той ролью, которую могут играть фундаментальные дальнедействующие кулоновские взаимодействия, которые не гаснут до конца, даже если использовать металлические затворы. Обсуждаются также уже упомянутые нами ранее эффекты голодания истока. Все такие эффекты особенно важно учитывать в III-V FET [23 – 25]. В этих работах подчеркивается также, что потенциальный барьер в районе виртуального истока не фиксирован, а подвижен. Он подвержен влиянию транспорта, а равновесный заряд навряд ли удерживается на вершине барьера при наличии тока. Роль транспортных эффектов на поведение заряда обсуждается в [23, 24]. Более подробное обсуждение вопросов, поднятых в [46, 47] в связи с моделью MVS/прохождения можно найти в [13].

Подведем итоги. Показано, что уже простейшая барьерная модель функционирования MOSFET позволяет эвристически оценить некоторые фундаментальные пределы работы нанотранзисторов в роли переключателей тока. Минимальная энергия переключения «ON»/«OFF» оказывается равной  $kT \ln 2$  и совпадающей с таким же результатом, следующим из принципа неопределенности Гейзенберга, а также соответствующей минимальной энергии, необходимой для стирания одного бита информации (Ландауэр). Оценены следующие фундаментальные пределы: минимальная длина канала проводимости и минимальное время пролета канала, минимальная плотность упаковки транзисторов и диссипация мощности. Достигнутые ныне минимальные размеры транзисторов и реальное время переключения уже того же порядка величин, что и известные фундаментальные пределы.

Представленные результаты моделирования свидетельствуют о том, что классическая модель прохождения MOSFET вполне оправдана применительно к Si транзисторам с длиной канала проводимости до 10 нм и даже еще несколько меньше.

Дается литературный анализ упрощающих допущений в модели прохождения, касающихся таких аспектов как неидеальные контакты, критическая длина при рассеянии назад, зависимость средней длины свободного пробега и подвижности от длины канала проводимости, критически

обсуждается роль рассеяния, вырождение носителей тока, плотность заряда, транспорт и другие вопросы.

Далее в завершающей главе полевой транзистор подробно рассматривается и обсуждается с позиций современной наноэлектроники «снизу – вверх» [6, 19, 21, 22].

## Литература

1. V. V. Zhirnov, R. K. Cavin III, J. A. Hutchby, G. I. Bourianoff, *Proc. IEEE*, **91**, 1934 – 1939 (2003).
2. R. Landauer, *IBM J. Res. Dev.*, **5**, 183 – 191 (1961).
3. C. Bennett, R. Landauer, *Scientific American*, **61**, 48 – 57 (1985).
4. C. H. Bennett, *History Phil. Mod. Phys.*, **34**, 501 – 510 (2003).
5. J. D. Meindl, J. A. Davis, *IEEE J. Solid State Circuits*, **35**, 1515 – 1516 (2000).
6. Ю. А. Кругляк, *Наноэлектроника «снизу – вверх»* (Одесса: ТЭС: 2015).
7. A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, E. Lutz, *Nature*, **483**, 187 – 189 (2012).
8. Y. Jun, M. Gavrilov, J. Bechhoefer, *Phys. Rev. Lett.*, **113**, 190601 (2014).
9. L. L. Yan, T. P. Xiong, K. Rehan, F. Zhou, D. F. Liang, L. Chen, J. Q. Zhang, W. L. Yang, Z. H. Ma, M. Feng, *Phys. Rev. Lett.*, **120**, 210601 (2018); [www.arxiv.org/abs/1803.10424v1](http://www.arxiv.org/abs/1803.10424v1).
10. M. Esposito, *Physics*, **11**, 49 (2018).
11. A. D. Franklin, M. Luisier, Shu-Jen Han, G. Tulevski, C. M. Breslin, L. Gignac, M. S. Lundstrom, W. Haensch, *Nano Lett.*, **12**, 758 – 762 (2012).
12. M. Luisier, M. S. Lundstrom, D. A. Antoniadis, J. Bokor, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 251 (2011); [www.doi.org/10.1109/IEDM.2011.6131531](http://www.doi.org/10.1109/IEDM.2011.6131531).
13. M. Lundstrom, *Fundamentals of Nanotransistors* (Singapore: World Scientific: 2018); [www.nanohub.org/courses/NT](http://www.nanohub.org/courses/NT).
14. J. Wang, M. Lundstrom, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 707 – 710 (2002).
15. R. Mehrotra, Sung Geun Kim, T. Kubis, M. Povolotskyi, M. S. Lundstrom, G. Klimeck, *IEEE Trans. Electron Dev.*, **60**, 2171 – 2177 (2013).
16. S. Datta, *Phys. Rev. B*, **40**, 5830 – 5833 (1989).
17. M. A. Alam, M. A. Stettler, M. S. Lundstrom, *Solid-State Electron.*, **36**, 263 – 271 (1993).
18. C. Jeong, R. Kim, M. Luisier, S. Datta, M. Lundstrom, *J. Appl. Phys.*, **197**, 023707 (2010).

