Ю.А. Кругляк

НАНОЭЛЕКТРОНИКА «СНИЗУ — ВВЕРХ»

$$I=\frac{2q}{h}\int T(E)M(E)(f_1(E;E_{F1},T_1)-f_2(E;E_{F2},T_2))dE$$

$$f_1(E)=\frac{1}{e^{(E-E_{F1})/kT_1}+1} \qquad f_2(E)=\frac{1}{e^{(E-E_{F2})/kT_2}+1}$$

$$T(E)=\frac{\lambda(E)}{\lambda(E)+L}\Rightarrow\begin{Bmatrix} L\gg\lambda\ ;\ T\ll 1\\ L\approx\lambda\ ;\ T<1\\ L\ll\lambda\ ;\ T\to 1\end{Bmatrix}$$

$$M(E)=\begin{Bmatrix} M_{1D}(E) = \frac{h}{4}\langle v_x^+(E)\rangle D_{1D}(E)\\ WM_{2D}(E) = W\frac{h}{4}\langle v_x^+(E)\rangle D_{2D}(E)\\ AM_{3D}(E) = A\frac{h}{4}\langle v_x^+(E)\rangle D_{3D}(E)\end{Bmatrix}$$

$$G\equiv\frac{I}{V}=\int_{-\infty}^{+\infty}dE\left(-\frac{\partial f_0}{\partial E}\right)G(E) \qquad R=\rho(L+\lambda)\left\{1,\ \frac{1}{W},\ \frac{1}{A}\right\}$$

$$D(E)v(E)p(E)=N(E)\cdot d$$



Ю. А. Кругляк

НАНОЭЛЕКТРОНИКА «СНИЗУ – ВВЕРХ»

Издание 2-ое, дополненное

Киев
Издательство Стрельбицкого
2016

ББК 22.332
УДК 537.32
PACS: 71.15.Mb, 71.20.-b, 73.22.Pr , 73.23.Ad, 84.32.Ff, 85.35.-p

Кругляк, Юрий Алексеевич

Наноэлектроника «снизу – вверх» / Ю. А. Кругляк

Издание 2-ое, дополненное

Киев: Издательство Стрельбицкого, 2016

544 стр.; 277 рис.; 3 табл.; 515 лит.

ISBN 978-617-7337-15-6

Книга посвящена современной наноэлектронике в концепции «снизу – вверх», основы которой были заложены еще Рольфом Ландауэром, а ныне активно развиваемой Суприе Даттой, Марком Лундстромом и другими исследователями. Подробно рассмотрена обобщенная модель транспорта электронов и переноса тепла в микро- и наноэлектронике, одинаково пригодная для 0D, 1D, 2D и 3D резисторов произвольного масштаба, характеризуемых любой наперед заданной дисперсией и работающих в баллистическом, диффузионном и промежуточных режимах. В первой части рассматривается классический транспорт, а во второй части – квантовый, включая начала спинтроники, магнетроники и молекулярной электроники, на основе метода неравновесных функций Грина в матричном представлении вплоть до квантовой природы классики.

Книга предназначена прежде всего для студентов, магистрантов, аспирантов и преподавателей университетов естественно-научного профиля, как физиков, химиков, так и будущих электроинженеров. Чтобы сделать книгу доступной даже студентам младших курсов, пожалуй, достаточно знать лишь основы линейной алгебры и дифференциальных уравнений, а в области физики понадобятся лишь самые начальные сведения из квантовой механики. Вместе с тем в книге используются новые подходы в преподавании весьма сложных разделов общей физики, что позволило включить в книгу новейшие результаты в области нанофизики, в частности, спинтроники, мало известные отечественной университетской аудитории.

ББК 22.332
УДК 537.32
PACS: 71.15.Mb, 71.20.-b, 73.22.Pr , 73.23.Ad, 84.32.Ff, 85.35.-p

ISBN 978-617-7337-15-6

Жене Наталье,
дочери Наташе, сыновьям Алеше и Жене,
внукам Варе, Паше, Ксюше, Антону и Денису,
правнуку Родиону

Предисловие ко второму изданию

Эта книга была написана мною по итогам прослушивания, решения задач, сдачи тестов и экзаменов он-лайн блестящих курсов лекций Суприё Датта (Supriyo Datta) «Fundamentals of Nanoelectronics, Part I: Basic Concepts» и «Fundamentals of Nanoelectronics, Part II: Quantum Models», прочитанных им в 2012 и 2015 годах, лекций Марка Лундстрома (Mark Lundstrom) «Near-Equilibrium Transport: Fundamentals and Applications» и «Fundamentals of Nanotransistors», прочитанных им в 2011 и 2016 годах, лекций Тимоти Фишера (Timothy Fisher) «Thermal Energy at the Nanoscale» и Ал Шакури (Ali Shakouri) «Thermoelectricity: From Atoms to Systems», прочитанных ими в 2013 году в рамках инициативы Purdue University, USA / nanoHUB-U (www.nanohub.org/u/).

Содержание книги доступно и на украинском языке пока в виде серии учебно-образовательных обзоров, написанных совместно с М. В. Стрихой и все еще публикуемых в каждом номере журнала «Сенсорная электроника и микросистемные технологии», начиная с № 4 за 2012 год. Несмотря на наши совместные усилия реальное положение дел с преподаванием в украинских университетах на специализациях физического и инженерного профилей, которое обычно сводится к использованию традиционных, десятилетиями установившихся концепций, обусловливает необходимость более глубокого ознакомления отечественных преподавателей, аспирантов и студентов с современной концепцией изложения физики наносистем «снизу – вверх», от атомов и молекул к атомным и молекулярным кластерам с использованием концепций и вычислительных методов квантовой химии, авторами которой являются Суприё Датта и Марк Лундстром из университета Пердью, в значительной степени альтернативной традиционной концепции преподавания физики электрических явлений и электроники, опирающейся на рассмотрении электрических явлений в массивных и объемных проводниках.

Во втором издании книги добавлен раздел «Анализ электрофизических характеристик транзисторов», литература к главам дополнена ссылками на новые публикации автора, а также исправлено пару замеченных опечаток.

Благодарю свою жену и соавтора совместных статей Наталью Ефимовну Кругляк за неоценимую помощь в работе и в подготовке книги к публикации, особенно за графическое оформление сотен и сотен рисунков на украинском, английском и русском языках.

Из предисловия к первому изданию

У каждого из нас есть мобильный телефон или смартфон или планшет, а многие, особенно студенты и офисные служащие пользуются ноутбуками и стационарными персональными компьютерами. Каждое из этих удивительных устройств сейчас имеет до четырех и более миллиардов нанотранзисторов (нанорезисторов), токопроводящая область которых составляет всего несколько сот атомов. В лабораториях уже созданы нанотранзисторы, работающие на одном атоме или на одной молекуле (от молекулы водорода до практически любой другой). Это – наше будущее. Предстоящие годы ознаменуются постепенным переходом от зарядовой архитектуры компьютеров к спиновой, появятся компьютеры со спиновой структурой хранения и передачи информации, а для этого нужны молекулярные магниты для создания спиновых конденсаторов и устройства для передачи информации от магнита к магниту.

Бурное развитие наноэлектроники в последние 15 – 20 лет привело не только к созданию и широкому использованию нанотранзисторов и других разнообразных наноразмерных устройств электроники, но и к более глубокому пониманию причин возникновения тока, обмена и рассеяния энергии и принципов работы электронных устройств в целом как наноразмерных, так и привычных электронных приборов. В наши дни фактически происходят революционные события в электронике, что влечет за собой необходимость пересмотра содержания университетского физического образования. Похожая революционная ситуация наблюдалась 65 лет назад после открытия транзистора, что привело не только к повсеместному использованию устройств микроэлектроники, но и к коренному пересмотру университетских и инженерных курсов общей физики, не говоря уже о специальных курсах в области электроники и смежных дисциплин.

С времен становления физики твердого тела используемые в электронике материалы характеризовались интегральными свойствами такими, например, как подвижность носителей тока, с дальнейшим их использованием для объяснения наблюдаемых физических явлений и моделирования различных электронных устройств (концепция «сверху – вниз»). С переходом в наши дни к нанотехнологиям нано- и молекулярные транзисторы требуют с самого начала для своего описания и моделирования применения законов квантовой механики и неравновесной статистической термодинамики, что неизбежно приведет к пересмотру физического образования уже на начальных университетских курсах.

Развитие концепции «снизу – вверх» Рольфом Ландауэром, Суприё Датта, Марком Лундстромом и другими привело в конечном итоге к созданию объединенной модели переноса электронов и тепла в электронных устройствах как нано-, так и микро- и макроскопических, произвольной размерности 1D, 2D и 3D и работающих в баллистическом, диффузионном и промежуточных режимах, изложению которой и посвящена настоящая книга.