19. S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport* (Singapore: World Scientific: 2012).
20. M. Lundstrom, C. Jeong, *Near-equilibrium transport. Fundamentals and Applications* (Singapore: World Scientific: 2013).
21. Ю. О. Кругляк, М. В. Стріха, *Укр. Фіз. Журн. Огляди*, **10**, 3 – 32 (2015).
22. S. Datta, *Lessons from Nanoelectronics. Part A: Basic Concepts* (Singapore: World Scientific: 2017).
23. S. Rakheja, M. Lundstrom, D. Antoniadis, *IEEE Trans. Electron Dev.*, **62**, 2786 – 2793 (2015).
24. S. Rakheja, M. Lundstrom, D. Antoniadis, *IEEE Trans. Electron Dev.*, **62**, 2794 – 2801 (2015).
25. U. Radhakrishna, T. Imada, T. Palacios, D. Antoniadis, *Phys. Status Solidi C*, **11**, 848 – 852 (2014).
26. J. Guo, S. Datta, M. Lundstrom, M. Brink, P. McEuen, A. Javey, H. Dai, H. Kim, P. McIntyre, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 711 – 714 (2002).
27. M. V. Fischetti, L. Wang, B. Yu, C. Sachs, P. M. Asbeck, Y. Taur, M. Rodwell, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 109 – 112 (2007).
28. R. Venugopal, S. Goasguen, S. Datta, M. S. Lundstrom, *J. Appl. Phys.*, **95**, 292 – 305 (2004).
29. D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **63**, 2650 – 2656 (2016).
30. M. Lundstrom, *Fundamentals of Carrier Transport* (Cambridge, U.K.: Cambridge Univ. Press: 2000).
31. M. S. Lundstrom, *IEEE Electron Dev. Lett.*, **18**, 361 – 363 (1997).
32. A. Majumdar, D. A. Antoniadis, *IEEE Trans. Electron Dev.*, **61**: 351 – 358 (2014).
33. G. Gildenblat, *J. Appl. Phys.*, **91**, 9883 – 9886 (2002).
34. R. Clerc, P. Palestri, L. Selmi, G. Ghibaudo, *J. Appl. Phys.*, **110**, 104502 (2011).
35. P. Palestri, D. Esseni, S. Eminent, C. Fiegna, E. Sangiorgi, L. Selmi, *IEEE Trans. Electron Dev.*, **52**, 2727 – 2735 (2005).
36. L. Lucci, P. Palestri, D. Esseni, L. Bergagnini, L. Selmi, *IEEE Trans. Electron Dev.*, **54**, 1156 – 1164 (2007).
37. J. Lusakowski, M. J. Martin Martinez, R. Rendal, T. Gonzalez, R. Tauk, Y. M. Meziani, W. Knap, F. Boef, T. Skotnicki, *J. Appl. Phys.*, **101**, 114511 (2007).
38. H. Tsuchiya, K. Fujii, T. Mori, T. Miyoshi, *IEEE Trans. Electron Dev.*, **53**, 2965 – 2971 (2006).



39. V. Barrel, T. Poiroux, S. Barrund, F. Andrieu, O. Faynot, D. Munteanu, J.-L. Autran, S. Deleonibus, *IEEE Trans. Nanotech.*, **8**, 167 – 173 (2009).
40. M. V. Fischetti, S. E. Laux, *J. Appl. Phys.*, **89**, 1205 – 1231 (2001).
41. K. Natori, H. Iwai, K. Kakushima, *J. Appl. Phys.*, **118**, 234502 (2015).
42. P. Palestri, D. Esseni, S. Eminente, C. Fiegna, E. Sangiorgi, L. Selmi, *Intern. Electron Dev. Mtg. (IEDM), Technical Digest*, 605 – 608 (2004).
43. А. Г. Самойлович, Л. Л. Коренблит, *Усп. физ. наук*, **LVII**, 577 – 630 (1955).
44. M. Lundstrom, X. Sun, *Some Useful Relations for Analyzing Nanoscale MOSFETs Operating in the Linear Region* (West Lafayette, Indiana: Purdue University, USA, 2016); [www.arxiv.org/abs/1603.03132](http://www.arxiv.org/abs/1603.03132).
45. J.-H. Rhew, Zhibin Ren, M. Lundstrom, *Solid-State Electron.*, **46**, 1899 – 1906 (2002).
46. M. V. Fischetti, S. Jin, T.-W. Tang, P. Asbeck, Y. Taur, S. E. Laux, M. Rodwell, N. Sano, *J. Comp. Electron.*, **8**, 60 – 77 (2009).
47. M. V. Fischetti, S. T. P. O`Regan, S. Narayanan, C. Sachs, S. Jin, J. Kim, Y. Zhang, *IEEE Trans. Electron Dev.*, **54**, 2116 – 2136 (2007).

## Глава 10. Полевой транзистор с позиций нанoeлектроники «снизу – вверх»

### 10.1. Введение

«Без-полевой» подход к электрической проводимости, характерный для нанoeлектроники «снизу – вверх» [1 – 5], вызывает удивление у многих, поскольку обычно и привычно полагают, что электрический ток возникает под действием внешнего приложенного электрического поля. Спешим сразу уточнить, что электрическое поле действительно играет важную роль как только мы приложим к проводнику достаточно большóую разность потенциалов. Целью этой главы станет обсуждение влияния электростатического потенциала и соответствующего внешнего электрического поля на вольт-амперную зависимость за пределами режима линейного отклика. Для иллюстрации этого влияния как ничто другое подходит полевой транзистор, который фактически представляет собой «варистор» (voltage-controlled resistor) (рис. 7/гл. 1), длина проводящего канала которого в нанотранзисторах уже составляет всего несколько сот атомов. Функциональность полевого транзистора определяется ходом всей его вольт-амперной характеристики (ВАХ), которая собственно и контролируется электростатически. Мы уже подробно обсудили физику процессов в полевых транзисторах. Сейчас же мы воспользуемся транзистором, чтобы еще раз, но теперь с позиций нанoeлектроники «снизу – вверх», проиллюстрировать роль электростатики в протекании тока.

Ток в модели упругого резистора определяется формулой (201/[2/гл.1]):

$$I = \frac{1}{q} \int_{-\infty}^{+\infty} dE G(E) (f_1(E) - f_2(E)),$$

где проводимость

$$G(E) = \frac{q^2 D(E)}{2\tau(E)}$$

выражена через плотность состояний  $D(E)$  и среднее время  $\tau(E)$ , которое электрон проводит в канале на его пути от левого контакта 1 к правому 2.

При учете внешнего электростатического потенциала  $U$

$$I = \frac{1}{q} \int_{-\infty}^{+\infty} dE G(E - U) (f_1(E) - f_2(E)); \quad (1)$$

на первых порах будем пользоваться точечной моделью канала переноса электронов, которая предполагает неизменность плотности состояний  $D(E)$

при движении вдоль проводника. Позже перейдем к стандартной расширенной модели проводника, учитывающей градиент  $D(E)$  при переходе вдоль проводника от истока к стоку.

## 10.2. Соотношение между током и напряжением

Рассмотрим простейший трехтерминальный транзистор: напряжение на затворе  $V_G$  позволяет контролировать ток  $I(V_D)$  в канале проводимости. Характерная вольт-амперная характеристика полевого транзистора показана на рис. 1.

Все ясно с низковольтной проводимостью: ничего нового, только закон Ома. Однако, с ростом напряжения на стоке  $V_D$  происходит нарушение закона Ома и появляются эффекты, требующие пересмотра и дополнения уже привычных представлений.

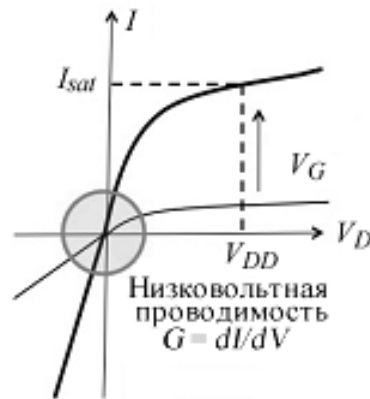


Рис. 1. Типичная вольт-амперная характеристика  $I(V_D)$  полевого транзистора и ее изменение с ростом потенциала затвора  $V_G$ . Кружком показана область режима линейного отклика с низковольтной проводимостью  $G = dI/dV$ .

Пусть потенциал на затворе нулевой, а в щели  $E_{F1} - E_{F2}$  нет электронных состояний проводника (рис. 2): тока нет.

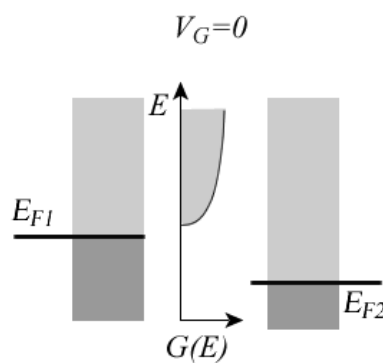


Рис. 2. В проводнике нет тока.

При подаче положительного потенциала на затвор  $V_G > 0$  все уровни энергии проводника понижаются, ток увеличивается тем больше, чем больше положительный потенциал на затворе (рис. 3), в чем можно убедиться подставляя  $U = -qV_G$  в ур-е (1).

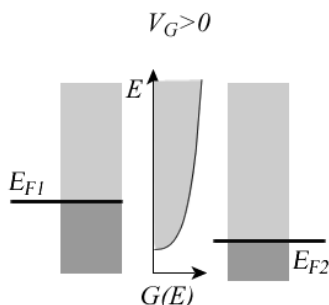


Рис. 3. Положительный потенциал на затворе понижает все уровни энергии проводника. Ток увеличивается тем больше, чем больше положительный потенциал на затворе.

Такова ситуация для  $n$ -проводников. Для проводников  $p$ -типа плотность состояний нисходящая с ростом энергии и возрастание положительного потенциала на затворе ведет к противоположному эффекту по току. Далее будем рассматривать только проводники  $n$ -типа.

Сосредоточим свое внимание на участке вольт-амперной характеристики за пределами низковольтной проводимости и зададимся вопросами – почему зависимость  $I \propto V_D$  становится нелинейной и «выпрямляющей», т. е. различной для положительных и отрицательных значений напряжения на стоке  $V_D$ .

### 10.3. Почему ток насыщается?

По мере роста напряжения  $V_D$  ток не возрастает линейно согласно закону Ома, а стремится к насыщению (рис. 1). Почему? Причина довольно очевидна. Как только электрохимический потенциал стока  $E_{F2}$  опускается ниже дна зоны проводимости (рис. 4), увеличение тока произойти уже не может.

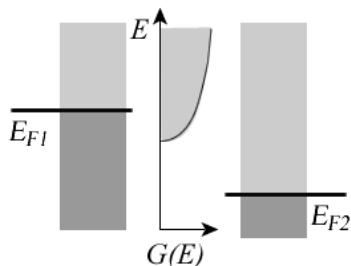


Рис. 4. Ток испытывает насыщение как только электрохимический потенциал стока  $E_{F2}$  оказывается ниже дна зоны проводимости.

Величина тока насыщения определяется из ур-я (1), а именно:

$$I_{sat} = \frac{1}{q} \int_{-\infty}^{+\infty} dE G(E-U) f_1(E), \quad (2)$$

где фермиевская функция на стоке опущена, поскольку предполагается, что для всех отличных от нуля мод проводимости фермиевская функция на стоке нулевая.

Если бы ток насыщения определялся бы только значением потенциала на затворе,

$$U^{(1)} = -qV_G,$$

то насыщение было бы полным (горизонтальная полочка) как только  $E_{F2}$  окажется на несколько  $kT$  ниже дна зоны проводимости. В реальных же экспериментах ток продолжает расти с увеличением  $V_D$  (рис. 5). Причина заключается в том, что увеличение потенциала на стоке  $V_D$  не только ведет к понижению электрохимического потенциала  $E_{F2}$ , но и к понижению всех уровней энергии проводника (рис. 5) подобно тому, как это происходит при увеличении потенциала на затворе  $V_G$ .

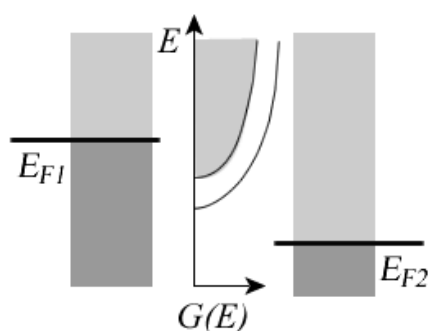


Рис. 5. При понижении  $E_{F2}$ , что соответствует росту потенциала на стоке  $V_D$ , понижаются также все уровни в зоне проводимости проводника.

В конечном результате ток продолжает несколько возрастать по мере того, как проводимость  $G(E)$  ползет вниз по энергии на некоторую долю  $\alpha$  ( $\alpha < 1$ ) от потенциала на стоке  $V_D$ , что и отражено ниже в более точной модели для тока насыщения:

$$U^{(2)} \rightarrow U_L \equiv \alpha(-qV_D) + \beta(-qV_G). \quad (3)$$

Задача создания хорошего полевого транзистора сводится к тому, чтобы величину  $\alpha$  сделать как можно меньшей и тем самым влияние потенциала на

стоке  $V_D$  свести к минимуму. Если величина  $\alpha$  окажется нулевой, ток выйдет на полное насыщение (рис. 6), что и соответствует идеальному полевому транзистору: в таком транзисторе ток насыщения полностью определяется потенциалом на затворе  $V_G$  и никак не зависит от потенциала на стоке  $V_D$ .

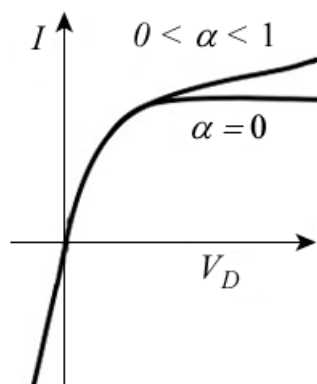


Рис. 6. Ток в полевом транзисторе испытывает полное насыщение, если он в режиме насыщения не подвержен влиянию потенциала на стоке  $V_D$ .