Содержание

Предисловие ко второму изданию

Из предисловия к первому изданию

I. Обобщенная модель транспорта электронов и переноса тепла в микро- и наноэлектронике

Глава 1. Обобщенная модель электронного транспорта

1.1. Концепция «снизу – вверх» в современной наноэлектронике 15

1.2. Причины возникновения электрического тока 19
1.2.1. Роль фермиевских функций 22
1.2.2. Квазиравновесный транспорт электронов 24

1.3. Модель электронного транспорта Ландауэра – Датты – Лундстрома 27
1.3.1. Число электронов и ток в проводнике 28
1.3.2. Моды проводимости 31
1.3.3. Коэффициент прохождения 36
1.3.4. Режим линейного отклика 39
1.3.5. Транспорт в массивных проводниках 40
1.3.6. Баллистический и диффузионный режимы проводимости 43
1.3.6.1. 2D Баллистические резисторы 44
1.3.6.2. Широкие 2D баллистические резисторы при $T = 0^{o}\ K$ 46
1.3.6.3. Широкие 2D баллистические резисторы при $T > 0^{o}\ K$ 46
1.3.6.4. Интегралы Ферми – Дирака 48
1.3.6.5. 2D диффузионные резисторы 49
1.3.6.6. Широкие 2D диффузионные резисторы при $T = 0^{o}\ K$ 49
1.3.6.7. Широкие 2D диффузионные резисторы при $T > 0^{o}\ K$ 50
1.3.7. О понятии подвижности электронов 51
1.3.8. Способы записи 2D проводимости 54
1.3.9. Диссипация мощности в баллистических резисторах 56
1.3.10. Падение напряжения в баллистических резисторах 57
1.3.11. 1D и 3D резисторы 58
1.3.12. Модель упругого резистора в концепции Ландауэра – Датты 61
1.3.12.1. Диссипация тепла упругим резистором 65
1.3.12.2. Модель упругого резистора и макропроводники 66
1.3.12.3. Баллистический и диффузионный режимы 67

1.3.12.4. О модах проводимости в общем случае 71
1.3.12.5. Фундаментальное соотношение 72
1.3.12.6. Дисперсии $E(p)$ для кристаллических твердых тел 75
1.3.12.7. Еще раз о формуле Друде 77
1.3.12.8. Пропорциональна ли электропроводимость электронной плотности? 80
1.3.12.9. Квантование проводимости и целочисленность мод проводимости 82

Литература 84

Глава 2. Обобщенная модель переноса тепла

2.1. Термоэлектрические явления и устройства 89
2.1.1. Эффект Зеебека 90
2.1.2. Эффект Пельтье 94
2.1.3. Термоэлектрические устройства 98

2.2. Термоэлектрические коэффициенты 104
2.2.1. «Движущие силы» электрического тока 105
2.2.2. Ток в массивных проводниках 109
2.2.3. Перенос тепла электронами 110
2.2.4. Обращенная форма транспортных уравнений 111
2.2.5. Транспортные коэффициенты для 1D проводников 112
2.2.6. Закон Видемана – Франца и числа Лоренца 117
2.2.7. Полупроводники p-типа и биполярная проводимость 118
2.2.8. Транспортные коэффициенты для 3D проводников в диффузионном режиме 120
2.2.9. Коэффициент Зеебека в вырожденном пределе: формула Мотта 127
2.2.10. Анализ экспериментальных данных для Ge: коэффициент Зеебека и термоэлектрические коэффициенты 123
2.2.11. Термоэлектрические коэффициенты для 1D, 2D и 3D проводников с параболической дисперсией 126
2.2.11.1. Термоэлектрические коэффициенты для 1D баллистических резисторов 127
2.2.11.2. Термоэлектрические коэффициенты для 1D диффузионных резисторов 127
2.2.11.3. Термоэлектрические коэффициенты для 2D баллистических резисторов 128

2.2.11.4. Термоэлектрические коэффициенты для 2D диффузионных резисторов 128
2.2.11.5. Термоэлектрические коэффициенты для 3D баллистических резисторов 129
2.2.11.6. Термоэлектрические коэффициенты для 3D диффузионных резисторов 129

2.3. Перенос тепла фононами 130

2.3.1. Элементарная физика фононов 130

2.3.2. Обобщенная модель переноса тепла фононами 133

2.3.2.1. Теплопроводность массивных проводников 137
2.3.2.2. Удельная теплоемкость 138
2.3.2.3. Средняя длина свободного пробега и коэффициент прохождения 139
2.3.2.4. Плотность фононных состояний и фононные моды 142
2.3.2.5. Дебаевская модель 146
2.3.2.6. Рассеяние фононов 148
2.3.2.7. Температурная зависимость теплопроводности 149
2.3.2.8. Различие между решеточной теплопроводностью и электрической проводимостью 150
2.3.2.9. Квантование теплопроводности 150

2.3.3. Термоэлектричество «снизу – вверх» в концепции Ландауэра – Датта
2.3.3.1. Еще раз о коэффициенте Зеебека 154
2.3.3.2. О показателях качества ТЭ устройств и материалов 156
2.3.3.3. Перенос тепла электронами в концепции «снизу – вверх» 157
2.3.3.4. Квазиравновесный режим 159
2.3.3.5. Дельта-термоэлектрик 160
2.3.3.6. Оптимизация силового коэффициента 163
2.3.3.7. О транспорте тепла фононами в концепции «снизу – вверх» 165
2.3.3.8. Баллистический транспорт фононов 166
2.3.3.9. Фононная теплопроводность 167

Литература 170

Глава 3. Графен

3.1. Элементарная теория зонной структуры графена 176
3.1.1. Обратная решетка и зона Бриллюэна графена 181
3.2. Плотность состояний и плотность носителей тока 183
3.3. Число мод и проводимость 184
3.4. Рассеяние в графене 185
3.5. Подвижность и формула Друде 187
3.6. Циклотронная частота в графене 187
3.7. Эффективная масса в графене 190
3.8. Плотность фононных состояний в графене 192
3.9. Вклад электронов и фононов в теплопроводность графена 193
3.10. Максимальная электрическая проводимость графена 195
3.11. Термоэлектрические коэффициенты графена 196
3.11.1. Баллистический режим 196
3.11.2. Диффузионный режим 197

Литература 198

Глава 4. Учет рассеяния в обобщенной транспортной модели

4.1. Качественные аспекты физики рассеяния электронов 201
4.2. Коэффициент прохождения и средняя длина свободного пробега 206
4.3. Средняя длина свободного пробега и время рассеяния 208
4.4. Оценка усредненного значения средней длины свободного пробега из экспериментальных измерений 209
4.5. Оценка длины свободного пробега из коэффициента диффузии 210
4.6. Связь длины свободного пробега с подвижностью 212
4.7. Усредненная длина среднего свободного пробега для степенного закона рассеяния 213
4.8. Подвижность при постоянном значении времени рассеяния 213
4.9. Анализ экспериментальных данных для Si MOSFET 214

Литература 218

Глава 5. Обобщенная модель переноса и транспортное уравнение Больцмана

5.1. Транспортное уравнение Больцмана 219
5.2. Решение уравнения Больцмана в режиме динамического равновесия 223
5.3. Вычисление транспортных коэффициентов 224

5.4. Учет магнитного поля 230

Литература 237

Глава 6. Измерение проводимости и анализ электрофизических характеристик транзисторов

6.1. Методы измерения сопротивления (проводимости) 241

6.2. Измерение эффекта Холла 243
6.2.1. Пример анализа результатов холловского эксперимента 245

6.3. Метод измерений ван дер Пау 246

6.4. Температурные измерения 251
6.4.1. Температурный артефакт при измерении эффекта Холла (эффект Нернста) 253

6.5. Измерения в сильных магнитных полях 255
6.5.1. Осцилляции Шубникова – де Гааза 259

6.6. Анализ электрофизических характеристик транзисторов 260
6.6.1. Вычисление подвижности из проводимости 262
6.6.2. Вычисление средней длины свободного обратного рассеяния из Подвижности 263
6.6.3. Учет рассеяния по степенному закону 265
6.6.4. К вопросу обоснования применимости правила Матиссена 266

Литература 268

Глава 7. Роль электростатики и контактов

7.1. Диффузия и дрейф 272

7.2. Уравнение диффузии для баллистического транспорта 273

7.3. Электрохимические потенциалы вдали от равновесия 279

7.4. Токи в режиме неравновесных потенциалов 282

7.5. Сопротивление R_{int} на границе контакта двух проводников с разным числом мод 283

7.6. Уравнение диффузии как следствие уравнения Больцмана 285

7.7. Эффекты при распространении тока по квантовой проволоке 290

7.8. Роль электростатики 295
7.8.1. Полевой транзистор 296
7.8.2. Ток стремится к насыщению 298
7.8.3. Роль заряжания проводника 300
7.8.3.1. Электростатический «выпрямитель» 303
7.8.4. Расширенная модель проводника 305

7.9. Контакты 308
7.9.1. $p - n$ Переходы 309
7.9.2. Вольт-амперные характеристики 311

Литература 315

Глава 8. Термодинамика проводника с током

8.1. Термодинамика проводника с током с учетом окружающей среды 317
8.2. Асимметрия процессов поглощения и излучения энергии 319
8.3. Энтропия 322
8.3.1. Энтропия всегда растет 324
8.3.2. Свободная энергия всегда уменьшается 325

8.4. Универсальный закон равновесия Больцмана 326

8.5. Состояния в пространстве Фока 328
8.5.1. Фермионы 328
8.5.2. Бозоны 329
8.5.3. Взаимодействующие электроны 330

8.6. Альтернативное выражение для энтропии 331
8.6.1. Равновесное распределение из минимума свободной энергии 333

8.7. Глобальная ценность информации 334
8.8. Информационно-управляемый аккумулятор 337
8.8.1. Важно знать детальную информацию 340

8.9. Принцип Ландауэра о минимуме энергии, необходимом для стирания одного бита информации 341

8.10. Демон Максвелла 342

Литература 344

II. Квантовый транспорт, спинтроника, магнетроника, молекулярная электроника и смежные вопросы

Глава 9. Спинтроника и магнетроника

9.1. Введение 347

9.2. Граничное сопротивление и несовпадение мод проводимости 351

9.3. Спиновые потенциалы 354
9.3.1. Разность нелокальных спин-потенциалов 357

9.4. Спиновый момент 360

9.5. Уравнение Ландау – Лифшица – Гильберта 361
9.5.1. Выделенная ось магнита 363

9.6. Обращение намагниченности спиновым током 365

9.7. Поляризаторы и анализаторы спинового тока 366

Литература 368

Глава 10. Метод неравновесных функций Грина

10.1. Квантовый транспорт 371

10.2. Уравнения метода НРФГ в матричном представлении 374

10.3. Одноуровневый резистор: полуклассический подход 377

10.4. Одноуровневый резистор: квантовый подход 379

10.5. Квантовое уширение 382

10.6. Интерференция источников поступления электронов в резистор 383

10.7. Квантовый транспорт по многоуровневому проводнику 384

10.8. Функция проводимости для когерентного транспорта 389

10.9. Модельные транспортные задачи 391
10.9.1. Моделирование 1D проводника 391
10.9.1.1. Баллистический 1D проводник 395
10.9.1.2. Плотность состояний 1D проводника 396
10.9.1.3. 1D проводник с одним рассеивающим центром 397

10.9.2. Моделирование однородного 2D проводника 400
10.9.2.1. Контактные собственные энергии для 2D проводника 406
10.9.3. Графен как пример неоднородного 2D проводника 408
10.9.4. Общий подход к построению контактных матриц 412
10.9.5. Баллистическая проводимость графена 416

10.10. Дефазировка 417
10.10.1. Упругая дефазировка 419
10.10.2. 1D проводник с двумя и более рассеивающими центрами 421
10.10.3. Скачок потенциала на дефектах 425
Литература 429

Глава 11. Эффекты Холла, измерение электрохимических потенциалов, транспорт спинов в модели НРФГ и квантовая природа классики

11.1. Эффекты Холла 433

11.2. Почему отличаются между собой *n*- и *p*-проводники? 437

11.3. Пространственный профиль электрохимического потенциала 438

11.4. Измерение электрохимических потенциалов 442
11.4.1. Формулы Ландауэра 446
11.4.2. Формула Бюттекера 449
11.4.2.1. Не тривиально ли уравнение Бюттекера? 453

11.5. Измерение потенциала Холла 454

11.6. Квантовый эффект Холла и краевые состояния в графене 457

11.7. Формула Кубо 465
11.7.1. Формула Кубо для упругого резистора 468

11.8. Невзаимные электрические цепи 471
11.8.1. Соотношения Онзагера 472

11.9. Спинорное представление 474
11.9.1. Одноуровневый спиновый вентиль 478
11.9.2. Вращение магнитных контактов 482

11.10. Спиновые гамильтонианы 484
11.10.1. Гамильтониан с зеемановским расщеплением 484
11.10.2. Гамильтониан Рашбы 485

11.11. Спиноры и векторы 486
11.12. Прецессия спина 490
11.13. Квантовый спиновый эффект Холла 493

11.14. Метод НРФГ и диффузионный транспорт 494
11.14.1. Электронная плотность в матричном представлении 495
11.14.2. Измерение спинового потенциала 496
11.14.3. Транспорт спинов в четырехкомпонентном формате 497

11.15. Квантовая природа классики 500
11.15.1. Спиновая когерентность 500
11.15.2. Псевдоспин 502
11.15.3. Квантовая энтропия 504
11.15.4. Сколько информации переносит один электрон? 505
11.15.5. Увеличивает ли энтропию взаимодействие частиц? 506
11.15.6. Спины и магниты 507

Литература 508

Глава 12. Кулоновская блокада и одноэлектронный нанотранзистор на молекуле бензола

12.1. За пределами метода НРФГ 513

12.2. Кулоновская блокада в одноэлектронике 514
12.2.1. Вольт-амперные характеристики 517
12.2.2. Моделирование заряжания проводника в пространстве Фока 518
12.2.3. Вычисление тока в фоковском пространстве 521

12.3. Одноэлектронный транзистор на молекуле бензола 523
12.3.1. Элементарная теория SET 525
12.3.2. Условия переноса электрона с истока на сток 528
12.3.3. Формализм теории функционала плотности 529
12.3.3.1. Учет внешнего электростатического поля 531
12.3.4. Результаты расчетов 533

Литература 540

I. Обобщенная модель транспорта электронов и тепла в микро- и наноэлектронике

Глава 1. Обобщенная модель электронного транспорта

1.1. Концепция «снизу – вверх» в современной наноэлектронике

Бурное развитие наноэлектроники в последние 15 – 20 лет привело не только к созданию и широкому использованию нанотранзисторов и других разнообразных наноразмерных устройств электроники, но и к более глубокому пониманию причин возникновения тока, обмена и диссипации энергии и принципов работы устройств в целом как наноразмерных, так и привычных электронных приборов [1 – 4]. В наши дни фактически происходят революционные изменения в электронике, что влечет за собой необходимость пересмотра содержания университетского физического образования. Похожая революционная ситуация наблюдалась 70 лет назад после открытия транзистора, что привело не только к повсеместному использованию устройств микроэлектроники, но и к коренному пересмотру университетских и инженерных курсов общей физики, не говоря уже о специальных курсах в области электроники и смежных дисциплин. С времен становления физики твердого тела используемые в электронике материалы и вещества характеризовались интегральными свойствами такими, например, как подвижность носителей тока или коэффициент оптического поглощения, с дальнейшим их использованием для объяснения наблюдаемых физических явлений и моделирования различных электронных устройств. С переходом в наши дни к мезо- и наноскопике нано- и молекулярные транзисторы требуют с самого начала для своего описания и моделирования применения законов квантовой механики и неравновесной статистической термодинамики, что неизбежно приведет к пересмотру физического образования уже на начальных университетских курсах.

Феноменологическая модель электронной проводимости Друде [5, 6] и поныне занимает центральное место и в преподавании и в экспериментальных измерениях электрических характеристик различных материалов [7]. В качестве причины, вызывающей электрический ток, принимается внешнее электрическое поле

$$\mathrm{E}_x = -\frac{dV}{dx} = \frac{V}{L}, \qquad (1)$$

возникающее как градиент разности потенциалов V, подаваемой на проводник (рис. 1).

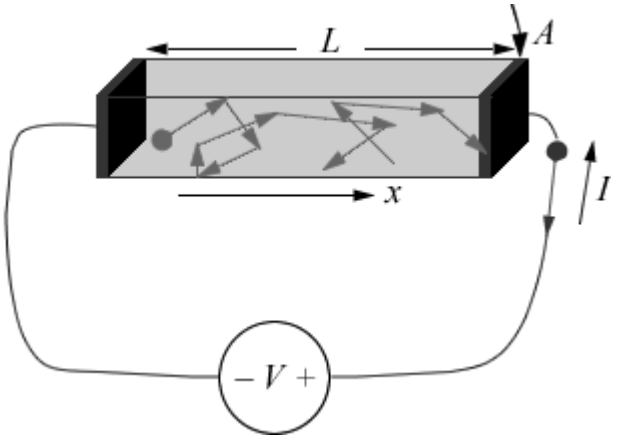


Рис. 1. Электронный транспорт в режиме диффузии через проводник длины *L* и площади поперечного сечения *A*. Используется традиционное соглашение о знаках: ток *I* и направление движения электронов противоположны.

Электрон испытывает действие силы

$$F_e = -q\mathrm{E}_x \tag{2}$$

со стороны электрического поля и движется с дрейфовой скоростью

$$v_d = -\mu \mathrm{E}_x, \tag{3}$$

характеризуемой подвижностью μ.

Ток равен суммарному заряду

$$Q = nqA \cdot L, \tag{4}$$

где n – объемная плотность электронов, перенесенная за время τ,

$$I = Q/\tau, \tag{5}$$

которое в свою очередь определяется скоростью дрейфа

$$\tau = L/v_d. \tag{6}$$

Для тока окончательно имеем

$$I = nqv_d A = nq\mu \mathrm{E}_x A \tag{7}$$

и

$$I = nq\mu \frac{A}{L} V, \tag{8}$$

а для проводимости *G* и ее удельного значения σ имеем:

$$G = nq\mu \frac{A}{L} = \sigma \frac{A}{L} \tag{9}$$

и

$$\sigma = nq\mu. \tag{10}$$

Согласно закону Ома, сопротивление

$$R \equiv V / I = \rho L / A, \qquad (11)$$

где удельное сопротивление ρ, как и обратная ему удельная проводимость σ в выражении для проводимости (9) не зависят от геометрии проводника и являются свойствами материала, из которого изготовлен проводник. Закон Ома утверждает, что если уменьшить длину проводника в несколько раз, то в это же число раз уменьшится его сопротивление. А если уменьшить длину канала проводимости до очень маленьких размеров, то сопротивление практически «занулится»?

При обычном «диффузионном» движении электронов по проводнику среднее значение длины свободного пробега в полупроводниках меньше 1 *мкм* и изменяется в широких пределах в ту и в другую стороны в зависимости от температуры и природы материала, из которого сделан проводник. Длина канала проводимости в современных полевых транзисторах $\sim 30-40\,нм$, что соответствует нескольким сотням атомов. Вполне уместно задаться вопросом: а если длина проводника меньше диффузионной длины свободного пробега, то движение электрона станет баллистическим? Будет ли сопротивление подчиняться закону Ома? А если уменьшить длину канала проводимости до нескольких атомов? Имеет ли смысл в этом случае говорить о сопротивлении как таковом? Все эти вопросы были предметом жарких дискуссий еще лет 15 – 20 назад. Сегодня ответы на эти вопросы даны и надежно подкреплены многочисленными экспериментальными данными. Измерено даже сопротивление молекулы водорода [8] и создан полевой транзистор на одиночном атоме фосфора [9].

Сейчас надежно экспериментально установлено, что сопротивление баллистического резистора

$$R^{ball} = \frac{h}{q^2}\frac{1}{M},$$

где фундаментальная константа $h/q^2 = 25.8128\,кОм$ и носит название постоянной Клитцинга в честь ученого, открывшего целочисленный квантовый эффект Холла и впервые доказавшего возможность квантования сопротивления, а целочисленная величина M есть число мод (каналов) проводимости резистора. Этот результат хорошо известен, однако, обычно считают, что он имеет отношение только к ультракоротким резисторам и является предметом обсуждения в спецкурсах по мезофизике и наноэлектронике. Что не очень широко известно, так это то, что сопротивление и для длинных и для коротких резисторов может быть записано в виде

$$R = \frac{h}{q^2}\frac{1}{M}\left(1+\frac{L}{\lambda}\right),$$

где λ есть среднее значение длины свободного пробега в проводнике.

Баллистические и диффузионные проводники – это не два разных мира, а один и тот же объект, плавно превращающийся из баллистического резистора в диффузионный проводник по мере того как его длина увеличивается. Баллистический предел очевиден при $L \ll \lambda$, тогда как при $L \gg \lambda$ получаем стандартную запись закона Ома (11). Действительно, сопротивление можно записать в виде

$$R = \frac{\rho}{A}(L + \lambda), \tag{12}$$

тогда для удельного сопротивления получим выражение

$$\rho = \frac{h}{q^2}\left(\frac{1}{M/A}\right)\frac{1}{\lambda},$$

позволяющее по новому интерпретировать сопротивление через удельное число мод проводимости и среднюю длину свободного пробега. Выражение (12) справедливо и для длинных массивных проводников и для ультракоротких нанорезисторов и известно как обобщенный закон Ома, а вытекающие из него следствия будут еще предметом обсуждения в этой книге.