Чтобы гарантировать существенно преобладающий контроль над каналом проводимости со стороны затвора по сравнению со стоковым терминалом, нужно толщину диэлектрической прокладки затвора сделать существенно меньшей длины проводника. При типичной длине проводника в современных нанотранзисторах в несколько сот атомов, толщина диэлектрической прокладки должна быть лишь несколько атомов, чтобы обеспечить нужное малое значение  $\alpha$ . Поразительно, как такая толщина с высокой надежностью контролируется в производстве современных ноутбуков.

#### 10.4. Роль заряжения проводника

Есть еще один эффект, который ведет к увеличению тока насыщения сверх того значения, которое получается при использовании (3) в ур-и (1). В проводнике под напряжением заселенность состояний проводника меньше по сравнению с заселенностью проводника в состоянии динамического равновесия, когда оба контакта стремятся максимально заполнить доступные состояния проводника. Тогда как под напряжением исток стремится пополнить электронами состояния проводника, а сток стремится электроны удалить из проводника. Уменьшение максимально возможной заселенности электронами состояний проводника под напряжением ведет к заряжению проводника положительным зарядом, что в свою очередь ведет к понижению всех состояний проводника (рис. 5) и даже в случае идеального транзистора ( $\alpha = 0$ ) увеличивает ток.

Учесть эффект заряджания в рамках точечной модели проводника можно путем добавления к потенциалу проводника  $U_L$  (3) еще одного слагаемого

$$U = U_L + U_0(N - N_0), \quad (4)$$

которое учитывает отличие равновесного числа электронов в проводнике  $N_0$  от его неравновесного значения  $N$  в условиях нахождения проводника под напряжением, где  $U_0$  есть вклад в потенциал проводника в пересчете на один электрон.

Равновесное число электронов в проводнике  $N_0$  вычисляется через равновесное значение фермиевской функции  $f_0(E)$ :

$$N_0 = \int_{-\infty}^{+\infty} dE D(E - U) f_0(E), \quad (5)$$

а соответствующее неравновесное значение

$$N = \int_{-\infty}^{+\infty} dE D(E - U) \frac{f_1(E) + f_2(E)}{2} \quad (6)$$

в предположении, что проводник с обоими контактами взаимодействует одинаково хорошо.

Как только найден потенциал проводника  $U$ , удовлетворяющий одновременно уравнениям (4) – (6), ток вычисляется по (1) или по эквивалентному ему уравнению

$$I = \frac{1}{q} \int_{-\infty}^{+\infty} dE G(E) (f_1(E + U) - f_2(E + U)). \quad (7)$$

Вычисление потенциала проводника  $U$ , сводится к итерационному процессу, принципиальная блок-схема которого показана на рис. 7.

Простая точечная модель проводника оказывается вполне приемлемой, как это показано в [6].

Результаты численного расчета вольт-амперных характеристик 2D проводника с параболической дисперсией и плотностью состояний [7 – 9]

$$D(E) = g \frac{mLW}{2\pi\hbar^2} H(E - E_c), \quad (8)$$

где  $H(E - E_c)$  – ступенчатая функция Хэвисайда, показаны на рис. 8.



Рис. 7. Самосогласованная процедура вычисления потенциала проводника  $U$  и тока в точечной модели проводника.

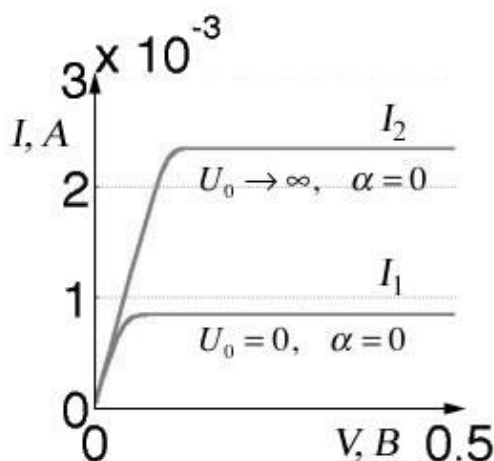


Рис. 8. Вольт-амперные характеристики 2D проводника с параболической дисперсией в точечной модели проводника в режиме баллистического транспорта [1, 5].

Вычисления проводились для  $g = 2$ ,  $m = 0.2 \times 9.1 \times 10^{-31} \text{ Кг}$ ,  $kT = 5 \text{ мэВ}$ ,  $\beta = 1$ ,  $\alpha = 0$  и  $U_0 = 0$  или  $U_0 = \infty$ ,  $L = W = 1 \mu\text{м}$  в модели баллистического транспорта, так что

$$G(E) = \frac{q^2}{h} M(E),$$

а число мод проводимости [8, 9]

$$M(E) = g \frac{2W}{h} \sqrt{2m(E - E_c)} H(E - E_c). \quad (9)$$



Вычисленные характеристики (рис. 8) имеют две характерных области: линейный рост и ток насыщения. Хотя эти графики и были получены численно, однако, и наклон и ток насыщения могут быть вычислены аналитически, особенно просто если принять низкотемпературное приближение, в котором фермиевские функции изменяются скачкообразно от единицы до нуля в узком промежутке, включающем  $E = E_F$ : действительно, в расчетах использовалось значение теплового кванта энергии  $kT = 5 \text{ мэВ}$  вместо его значения при комнатной температуре  $kT = 25 \text{ мэВ}$ , так что приведенные численные характеристики ближе сопоставимы с низкотемпературными ВАХ.

Этими графиками мы хотим проиллюстрировать два важных результата. Начальный наклон ВАХ не зависит от энергии заряжания  $U_0$  в (4). Он определяется низковольтной проводимостью. Именно по этой причине теорию проводимости в режиме линейного отклика мы имеем право строить без учета электростатики.

Во-вторых, ток насыщения сильно зависит от электростатики и увеличивается в  $\sim 2.8$  раз при переходе от модели с нулевой энергией заряжания  $U_0$  к модели с очень большим значением  $U_0$ . Причина этого фактически уже упоминалась выше. При нулевой энергии заряжания спектр состояний канала проводимости остается неизменным и число электронов  $N$  в канале проводимости равно  $N_0/2$ , поскольку  $f_1 = 1$  и  $f_2 = 0$ . При очень большом значении  $U_0$  во избежание доминирования  $U_0(N - N_0)$  значение  $N$  должно быть практически равным его равновесному значению  $N_0$  даже при наполовину заполненных состояниях. С этой целью состояния в проводнике должны опуститься ниже при увеличении тока, как это показано на рис. 5.