Обращает на себя внимание то обстоятельство, что впечатляющие успехи экспериментальной наноэлектроники практически не повлияли на то как мы думаем, обучаем и объясняем понятия сопротивления, проводимости и работу электронных устройств в целом. И поныне, по-видимому, по историческим причинам доминирует привычная концепция «сверху – вниз», от массивных проводников до атомов и молекул. Такой подход был вполне приемлем до тех пор, пока не было достаточного массива экспериментальных данных по измерению проводимости наноразмерных проводников. В последнее десятилетие ситуация изменилась. Накоплены обширные экспериментальные данные и для больших и для предельно малых проводников. Началась разработка концепции проводимости «снизу – вверх» [10 – 13], которая не только оказалась комплементарной концепции «сверху – вниз», но и привела к переосмысливанию принципов работы обычных электронных устройств. Вспомним, что в квантовой механике с самого начала доминировала концепция «снизу – вверх»: от атома водорода в направлении твердого тела.

Есть еще один круг задач в наноэлектронике, для решения которых концепция «снизу – вверх» особенно привлекательна. Это – транспортные задачи. В обычной электронике транспорт частиц описывается законами механики – классической или квантовой. Транспорт по массивному проводнику сопровождается выделением тепла, что описывается законами термодинамики – обычной или статистической. Процессы в механике обратимы, а в термодинамике необратимы. Разделить эти два процесса – движения и выделения тепла – строго говоря, невозможно. Совсем другая ситуация в наноэлектронике. Здесь процессы движения электронов и выделения тепла

пространственно разделены: электроны движутся упруго, баллистически («упругий резистор»), а выделение тепла происходит лишь на границах контакта проводника с электродами. Концепция «упругого резистора» была предложена Ландауэром в 1957 году [14–16] задолго до ее экспериментального подтверждения в нанотранзисторах. Концепция «упругого резистора», строго говоря, является идеализацией, но она надежно подтверждена многочисленными экспериментальными данными для ультрамалых нанотранзисторов. Развитие концепции «снизу – вверх» Рольфом Ландауэром, Суприе Даттой и Марком Лундстромом (ЛДЛ) привело в конечном итоге к созданию обобщенной модели транспортных явлений в электронных устройствах как наноскопических, так и микро- и макроскопических, произвольной размерности 1D, 2D и 3D и работающих в баллистическом, квазибаллистическом или диффузионном режимах, изложению которой и посвящено настоящая глава, как и книга в целом. В своем изложении мы будем ближе всего следовать лекциям Марка Лундстрома «Near-Equilibrium Transport: Fundamentals and Applications» [11], прочитанных им в 2011 году, и его лекциям «Nanotechnology: Fundamentals of Nanotransistors», прочитанных в 2016 году, а также лекциям Суприе Датты «Fundamentals of Nanoelectronics, Part 1: Basic Concepts» и «Fundamentals of Nanoelectronics, Part 2: Quantum Transport», прочитанных им в 2012 и в 2015 годах в рамках инициативы Purdue University / nanoHUB-U (www.nanohub.org/u) [11,12, 19].

Начнем мы с казалось бы очевидного вопроса о причинах возникновения электрического тока.

1.2. Причины возникновения электрического тока

На вопрос о причине возникновения тока при приложении разности потенциалов к концам проводника обычно ссылаются на связь плотности тока $\vec{J}$ с внешним приложенным электрическим полем $\vec{\mathrm{E}}$

$$\vec{J} = \sigma \vec{\mathrm{E}},$$

что следует из (7); другими словами, причиной возникновения тока обычно считают электрическое поле. Ответ, в лучшем случае, не полный. Еще до подключения проводника к клеммам источника напряжения на электроны проводника действуют сильные электрические поля, создаваемые ядрами атомов, а ток тем не менее не возникает. Почему сильные внутренние электрические поля не вызывают движение электронов, а намного более слабое внешнее электрическое поле батареи вызывает движение электронов? Обычно говорят, что внутренние микроскопические поля якобы не могут вызвать

движение электронов, необходимо приложить внешнее макроскопическое поле. Трудно признать такое объяснение удовлетворительным. В современных экспериментах по измерению проводимости отдельных молекул невозможно с определенностью вычленить отдельно внутренние и внешние электрические поля. Приходится считаться с этим уроком, преподнесенным нам современной экспериментальной наноэлектроникой, и заново задаться вопросом почему же движутся электроны при подключении батареи к концам проводника.

Для ответа на вопрос о причине возникновения тока нам с самого начала потребуются два понятия – плотности состояний проводника на единицу энергии $D(E)$, занятых электронами и свободных и электрохимического потенциала E_{F0} (рис. 2).

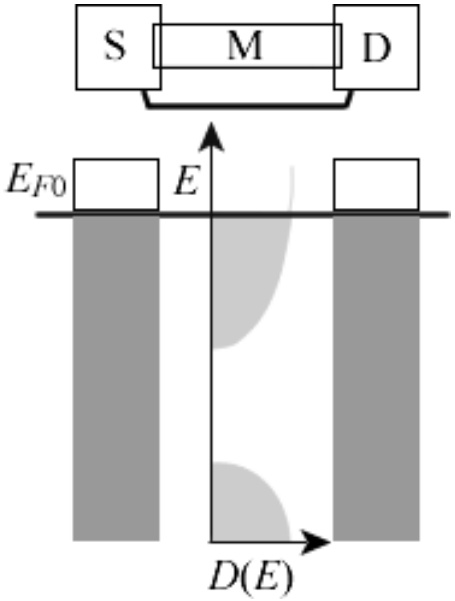


Рис. 2. Первым шагом в объяснении работы электронного устройства должно быть задание плотности состояний $D(E)$ внутри проводника M и определение равновесного значения электрохимического потенциала E_{F0}, отделяющего заполненные электронами состояния от пустых состояний.

Простоты ради, что никак не скажется на окончательных выводах, будем пользоваться так называемой точечной моделью проводника, которая предполагает неизменность плотности состояний $D(E)$ при движении вдоль проводника. Если система, включающая истоковый электрод (S/Source), проводник M и стоковый электрод (D/Drain), находится в равновесии (закорочена), то электрохимический потенциал E_{F0} везде одинаков и все состояния с $E < E_{F0}$ заполнены электронами, а состояния с $E > E_{F0}$ пустые (рис. 2).

При включении в цепь источника напряжения (рис. 3) разность потенциалов V понижает все энергии на положительном электроде D на

величину qV, где q – заряд электрона, в результате чего на электродах создается разность электрохимических потенциалов

$$E_{F1} - E_{F2} = qV\,. \tag{13}$$

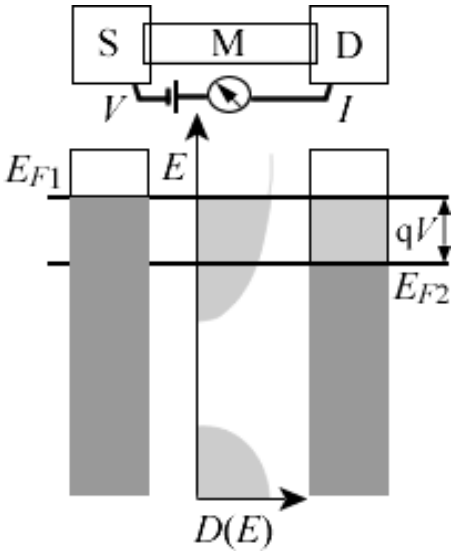


Рис. 3. При подаче напряжения V на клеммы проводника потенциал анода D понижается на величину qV, создавая разность электрохимических потенциалов на концах проводника.

Точно так же как разность температур приводит к потоку тепла, а различие в уровнях жидкости ведет к ее перетоку, так и разность электрохимических потенциалов является причиной возникновения тока. Только состояния проводника в окне $E_{F1} - E_{F2}$ и находящиеся достаточно близко к значениям E_{F1} и E_{F2} дают вклад в поток электронов, тогда как все состояния, лежащие значительно выше E_{F1} и ниже E_{F2}, не играют никакой роли. Причина этого кроется в следующем.

Каждый контакт стремится привести токовый канал в равновесие с собой путем заполнения электронами всех состояний канала с энергией, меньшей электрохимического потенциала E_{F1}, и опорожнения состояний канала с энергией, большей потенциала E_{F2}. Рассмотрим токовый канал с состояниями, энергия которых меньше E_{F1}, но больше E_{F2}. Контакт 1 стремится заполнить эти состояния, поскольку их энергия меньше E_{F1}, а контакт 2 стремится опорожнить эти состояния, поскольку их энергия больше E_{F2}, что и приводит к непрерывному движению электронов от контакта 1 к контакту 2.

Рассмотрим теперь состояния канала с энергией, большей E_{F1} и E_{F2}. Оба контакта стремятся опорожнить эти состояния, но они и так пустые и не дают вклада в электрический ток. Аналогична ситуация с состояниями, энергия которых одновременно меньше обоих потенциалов E_{F1} и E_{F2}. Каждый из контактов стремится заполнить их электронами, но они уже заполнены, и

вклада в ток дать не могут, а точнее не могут в пределах нескольких kT от ширины окна, в чем мы убедимся позже.

Подобная картина выглядит почти очевидной, если бы не привычное утверждение о том, что электроны движутся под действием электрического поля внутри проводника. Если бы это было так, то вклад в ток должны были бы дать все электроны, а не только те, энергия состояний которых лежит в пределах разности потенциалов на концах проводника.

1.2.1. Роль фермиевских функций

Итак, утверждалось, что в равновесии все состояния с энергией $E < E_{F0}$ заполнены электронами, а состояния с энергией $E > E_{F0}$ пустые (рис. 2). Это справедливо только в пределе абсолютного нуля температуры. Более точно, переход от полностью заполненных состояний к пустым совершается в зазоре $\sim \pm 2\,kT$, охватывающем значение $E = E_{F0}$, где k – постоянная Больцмана, T – абсолютная температура. Математически этот переход описывается функцией Ферми

$$f\left(E\right) = \frac{1}{e^{(E-E_F)/kT} + 1}. \tag{14}$$

График функции Ферми показан на рис. 4 слева, возможно, в несколько непривычном виде с энергией в безразмерных единицах по вертикальной оси, что позволит нам позже при объяснении причины токообразования совместить функцию Ферми с графиком плотности состояний $D(E)$.