## 10.5. Квантовая емкость

Нас обычно интересовал профиль вольт-амперной характеристики  $I \propto V_D$ , когда разность потенциалов приложена между истоком и стоком. Это неравновесная задача. Давайте обсудим одну равновесную задачу, в которой основным эффектом является заряжание. Пусть исток и сток находятся при одном и том же потенциале, а на затвор подается напряжение. Как при этом изменяется число электронов  $N$  в канале проводника?

Положительное напряжение, поданное на затвор, понижает плотность состояний  $D(E)$ , в результате чего число электронов  $N$  растет. Из (6) имеем:

$$N = \int_{-\infty}^{+\infty} dE D(E) f_0(E + U). \quad (10)$$

Предполагая, что электрохимический потенциал находится вне зоны проводимости (рис. 9), допустимо пользоваться больцмановским приближением функции Ферми:

$$f_0(E) \approx e^{-(E-E_F)/kT},$$

так что

$$N = e^{-U/kT} \times \int_{-\infty}^{+\infty} dE D(E) f_0(E).$$

Чтобы изменить число электронов в 10 раз, потенциал  $U$  в канале нужно изменить на  $kT \ln 10 \approx 60 \text{ мэВ}$ .

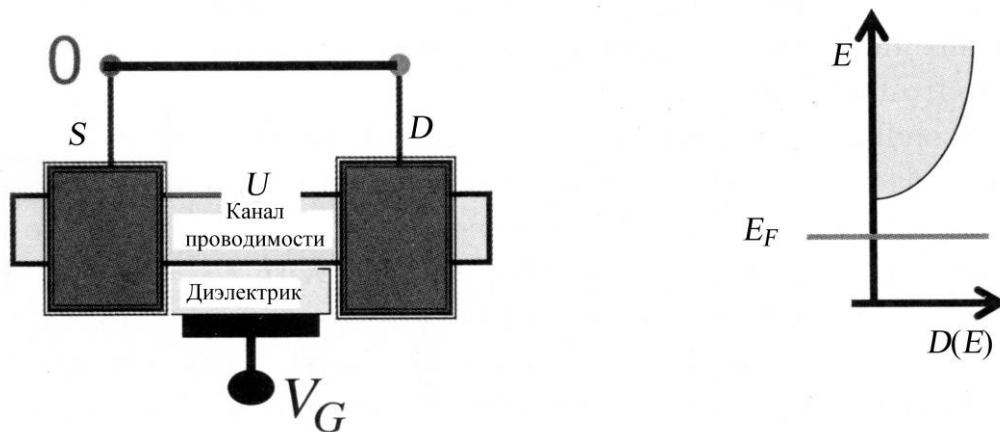


Рис. 9. Исток и сток находятся при одном и том же потенциале, а на затвор подано напряжение при электрохимическом потенциале вне зоны проводимости.

Для изменения потенциала в канале  $U$  на  $60 \text{ мэВ}$  на затвор нужно подать напряжение не менее  $60 \text{ мВ}$ , что собственно и ведет к привычному утверждению, что каждое увеличение числа электронов в канале в 10 раз требует подачи напряжения на затвор не менее  $60 \text{ мВ}$ .

Подобные рассуждения имеют смысл только тогда, когда число электронов в канале относительно невелико. По мере приближения электрохимического потенциала к зоне проводимости становится все более важным учитывать эффект зарядания.

Подача положительного напряжения на затвор понижает плотность состояний (сплошная кривая на рис. 10 слева), что ведет к увеличению числа

электронов в канале, а это в свою очередь ведет к подъему плотности состояний (пунктирная кривая там же). Для подсчета результирующего числа электронов в канале этот эффект зарядания нужно учесть.

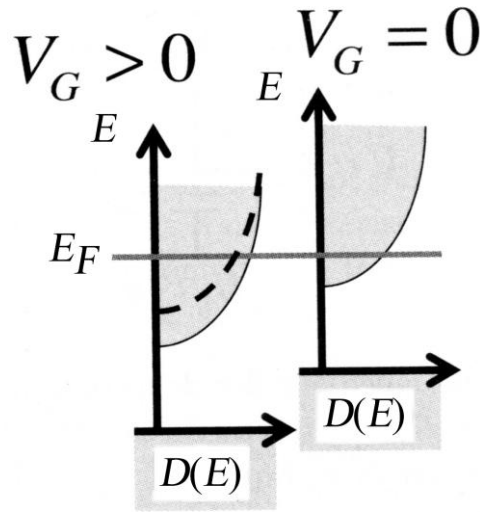


Рис. 10. К влиянию зарядания на число электронов в канале проводимости [5].

Разобьем эту задачу на две части – 1) изменение числа электронов, вызванное изменением потенциала  $U$  в канале и 2) изменение потенциала  $U$ , связанное с подачей напряжения  $V_G$  на затвор, а именно:

$$C = q \frac{dN}{dV_G} = q \frac{dN}{(dU/-q)} \times \frac{dU}{d(-qV_G)}. \quad (11)$$

Первый множитель – это квантовая емкость, которую можно связать с плотностью состояний (8), но уже не прибегая к больцмановскому приближению функции Ферми:

$$C_Q = q \frac{dN}{(dU/-q)} = q^2 \int_{-\infty}^{+\infty} dE D(E) \left( -\frac{\partial f_0}{\partial E} \right) = q^2 D_0, \quad (12)$$

где  $D_0$  представляет собой усредненное значение плотности состояний  $D(E)$  вблизи электрохимического потенциала  $E_F = E + U$ , т. е. вокруг  $E = E_F - U$ .

Для оценки второго множителя в (11) воспользуемся выражением (4) с привлечением (12), а именно:

$$\frac{dU}{d(V_G)} = \frac{dU_L}{d(V_G)} + U_0 \frac{dN}{d(V_G)} = \frac{dU_L}{d(V_G)} - U_0 D_0 \frac{dU}{d(V_G)}.$$

Следовательно, второй сомножитель в (11)

$$\frac{dU}{d(-qV_G)} = \frac{1}{1+U_0D_0} \times \frac{dU_L}{d(-qV_G)} \quad (13)$$

характеризует изменение потенциала  $U_L$  (3) с напряжением на затворе, которое в идеале должно стремиться к единице. Однако, действительное изменение потенциала  $U$  (13), связанное с подачей напряжения  $V_G$  на затвор, может быть намного меньше, согласно (13), что зависит от произведения средней плотности состояний  $D_0$  и энергии заряжания  $U_0$  в пересчете на один электрон.

Соотношение (13) можно как бы визуализировать (рис. 11) с привлечением квантовой емкости (12) и обычной электростатической емкости  $C_E$ , связанной с энергией заряжания  $U_0$ , а именно:

$$U_0 = \frac{q^2}{C_E},$$

так что

$$U_0D_0 = \frac{C_Q}{C_E}. \quad (14)$$

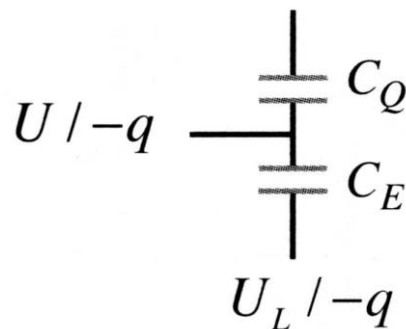


Рис. 11. К «визуализации» соотношения (13) [5].

Электростатическая емкость характеризует энергию заряжания, связанную с увеличением потенциала в канале  $q/C_E$  при добавлении в него еще одного электрона.

В наноэлектронике «снизу – вверх» широко используется концепция плотности состояний  $D$ , однако, почти не обсуждаются эффекты заряжания  $U$ . Связано это с тем, что эти эффекты не существенны для низковольтной проводимости в однородных проводниках, но в целом нужно иметь в виду, что эффекты заряжания требуют внимания при анализе электронных устройств как в равновесных условиях, так и за пределами равновесия.

## 10.6. Электростатический «выпрямитель»

Рассмотрим высококачественный полевой транзистор (ПТ) с  $\alpha = 0$  как в стандартной конфигурации (рис. 12а), так и с плавающим затвором (рис. 12б).

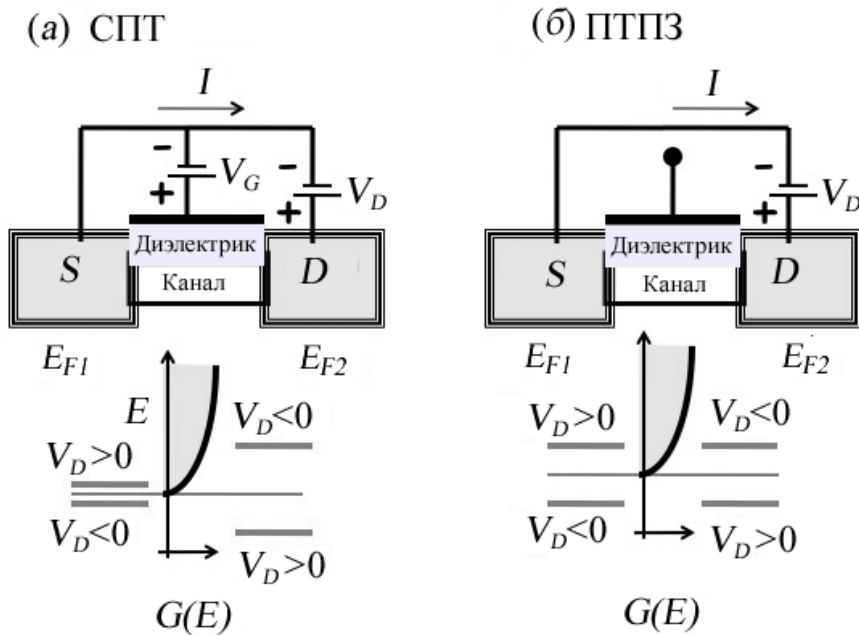


Рис. 12. Полевой транзистор: а) стандартный (СПТ); б) с плавающим затвором (ПТПЗ).

Пусть контакты  $S$  и  $D$  в транзисторе одинакового качества, иначе говоря, симметричны. В реальной ситуации это не так, однако, для нас это сейчас не важно, поскольку мы хотим качественно продемонстрировать два полезных результата. В стандартной конфигурации транзистора (рис. 12а) имеет место электростатическая асимметрия, поскольку затвор  $G$  поддерживается при фиксированном потенциале  $V_G$  относительно истока  $S$  (рис. 5/гл. 1), но не относительно стока  $D$ . В транзисторе же с плавающим затвором (рис. 12б) имеет место электростатическая симметрия. Мы хотим проиллюстрировать один существенный результат, а именно: контактно-симметричный полевой транзистор может иметь асимметричную характеристику (условно называемый «выпрямитель»), иначе говоря,

$$I(+V_D) \neq I(-V_D)$$

просто в результате электростатической асимметрии.

На рис. 13 показаны характеристики, рассчитанные для приведенной выше точечной модели проводника (рис. 8) за исключением того что

электрохимический потенциал расположен точно на дне зоны проводимости, как это показано на рис. 12:  $E_{F0} = E_C$ .

В случае стандартного транзистора полагалось  $\alpha = 0$ , что соответствует высококачественному полевому транзистору, тогда как расчет ПТПЗ проводился с  $\alpha = 0.5$ .

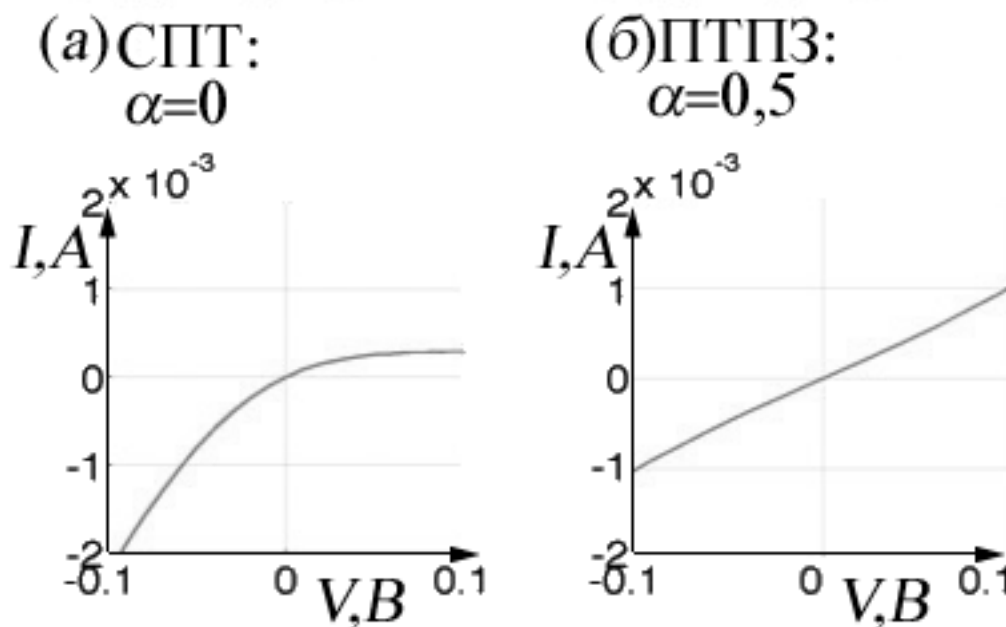


Рис. 13. Выходные характеристики для: а) идеального ( $\alpha = 0$ ) стандартного полевого транзистора (СПТ); б) полевого транзистора с плавающим затвором (ПТПЗ) при  $\alpha = 0.5$  [1, 2, 4, 5].