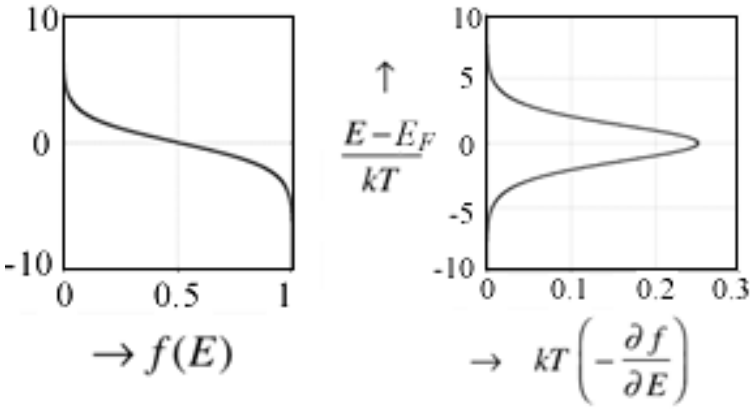


Рис. 4. Графики функции Ферми и нормированной функции теплового уширения.

Функция Ферми играет ключевую роль в статистической механике, однако же для наших целей достаточно понимать, что состояния с низкой энергией всегда заняты ($f = 1$), тогда как состояния с высокой энергией всегда пустые ($f = 0$), а переход от $f = 1$ к $f = 0$ происходит в узком интервале энергии $\sim \pm 2\,kT$, охватывающем значение $E = E_{F0}$.

Действительно, на рис. 4 показана производная от функции Ферми, помноженная на kT с тем, чтобы сделать ее безразмерной:

$$F_T\left(E,E_F\right)=kT\left(-\frac{\partial f}{\partial E}\right). \tag{15}$$

Подставляя выражение для фермиевской функции (14), видно, что

$$F_T\left(E,E_F\right)=\frac{e^x}{\left(e^x+1\right)^2}, \tag{16}$$

где $x\equiv (E-E_F)/kT$. Из (16) видно, что

$$F_T\left(E,E_F\right)=F_T\left(E-E_F\right)=F_T\left(E_F-E\right), \tag{17}$$

а из уравнений (16) и (14) следует, что

$$F_T=f\left(1-f\right). \tag{18}$$

Интегрирование функции (18) во всем интервале изменения энергии дает площадь, равную kT:

$$\int_{-\infty}^{+\infty}dEF_T\left(E,E_F\right)=kT\int_{-\infty}^{+\infty}dE\left(-\frac{\partial f}{\partial E}\right)=kT\left[-f\right]_{-\infty}^{+\infty}=kT\left(1-0\right)=kT. \tag{19}$$

Так что функцию F_T можно приблизительно представить себе в виде прямоугольного «импульса», центрированного относительно значения $E=E_F$ с высотой, равной ¼ и шириной $\sim 4\,kT$. Функция $(-\partial f/\partial E)$ получила название фермиевского окна проводимости.

Когда рассматриваемая система находится в равновесии (рис. 2), электроны распределяются по имеющимся состояниям в соответствии с функцией Ферми. При выходе из равновесия нет простых правил для вычисления функции распределения электронов. Все зависит от конкретной задачи, решать которую нужно методами неравновесной статистической механики.

В нашем специальном случае выхода из равновесия (рис. 3) можно надежно утверждать, что оба контакта S и D настолько велики по сравнению с каналом переноса электронов, что они не могут выйти из равновесия. Каждый из контактов локально находится в равновесии со своим собственным электрохимическим потенциалом, порождая две фермиевских функции (рис. 5):

$$f_1\left(E\right)=\frac{1}{e^{(E-E_{F1})/kT}+1} \tag{20}$$

и

$$f_2(E)=\frac{1}{e^{(E-E_{F2})/kT}+1}. \tag{21}$$

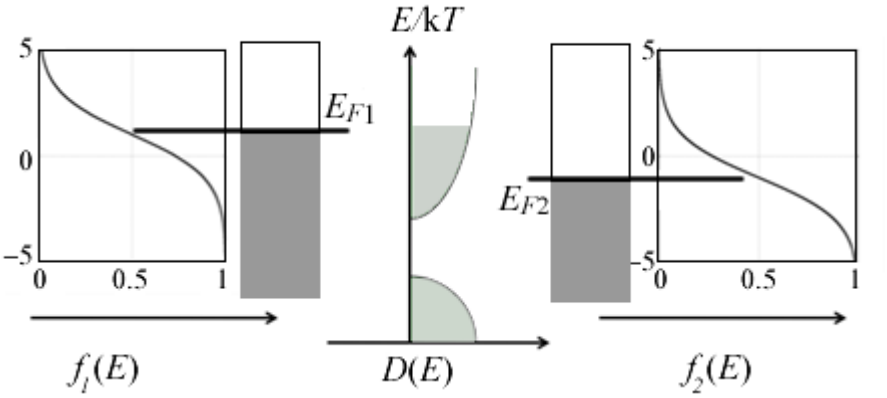


Рис. 5. При выходе из равновесия электроны в контактах занимают доступные им состояния в соответствии с фермиевскими распределениями и значениями электрохимических потенциалов.

Подводя итоги, утверждается, что *причиной появления тока является различие в подготовке равновесных состояний контактов, отображаемое их соответствующими фермиевскими функциями* $f_1(E)$ *и* $f_2(E)$. Качественно это справедливо для любых проводников – и наноразмерных и макроскопических. Однако, для наноразмерных проводников ток при любых значениях энергии электронных состояний в проводнике, как будет показано позже, пропорционален разности фермиевских распределений в обоих контактах: $I(E)\sim f_1(E)-f_2(E)$. Эта разность зануляется, если энергия E больше E_{F1} и E_{F2}, поскольку в этом случае обе функции Ферми равны нулю. Эта разность также зануляется, если энергия E меньше и E_{F1} и E_{F2},поскольку в этом случае обе фермиевские функции равны единице. Ток возникает лишь в окне $E_{F1}-E_{F2}$, если в этом окне есть хотя бы одно электронное состояние проводника.

1.2.2. Квазиравновесный транспорт электронов

Вольт-амперная характеристика обычно нелинейная, но из нее можно вычленить участок «линейного отклика», под которым подразумевается проводимость dI/dV при $V\to 0$.

Построим функцию разности двух фермиевских функций, нормированную к приложенному напряжению,

$$F(E)=\frac{f_1(E)-f_2(E)}{qV/kT}, \tag{22}$$

где

$$\begin{aligned} E_{F1}&=E_{F0}+(qV/2),\\ E_{F2}&=E_{F0}-(qV/2). \end{aligned} \tag{23}$$

Функция разности $F(E)$ сужается по мере того, как напряжение V, помноженное на заряд электрона, становится все меньше величины kT (рис. 6).

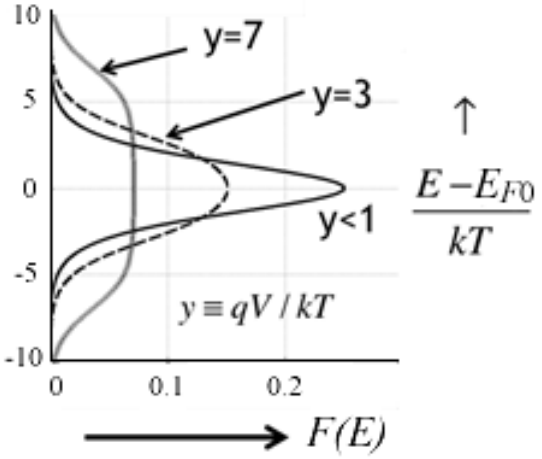


Рис. 6. График функции разности $F(E)$ в зависимости от значения $(E - E_{F0})/kT$ для различных значений $qV/kT \equiv y$.

Отметим также, что по мере того, как kT начинает превосходить энергию qV, функция $F(E)$ все больше приближается к функции теплового уширения:

$$F(E) \to F_T(E) \text{ при } qV/kT \to 0,$$

так что из ур-я (22) следует, что

$$f_1(E) - f_2(E) = \frac{qV}{kT} F_T(E, E_{F0}) = \left(-\frac{\partial f_0}{\partial E} \right) qV, \tag{24}$$

если приложенное напряжение, помноженное на заряд электрона, $qV = E_{F1} - E_{F2}$ становится намного меньшим kT.

Нам потребуется также следующее выражение

$$f(E) - f_0(E) = -\frac{\partial f_0}{\partial E}(E_F - E_{F0}), \tag{25}$$

которое, как и уравнение (24), можно получить следующим образом.

Для функции Ферми

$$f(x) \equiv \frac{1}{e^x + 1}, \quad x \equiv \frac{E - E_F}{kT} \tag{26}$$

имеем

$$\begin{aligned} \frac{\partial f}{\partial E} &= \frac{df}{dx}\frac{\partial x}{\partial E} = \frac{df}{dx}\frac{1}{kT}, \\ \frac{\partial f}{\partial E_F} &= \frac{df}{dx}\frac{\partial x}{\partial E_F} = -\frac{df}{dx}\frac{1}{kT}, \\ \frac{\partial f}{\partial T} &= \frac{df}{dx}\frac{\partial x}{\partial T} = -\frac{df}{dx}\frac{E - E_F}{kT^2}, \end{aligned} \tag{27}$$

откуда

$$\frac{\partial f}{\partial E_F} = -\frac{\partial f}{\partial E},$$
$$\frac{\partial f}{\partial T} = -\frac{E - E_F}{T}\frac{\partial f}{\partial E}. \qquad (28)$$

Уравнениее (25) получается из разложения функции Ферми вблизи точки равновесия в ряд Тейлора

$$f\left(E, E_F\right) \cong f\left(E, E_{F0}\right) + \left(\frac{\partial f}{\partial E_F}\right)_{E_F = E_{F0}} \left(E_F - E_{F0}\right). \qquad (29)$$

Из уравнения (28) следует

$$\left(\frac{\partial f}{\partial E_F}\right)_{E_F = E_{F0}} = \left(-\frac{\partial f}{\partial E}\right)_{E_F = E_{F0}}. \qquad (30)$$

Пусть $f(E)$ соответствует $f(E, E_F)$, а $f_0(E)$ соответствует $f(E, E_{F0})$, тогда

$$f(E) \approx f_0(E) + (E_F - E_{F0})\left(-\frac{\partial f}{\partial E}\right), \qquad (31)$$

что после перегруппировки дает искомое уравнение (25), которое справедливо при $E_F - E_{F0} << kT$.