В первом случае имеет место «выпрямляющая» выходная характеристика (рис. 13а), а во втором случае – линейная, т.е. «омическая» характеристика (рис. 13б). Вывод можно сделать следующий: для получения асимметричной выходной характеристики нет нужды конструировать полевой транзистор с асимметричными контактами. Даже симметричный по контактам транзистор может иметь несимметричную «выпрямляющую» характеристику, если имеет место электростатическая асимметрия.

Обратим также внимание на то обстоятельство, что наклон выходной характеристики  $dI/dV|_{V \rightarrow 0}$  не зависит от  $\alpha$ , что видно из сравнения наклона характеристик при  $V = 0$  на рис. 13а и рис. 13б, и может быть предсказан без учета энергии заряжания, т.е. без учета электростатики.

## 10.7. Расширенная модель проводника

Точечная модель проводника не учитывает градиент электрического поля в канале проводимости и предполагает, что плотность состояний  $D(E)$  остается постоянной на всем протяжении проводника от истока до стока. В реальности электрическое поле понижает уровни энергии проводника вблизи стока в большей степени по сравнению с истоком, в результате чего наблюдается градиент плотности состояний  $D(z, E)$  (рис. 14). Как это влияет на ток?

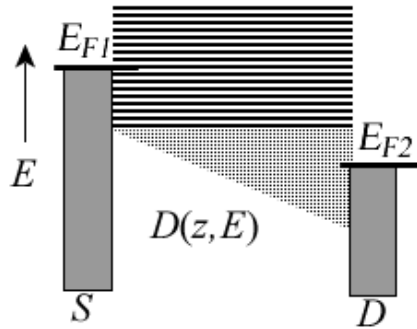


Рис. 14. Ненулевое электрическое поле в канале проводимости больше понижает состояния проводника ближе к стоку по сравнению с истоком.

На первый взгляд кажется, что в случае упругого резистора «косоидущие состояния» не могут дать вклад в проводимость, поскольку в упругом резисторе вклад в проводимость могут дать лишь состояния с постоянной энергией, начинающиеся на истоке и заканчивающиеся на стоке. Однако, даже для идеального упругого резистора должно иметь место увеличение тока, поскольку число мод вблизи стока больше числа мод у истока (рис. 15).

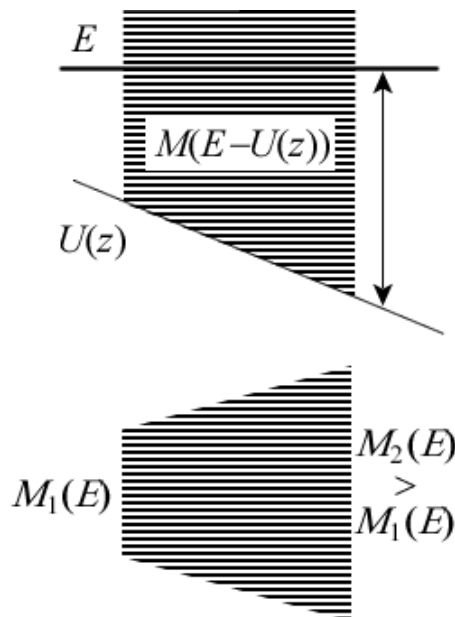


Рис. 15. Градиент электрического поля в канале проводимости приводит к большему числу мод вблизи стока по сравнению с истоком.

Вызвано это тем, что число мод с энергией  $E$  зависит от того, насколько далеко эта энергия находится от дна зоны проводимости, положение которого определяется потенциалом  $U(z)$ , значение которого ниже у стока по сравнению с истоком (рис. 15).

Структура спектра состояний проводника выглядит как бы более растянутой у стока по сравнению с истоком. Это не существенно для баллистического проводника, а вот для проводника, длина которого  $L$  намного больше средней длины свободного пробега  $\lambda$ , уширение спектра на стоке может увеличить проводимость.

В общем случае можно написать, что

$$\frac{q^2}{h} \frac{M_1 \lambda}{L + \lambda} \leq G(E) \leq \frac{q^2}{h} M_1. \quad (15)$$

Этот эффект не существенен для квазибаллистических упругих каналов, поскольку для таких каналов минимальные и максимальные значения проводимости фактически одинаковы. На самом деле увеличение проводимости можно приписать зависимости средней длины свободного пробега от электрического поля; этим можно пренебречь в режиме линейного отклика, чем мы фактически и пользовались.

### 10.8. Уравнение диффузии и расширенная модель проводника

Как учесть градиент электрического поля количественно? Для этого в диффузионном уравнении

$$\frac{I}{A} = -\frac{\sigma}{q} \frac{dE_F}{dz}$$

учтем пространственное изменение удельной проводимости вдоль длины проводника:

$$\frac{I}{A} = -\frac{\sigma(z)}{q} \frac{dE_F}{dz}. \quad (16)$$

Если ранее удельная проводимость усреднялась по энергии в пределах фермиевского окна проводимости шириной в несколько  $kT$  в окрестности  $E = E_{F0}$ :

$$\sigma = \int_{-\infty}^{+\infty} dE \sigma(E) \left( -\frac{\partial f}{\partial E} \right)_{E=E_{F0}}, \quad (17)$$



то теперь переменный вдоль длины проводника электростатический потенциал  $U(z)$  сдвигает состояния проводника по энергии, в результате чего усреднение нужно вести в окрестности  $E = E_F(z) - U(z)$ :

$$\sigma = \int_{-\infty}^{+\infty} dE \sigma(E) \left( -\frac{\partial f}{\partial E} \right)_{E=E_F(z)-U(z)}. \quad (18)$$