Подведем итоги. Проводимость веществ может меняться более чем в 10^{20} раз, переходя, например, от серебра до стекла – веществ, весьма удаленных друг от друга в шкале проводимости. Стандартное объяснение различия в их проводимости якобы состоит в том, что плотность «свободных электронов» в этих веществах существенно отличается. Подобное объяснение немедленно требует пояснения, какие электроны свободные, а какие нет. Это различие становится все более абсурдным по мере перехода к наноразмерным проводникам.

Концепция «снизу – вверх» предлагает следующий простой ответ. Проводимость зависит от плотности состояний в окне шириной несколько kT, охватывающих равновесный электрохимический потенциал E_{F0}, определяемый функцией F_T (ф-ла (15), рис. 4), которая отлична от нуля в небольшом промежутке шириной несколько kT вокруг равновесного значения электрохимического потенциала.

Дело не в суммарном числе электронов, которое одного порядка как в серебре, так и в стекле. Ключевым моментом является наличие электронных состояний в области значений электрохимического потенциала E_{F0} (фермиевское окно проводимости), что в корне отличает одно вещество от другого.

Настоящий ответ нельзя назвать новым, и он хорошо известен специалистам в области микро- и наноэлектроники. Тем не менее обсуждение проводимости и поныне обычно начинается с теории Друде [7], которая сыграла важнейшую историческую роль в понимании природы электрического тока. К сожалению, подход Друде породил два недоразумения, которые следовало бы преодолеть, а именно:

(1) ток порождается электрическим полем и

(2) ток зависит от числа электронов.

Оба недоразумения связаны друг с другом, поскольку если бы ток действительно порождался бы электрическим полем, то все электроны были бы подвержены влиянию поля.

Уроки, преподнесенные нам экспериментальной наноэлектроникой, показывают, что ток порождается «подготовкой» двух контактов $f_1(E) - f_2(E)$, и эта разница не нулевая только в фермиевском окне проводимости вокруг равновесного электрохимического потенциала E_{F0}. Проводимость канала высокая или низкая зависит только от наличия электронных состояний в этом окне. К этому выводу обычно приходят на основе транспортного уравнения Больцмана [22] или формализма Кубо [23], тогда как предлагаемая концепция «снизу – вверх» сразу дает корректную картину возникновения тока.

1.3. Модель электронного транспорта Ландауэра – Датты – Лундстрома

Мы рассмотрим модель электронного транспорта довольно простую и вместе с тем удивительно полезную в прикладном отношении и нашедшую широкое применение при анализе как электронного, так и фононного транспорта не только в режиме линейного отклика, но и в высоковольтном режиме горячих электронов, нелокального и квантового транспорта, транспорта в неупорядоченных и наноструктурированных материалах как наноразмерных, так и протяженных, в которых проводимость определяется только свойствами проводника. Свое начало она берет в работах Ландауэра [14 – 16], переосмысленных Даттой [12, 17, 18] и далее развитых Лундстромом применительно к самым различным материалам [13, 19].

Центральное место в любом электронном устройстве играет канал проводимости, характеризуемый плотностью состояний $D(E-U)$, где E – энергия состояний проводника, а U – самосогласованный электростатический потенциал затвора, позволяющий смещать состояния резистора вверх или вниз по шкале энергии (рис. 7).

Далее будем рассматривать двухтерминальное устройство $(U = 0)$. Канал проводимости связан с двумя «идеальными» контактами, способными к быстрому восстановлению равновесия в процессе электронного транспорта и характеризуемыми фермиевскими функциями (20) и (21) с соответствующими электрохимическими потенциалами E_{F1} и E_{F2}. При подаче напряжения V электрохимический потенциал E_{F2} понижается относительно потенциала E_{F1} на величину qV.

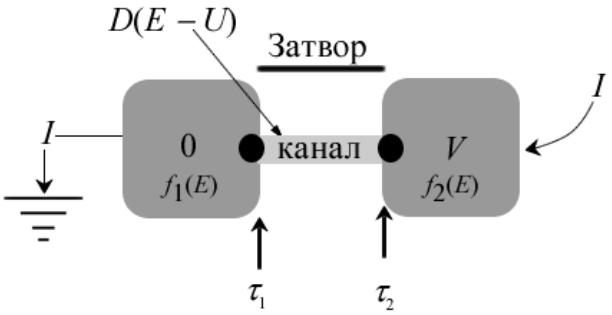


Рис. 7. Типичное наноразмерное электронное устройство с контактами, характеризуемыми временами пролета τ.

Связь контактов с каналом проводимости характеризуется временами пролета τ, дающими представление о том, как быстро электроны могут покинуть контакт или проводник. Для наноразмерных резисторов, например, молекул, времена τ контролируются контактами. Для протяженных резисторов с хорошими контактами, как мы убедимся позже, времена τ становятся сопоставимыми с временами пролета всей протяженности проводника. В общем случае качество обоих контактов может быть различным, что ведет к различным значениям τ_1 и τ_2. Иногда времена τ удобнее выражать в единицах энергии $\gamma = \hbar/\tau$. Если в роли канала проводимости выступает одиночная молекула, величина γ приобретает простой физический смысл уширения энергетических уровней молекулы за счет конечности времени жизни электронов на молекулярных уровнях.

Предполагается, что канал имеет зонную структуру $E(k)$. Это требование, однако, не является обязательным [17]. Обсуждение ограничений и применимости модели ЛДЛ можно найти в [12, 13, 17, 18].

Перейдем к построению математической модели транспорта ЛДЛ в концепции «снизу – вверх». Начать надо с нахождения зависимости плотности электронов и тока в проводнике в зависимости от фермиевских распределений и соответствующих электрохимических потенциалов, плотности состояний проводника и характеристических времен τ пролета.

1.3.1. Число электронов и ток в проводнике

Пусть пока только левый контакт 1 связан с каналом проводимости. Он стремится пополнить проводник электронами в соответствии со своим электрохимическим потенциалом E_{F1}. В конце концов равновесие между контактом и каналом проводимости наступит тогда, когда число электронов с энергиями между E и $E + dE$ станет равным

$$N'_{01}(E)dE = D(E)dE\, f_1(E), \tag{32}$$

где $D(E)$ – плотность состояний с энергией E (пусть каждое состояние заполнено двумя электронами с противоположными спинами), а $f_1(E)$ есть равновесная функция Ферми контакта 1.

Процесс установления равновесия между контактом 1 и каналом описывается простым кинетическим уравнением

$$F_1 = \left.\frac{dN'(E)}{dt}\right|_1 = \frac{N'_{01}(E) - N'(E)}{\tau_1(E)}, \tag{33}$$

где скорость подачи электронов в канал dN'/dt положительна, если число электронов в канале N' меньше его равновесного значения N'_{01}, и отрицательна в противном случае. Если канал первоначально пустой, то он заполняется электронами до их равновесного числа; если же канал переполнен электронами, то он опустошается до тех пор, пока в канале не будет достигнуто равновесное число электронов.

Если канал проводимости связан только с контактом 2, аналогичные рассуждения приводят к следующей паре уравнений:

$$N'_{02}(E)dE = D(E)dEf_2(E), \tag{34}$$

$$F_2 = \left.\frac{dN'(E)}{dt}\right|_2 = \frac{N'_{02}(E) - N'(E)}{\tau_2(E)}. \tag{35}$$

Когда канал проводимости связан одновременно с двумя контактами и оба контакта одновременно пополняют электронами канал и опорожняют его, имеем

$$\left.\frac{dN'(E)}{dt}\right|_{1+2} = F_1 + F_2 = \left.\frac{dN'(E)}{dt}\right|_1 + \left.\frac{dN'(E)}{dt}\right|_2. \tag{36}$$

В состоянии динамического равновесия левые части кинетического уравнения (36) равны нулю. Приравнивая нулю правую часть равенства (36) и подставляя выражения для скоростей на обоих контактах 1 и 2 по уравнениям (33) и (35), для числа электронов в канале проводимости получим:

$$N'(E) = \frac{\left(1/\tau_1\right)}{\left(1/\tau_1\right)+\left(1/\tau_2\right)} N'_{01}(E) + \frac{\left(1/\tau_2\right)}{\left(1/\tau_1\right)+\left(1/\tau_2\right)} N'_{02}(E). \tag{37}$$

Простоты ради, положим $\tau_1 = \tau_2$ и подставим равновесные значения числа электронов N'_{01} и N'_{02} по уравнениям (32) и (34). Тогда для числа электронов в канале в состоянии динамического равновесия в дифференциальной форме получим:

$$N'(E)dE = \frac{D(E)dE}{2} f_1(E) + \frac{D(E)dE}{2} f_2(E). \tag{38}$$

Остается проинтегрировать по всему спектру энергий и получим число электронов в канале проводимости в состоянии динамического равновесия двух

контактов с электрохимическими потенциалами E_{F1} и E_{F2}, разница между которыми пропорциональна напряжению, поданному на концы проводника, а именно:

$$N = \int N'(E)dE = \int \left[\frac{D(E)}{2} f_1(E) + \frac{D(E)}{2} f_2(E) \right] dE . \qquad (39)$$

Получен ответ на первый из поставленных нами двух вопросов: найдено число электронов, или иначе плотности электронов, если известны геометрические параметры проводника, в зависимости от плотности состояний проводника и фермиевских распределений на контактах. Когда электронное устройство (рис. 7) находится в равновесии $(f_1 = f_2 \equiv f_0)$, получаем стандартное выражение для равновесного числа электронов в проводнике [20]

$$N_0 = \int D(E) f_0(E)dE , \qquad (40)$$

в отличие от которого уравнение (39) пригодно как для равновесного состояния устройства (рис. 7), так и для случая, когда электронное устройство выведено далеко за пределы равновесия.

Напомним, что для 1D-, 2D- и 3D-проводников плотность состояния пропорциональна, соответственно, длине L проводника, площади его поперечного сечения A и объему проводника Ω:

$$D(E) \approx \{L, A, \Omega\}, \qquad (41)$$

а плотность электронов определяется, соответственно, как

$$n_L = N/L,\ n_S = N/A,\ n = N/\Omega . \qquad (42)$$

Перейдем к вычислению тока в состоянии динамического равновесия. В этом состоянии один из контактов пополняет канал проводимости электронами, а другой контакт опустошает его. Если $E_{F1} > E_{F2}$, контакт 1 инжектирует электроны, а контакт 2 поглощает их, и наоборот, если $E_{F1} < E_{F2}$.

Скорости, с которыми электроны покидают или поглощаются контактами 1 и 2, даются уравнениями (33) и (35). В состоянии динамического равновесия

$$F_1 + F_2 = 0 . \qquad (43)$$

По договоренности, ток считается положительным, если он входит в контакт 2 (рис. 7), так что

$$I' = qF_1 = -qF_2 . \qquad (44)$$

Выразив ток один раз через F_1, а второй раз через F_2 согласно (44), и сложив эти одинаковые токи, сразу получим, что

$$I' = \frac{q}{2}(F_1 - F_2). \qquad (45)$$

Остается подставить значения для скоростей (33) и (35) в предположении одинаковости контактов $(\tau_1 = \tau_2 \equiv \tau)$, затем воспользоваться (32) и (34), и мы имеем:

$$I'(E) = \frac{q}{2\tau(E)}\left(N'_{01} - N'_{02}\right) = \frac{2q}{h}\frac{\gamma(E)}{2}\pi D(E)(f_1 - f_2), \tag{46}$$

где характеристическое время пролета τ выражено в единицах энергии, а именно:

$$\gamma \equiv \frac{\hbar}{\tau(E)}. \tag{47}$$

Интегрируя (46) по всему спектру энергий, окончательно для тока в состоянии динамического равновесия получаем

$$I = \int I'(E)dE = \frac{2q}{h}\int \gamma(E)\pi\frac{D(E)}{2}\left(f_1 - f_2\right)dE. \tag{48}$$

Согласно (48), ток возникает лишь тогда, когда фермиевские распределения на контактах различны, то ли по причине различия электрохимических потенциалов, то ли по причине различия в температуре контактов, то ли сразу по этим обеим причинам. При различии в электрохимических потенциалах контактов один из контактов стремится передать электроны проводнику, а другой стремится принять их на себя. Ток при различии в температуре контактов рассматривается в главе 2.

Подведем итоги. Получены два основных уравнения транспортной модели ЛДЛ:

$$N = \int \frac{D(E)}{2}\left(f_1 + f_2\right)dE, \tag{49}$$

$$I = \frac{2q}{h}\int \gamma(E)\pi\frac{D(E)}{2}\left(f_1 - f_2\right)dE. \tag{50}$$

Первое из них выражает число электронов в проводнике в состоянии динамического равновесия через плотность состояний проводника и фермиевские функции контактов, а второе уравнение дает ток через эти же характеристики электронного устройства и через характеристическое время пролета. Дальнейшее построение модели ЛДЛ фактически сводится к использованию и развитию этих двух основных уравнений.

1.3.2. Моды проводимости

Как и ожидалось, ток пропорционален разности фермиевских функций контактов. Позже мы убедимся, что комбинация фундаментальных констант $2q/h$ играет важную роль в выражении для тока (50). А каков смысл произведения $\gamma\pi D/2$?

Параметр γ (47) имеет размерность энергии, а размерность плотности состояний D есть (энергия)$^{-1}$, так что интересующее нас произведение $\gamma\pi D/2$

не имеет размерности. В отношении размерности плотности состояний обратим внимание на то, что мы оперируем с числом электронов, а не с плотностями электронов (42), как это обычно принято в физике полупроводников, где плотность состояний для 3D-образца имеет размерность (энергия·объем)$^{-1}$. Мы сейчас увидим, что произведение $\gamma(E)\pi\dfrac{D(E)}{2}=M(E)$ имеет физический смысл числа мод (иначе, каналов) проводимости резистора.

Рассмотрим 2D баллистический резистор длины L, меньшей средней длины свободного пробега λ, вытянутый вдоль оси x и шириной W вдоль оси y. Полная плотность состояний

$$D(E) = D_{2D}(E)L\cdot W, \tag{51}$$

$$D_{2D}(E)=g_v\frac{m^*}{\pi\hbar^2}, \tag{52}$$

где D_{2D} есть удельная плотность состояний (на единицу площади), выписанная для параболической зонной структуры с эффективной массой m^* и долинным вырождением g_v [20].

Определим характеристическое время τ. Из (38) и (46) имеем:

$$\frac{qN'(E)dE}{I'(E)dE}=\frac{\hbar}{\gamma}\frac{\left(f_1+f_2\right)}{\left(f_1-f_2\right)}. \tag{53}$$

Приложим к концам проводника достаточно большое напряжение такое, чтобы имело место $E_{F2}<<E_{F1}$, так что $f_2<<f_1$. Тогда из (53) имеем

$$\frac{qN'(E)dE}{I'(E)dE}=\frac{\text{Накопленный заряд}}{\text{Ток}}=\frac{\hbar}{\gamma}=\tau(E). \tag{54}$$

Число электронов в канале

$$N'(E)=n_S'(E)L\cdot W, \tag{55}$$

где n_S' – удельная электронная плотность (на единицу площади), а ток

$$I'(E)=qWn_S'(E)\left\langle v_x^+(E)\right\rangle, \tag{56}$$

так что из (54) имеем:

$$\tau(E)=\frac{L}{\left\langle v_x^+(E)\right\rangle}, \tag{57}$$

что есть просто среднее время пролета электрона через всю длину упругого резистора.

Для оценки τ нужно вычислить среднюю скорость пролета от контакта 1 до контакта 2 в направлении $+x$ (рис. 1). Эта скорость в случае баллистического транспорта, т.е. без изменения направления движения

$$\left\langle v_x^+(E)\right\rangle = v(E)\left\langle\cos\theta\right\rangle, \tag{58}$$

где угол θ отсчитывается от положительного направления оси x. Поскольку

$$\left\langle\cos\theta\right\rangle = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\cos\theta d\theta = \frac{2}{\pi}, \tag{59}$$

для средней скорости в случае параболического дисперсионного соотношения и изотропии скорости имеем:

$$\left\langle v_x^+(E)\right\rangle = \frac{2}{\pi}v = \frac{2}{\pi}\sqrt{\frac{2\left(E-E_C\right)}{m^*}}. \tag{60}$$

Определяя число мод проводимости как

$$M(E) = \gamma(E)\pi\frac{D(E)}{2} \tag{61}$$

и используя определение γ по (47) и плотности состояний для 2D проводника по (51) для числа мод 2D проводника окончательно получаем

$$M(E) = WM_{2D}(E) = W\frac{h}{4}\left\langle v_x^+(E)\right\rangle D_{2D}(E). \tag{62}$$

Аналогичные соображения для 1D и 3D проводников дают

$$\begin{aligned} M(E) &= M_{1D}(E) = \frac{h}{4}\left\langle v_x^+(E)\right\rangle D_{1D}(E), \\ M(E) &= WM_{2D}(E) = W\frac{h}{4}\left\langle v_x^+(E)\right\rangle D_{2D}(E), \\ M(E) &= AM_{3D}(E) = A\frac{h}{4}\left\langle v_x^+(E)\right\rangle D_{3D}(E). \end{aligned} \tag{63}$$

Обратим внимание, что в случае 2D проводников число мод проводимости пропорционально ширине проводника W, а для 3D проводников – площади их поперечного сечения A.

Заслуживает внимания физический смысл полученных результатов (63). Так, для 2D проводника с учетом (60) и (52) число мод проводимости

$$WM_{2D}(E) = g_v W\frac{\sqrt{2m^*(E-E_C)}}{\pi\hbar}. \tag{64}$$

Для частного случая параболической дисперсии

$$E\left(k\right)=E_C+\frac{\hbar^2k^2}{2m^*} \tag{65}$$

и выражая волновой вектор через де бройлевскую длину волны

$$\lambda_B=2\pi/k\,, \tag{66}$$

вместо общего выражения (64) для числа мод проводимости 2D проводника имеем:

$$WM_{2D}(E)=g_v\frac{Wk}{\pi}=g_v\frac{W}{\lambda_B(E)/2}, \tag{67}$$

что открывает возможность их физической интерпретации, а именно: число мод проводимости 2D проводника (с учетом долинного вырождения) показывает сколько де бройлевских полуволн энергии E укладывается по ширине 2D проводника. Целочисленность мод обеспечивается граничными условиями, согласно которым волновые функции электронов на обеих сторонах 2D проводника должны обращаться в нуль. Сам термин «моды» заимствован из теории волноводов.

Теперь мы можем базовые уравнения модели ЛДЛ (49) и (50) переписать в виде

$$N=\int\frac{D(E)}{2}\left(f_1+f_2\right)dE\,,$$

$$I=\frac{2q}{h}\int\gamma(E)\pi\frac{D(E)}{2}\left(f_1-f_2\right)dE\,. \tag{68}$$

Таким образом, для вычисления числа электронов и тока в проводнике кроме фермиевских функций контактов нужно знать плотность состояний $D(E)$ и число мод $M(E)$ проводника.

Для параболической дисперсии (65) плотности состояний хорошо известны [20]:

$$D_{1D}(E)L=\frac{L}{\pi\hbar}\sqrt{\frac{2m^*}{\left(E-E_C\right)}}H\left(E-E_C\right), \tag{69}$$

$$D_{2D}(E)A=A\frac{m^*}{\pi\hbar^2}H\left(E-E_C\right), \tag{70}$$

$$D_{3D}(E)\Omega=\Omega\frac{m^*\sqrt{2m^*(E-E_C)}}{\pi^2\hbar^3}H\left(E-E_C\right), \tag{71}$$

где $H(E-E_c)$ – ступенчатая функция Хевисайда, а соответствующие моды проводимости таковы:

$$M_{1D}(E) = H\left(E - E_C\right), \tag{72}$$

$$M_{2D}(E) = g_v \frac{\sqrt{2m^*(E - E_C)}}{\pi\hbar} H\left(E - E_C\right), \tag{73}$$

$$M_{3D}(E) = g_v \frac{m^*(E - E_C)}{2\pi\hbar^2} H\left(E - E_C\right). \tag{74}$$

Поведение плотности состояний $D(E)$ и числа мод $M(E)$ для 1D-, 2D- и 3D-проводников с параболической дисперсией качественно показано на рис. 8.