Обратим внимание на то, что использование уравнений (16) и (18) требует вычисления  $E_F(z) - U(z)$  из итерационной процедуры решения уравнения Пуассона

$$\frac{d}{dz} \left( \varepsilon \frac{dU}{dz} \right) = q^2 (n - n_0), \quad (19)$$

где  $\varepsilon$  – относительная диэлектрическая проницаемость прокладки затвора, а  $n_0$  и  $n(z)$  – электронная плотность на единицу длины в канале проводимости в состоянии равновесия и вне равновесия, которые можно вычислить согласно определению, а именно:

$$n(z) \equiv \int_{-\infty}^{+\infty} dE \frac{D(E - U(z))}{L} \frac{1}{e^{(E - E_F(z))/kT} + 1}, \quad (20)$$

$$n_0 = \int_{-\infty}^{+\infty} dE \frac{D(E)}{L} \frac{1}{e^{(E - E_{F0})/kT} + 1}. \quad (21)$$

Принципиальная блок-схема вычисления тока в расширенной модели проводника показана на рис. 16.



Рис. 16. Самосогласованная процедура вычисления потенциала проводника  $U$  и тока в расширенной модели проводника.

Нужно иметь в виду, что изложенный выше подход основан на модели локальных электрохимических потенциалах  $E_F^\pm(z)$ , описывающих транспорт электронов направо и налево, среднее значение которых фигурирует в уравнении диффузии (71/[2/гл.7]). В общем же случае распределение электронов может настолько сильно отличаться от фермиевского, что концепция электрохимических потенциалов окажется неадекватной и потребуются решать транспортное уравнение Больцмана, что является несравненно более сложной задачей. Несмотря на значительные успехи в этом направлении, уравнение диффузии, основанное на концепции локальных электрохимических потенциалов, по-прежнему остается главным инструментом в моделировании электронных устройств.

Концепция «снизу – вверх» расширяет область применимости уравнений диффузии (71/[2/гл.7]) и (16). Ими можно пользоваться и для моделирования баллистических каналов, если только должным образом модифицировать граничные условия с целью учесть граничные сопротивления.

## Литература

1. S. Datta, *Lessons from Nanoelectronics: A New Perspective on Transport* (Singapore: World Scientific: 2012).
2. Ю. А. Кругляк, *Нанoeлектроника «снизу – вверх»* (Одесса: ТЭС: 2015).
3. Ю. О. Кругляк, М. В. Стриха, *Укр. Фіз. Журн., Огляди*, **10**, 3 – 32 (2015).
4. Ю. А. Кругляк, *Нанoeлектроника «снизу – вверх»*, изд. 2-ое дополненное (Киев: Изд-во Стрельбицкого: 2016).
5. S. Datta, *Lessons from Nanoelectronics. Part A: Basic Concepts* (Singapore: World Scientific: 2017).
6. A. Rahman, Guo Jing, S. Datta, M. Lundstrom, *IEEE Trans. Elect. Dev.*, **50**: 9, 1853 – 1864 (2003).
7. R. F. Pierret. *Semiconductor Device Fundamentals*. (Reading, MA: Addison–Wesley: 1996).
8. Ю. А. Кругляк. *ScienceRise*, т. 5: 3(5), 21 – 38 (2014).
9. Yu. A. Kruglyak, M. V. Strikha, *Proc. 2015 IEEE 35th Intern. Conf. Electronics Nanotech. (ELNANO)*, April 21 – 24, 2015, Kyiv, Ukraine, pp. 70 – 74; DOI: 10.1109/ELNANO.2015.7146837.

Підписано до друку 14.08.2018 р. Формат 60x84/16  
Папір офсетний. Ум. друк. арк. 18,37  
Наклад 100 прим. Замовлення 840  
Видавництво та друкарня "ТЕС"  
(Свідоцтво ДК № 771) Одеса, Канатна 81/2.  
Тел.:(0482)42-90-98, (0482)42-89-72



**Юрий Алексеевич Кругляк**  
 Профессор, д. ф.-м. н., д. хим. н.,  
 Действительный член Академии наук  
 высшей школы Украины по  
 Отделению физики и астрономии

Ю. А. Кругляк

НАНОЭЛЕКТРОНИКА  
 «СНИЗУ – ВВЕРХ»

$$I = \frac{2q}{h} \int T(E) M(E) (f_1(E; E_{F1}, T_1) - f_2(E; E_{F2}, T_2)) dE$$

$$f_1(E) = \frac{1}{e^{(E-E_{F1})/kT_1} + 1} \quad f_2(E) = \frac{1}{e^{(E-E_{F2})/kT_2} + 1}$$

$$T(E) = \frac{\lambda(E)}{\lambda(E) + L} \rightarrow \begin{cases} L \gg \lambda; & T \ll 1 \\ L = \lambda; & T < 1 \\ L \ll \lambda; & T \rightarrow 1 \end{cases}$$

$$M(E) = \begin{cases} M_{1D}(E) = \frac{h}{4} v_x^*(E) D_{1D}(E) \\ M_{2D}(E) = W \frac{h}{4} v_x^*(E) D_{2D}(E) \\ M_{3D}(E) = A \frac{h}{4} v_x^*(E) D_{3D}(E) \end{cases}$$

$$G = \frac{I}{V} = \int_{-\infty}^{+\infty} dE \left( -\frac{\partial f_0}{\partial E} \right) G(E) \quad R = \rho(L + \lambda) \left[ 1, \frac{1}{W}, \frac{1}{A} \right]$$

$$D(E) v(E) p(E) = N(E) \cdot d$$

Одесса  
 ТЭС  
 2015

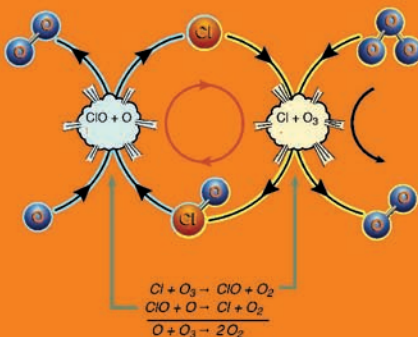
Ю. А. Кругляк  
 КВАНТОВАЯ ХИМИЯ  
 Киев: 1963 – 1991



Одесса  
 ТЭС  
 2016

Ю. А. Кругляк

Прогнозирование свойств молекулярных веществ  
 Критические свойства фреонов и их Si- и Ge-аналогов  
 Озоноразрушающие фреоны и механизмы истощения  
 озонового слоя Земли



Одесса  
 ТЭС  
 2018