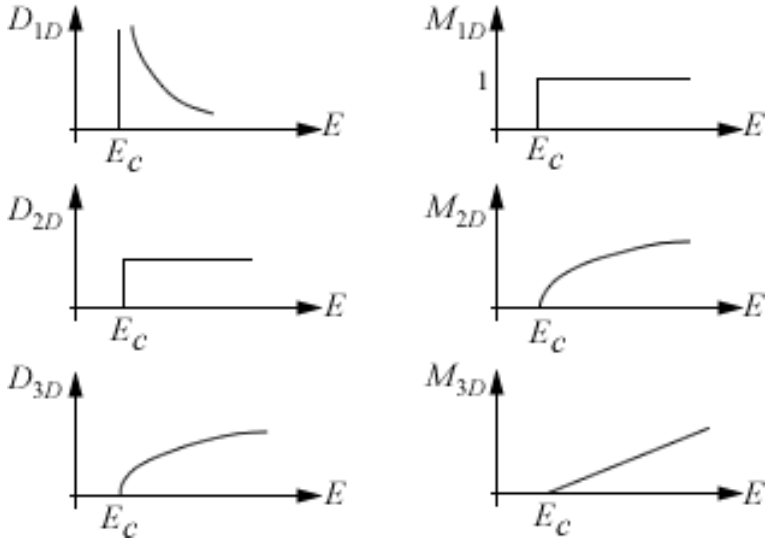


Рис. 8. Сравнительное поведение плотности состояний $D(E)$ и числа мод $M(E)$ для 1D-, 2D- и 3D-проводников с параболической дисперсией (65).

Плотность состояний для 1D проводника обратно пропорциональна $\sqrt{E}$, для 2D проводника постоянна, для 3D – прямо пропорциональна $\sqrt{E}$. Что касается мод проводимости, то 1D проводник характеризуется единственной модой проводимости, равной функции Хевисайда, что позволяет определить ее калибровку для использования ее в формулах (69) – (74). Число же мод 2D проводника прямо пропорциональна $\sqrt{E}$, а 3D проводника растет с энергией линейно.

Заключая этот раздел, подведем итоги:

1. Плотность состояний $D(E)$ нужна для вычисления плотности носителей тока.
2. Число мод $M(E)$ используется для вычисления тока.
3. Число мод $M(E)$ пропорционально средней скорости носителей тока в направлении их движения, помноженной на плотность состояний $D(E)$.
4. Число мод $M(E)$ зависит от зонной структуры и размерности проводника.

Хотя выше в качестве примера и рассмотрены проводники с параболической дисперсией, формулы (63) носят общий характер. Мы воспользуемся ими в главе 3, посвященной графену, дисперсия у которого линейная [21]. Для произвольной дисперсионной зависимости $E(k)$ разработаны численные методы расчета числа мод проводимости [22].

1.3.3. Коэффициент прохождения

До сих пор мы рассматривали баллистический транспорт. Перейдем к рассмотрению диффузионного транспорта, когда $L >> \lambda$. Электроны, инжектируемые контактами 1 и 2 , подвержены случайным блужданиям. Некоторые из них заканчиваются на инжектирующем контакте, а иные – на другом контакте. Если на контакт 2 подан положительный потенциал, то большая часть блужданий закончится на этом контакте.

Ключевым параметром в модели ЛДЛ есть величина $\gamma \pi D(E)/2$, которая в случае баллистического транспорта оказывается равной числу мод проводимости $M(E)$. Уширение γ и время пролета τ связаны между собой: $\gamma = \hbar/\tau$. В режиме диффузионного транспорта следует ожидать увеличения времени τ, что влечет за собой уменьшения величины $\gamma \pi D(E)/2$. Сейчас покажем, что в случае диффузионного транспорта $\gamma \pi D(E)/2 = M(E) \cdot T(E)$, где $T(E) \le 1$ получил название коэффициента прохождения (transmission coefficient).

В режиме баллистического транспорта электроны инжектируются в проводник под разными углами, в результате чего время пролета описывается некой функцией распределения. Величину γ мы оценивали через среднее время пролета

$$\gamma(E) = \frac{\hbar}{\langle \tau(E) \rangle}, \tag{75}$$

где

$$\langle \tau(E) \rangle = \frac{L}{\langle v_x^+(E) \rangle} = \frac{L}{v(E)\langle \cos\vartheta \rangle} = \frac{L}{v(E)(2/\pi)}. \tag{76}$$

Наша задача сейчас оценить $\langle \tau(E) \rangle$ в режиме диффузионного транспорта.

Воспользуемся первым законом диффузии Фика применительно к потоку электронов в 2D проводнике:

$$J = q\bar{D}\frac{dn_S}{dx}, \tag{77}$$

где плотность тока J пропорциональна градиенту плотности электронов с коэффициентом диффузии $\bar{D}$ в роли коэффициента пропорциональности.

Пусть электроны в проводник с подавляющим преимуществом поставляет только контакт 1, т.е. $f_1 \approx 1$, а $f_2 \approx 0$. На левом конце проводника $(x = 0)$ есть конечная плотность электронов $\Delta n_S(0)$ (рис. 9).

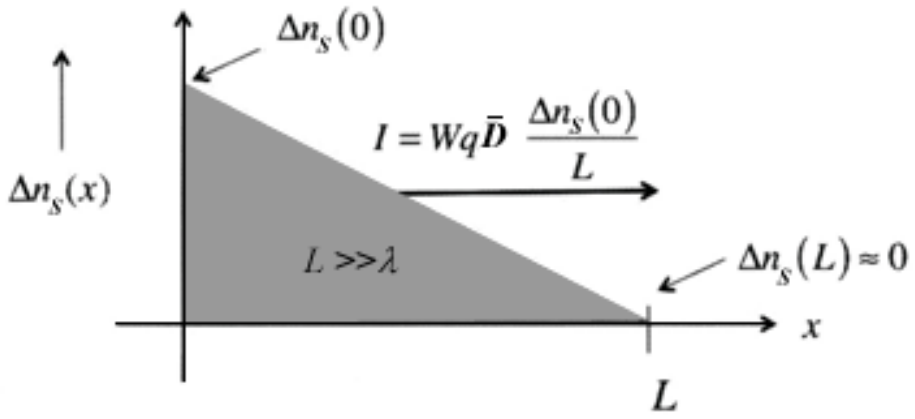


Рис. 9. К выводу времени пролета τ_D в диффузионном режиме.

Поскольку процессы рекомбинации/рождения носителей тока не учитываются, то профиль плотности электронов на всем протяжении проводника будет линейным с практически нулевым значением плотности на правом конце проводника: $\Delta_S(L) \to 0$ за счет $f_2 \approx 0$. Время пролета, согласно (54), есть отношение накопленного в проводнике заряда к току, а именно:

$$\tau_D = \frac{qN}{I} = \frac{Wq\Delta n_S(0)L/2}{Wq\bar{D}\Delta n_S(0)/L} = \frac{L^2}{2\bar{D}}, \tag{78}$$

где накопленный в проводнике заряд qN есть половина площади прямоугольника (рис. 9), построенного для 2D проводника на его длине L и ширине W, а ток $I = J \cdot W$ и, согласно (77), $dn_S/dx = \Delta n_S(0)/L$. Таким образом, время пролета в диффузионном режиме

$$\tau_D = \frac{L^2}{2\bar{D}}, \tag{79}$$

тогда как в баллистическом режиме

$$\tau_B = \frac{L}{\langle v_x^+ \rangle}. \tag{80}$$

Собирая все вместе, имеем:

$$\gamma(E)\pi\frac{D(E)}{2} = \frac{\hbar}{\tau_D}\pi\frac{D(E)}{2} = \frac{\hbar}{\tau_B}\pi\frac{D(E)}{2}\frac{\tau_B}{\tau_D} \equiv M(E)\cdot T(E), \tag{81}$$

где коэффициент прохождения есть отношение двух времен пролета – в баллистическом режиме и в диффузионном:

$$T(E) = \frac{\tau_B}{\tau_D}. \tag{82}$$

Другими словами, при наличии рассеяния число мод $M(E)$ нужно заменить на $M(E)\cdot T(E)$.

Для оценки $T(E)$ воспользуемся определением коэффициента прохождения (82) и выражениями (79) и (80). Тогда

$$T(E)=\frac{2\bar{D}}{L\left\langle v_x^+\right\rangle}. \tag{83}$$

Опираясь на теорию случайных блужданий [23], можно показать, что коэффициент диффузии

$$\bar{D}=\frac{\left\langle v_x^+\right\rangle\lambda}{2}, \tag{84}$$

что вместе с (83) дает простое выражение для коэффициента прохождения

$$T(E)=\frac{\lambda}{L}\ll 1. \tag{85}$$

Как и ожидалось, число мод проводимости в диффузионном режиме $M(E)\cdot T(E)$ резко уменьшается.

Выражение для коэффициента прохождения (85) верно для диффузионного режима. Можно показать [24], что в общем случае коэффициент прохождения

$$T(E)=\frac{\lambda(E)}{\lambda(E)+L}. \tag{86}$$

Это выражение справедливо как для диффузионного режима $(L>>\lambda)$, так и для баллистического $(L<<\lambda)$, приближаясь к единице, так и для промежуточных случаев.

Подводя итоги, в общем случае имеет место соотношение

$$\gamma(E)\pi\frac{D(E)}{2}=M(E)\cdot T(E), \tag{87}$$

в котором число мод определяется общими выражениями (63), а в отношении коэффициента прохождения в реальных экспериментах обычно рассматривают три режима:

$$\begin{gathered}\text{Диффузионный: } L>>\lambda; T=\lambda/L<<1,\\ \text{Баллистический: } L<<\lambda; T\to 1,\\ \text{Квази-баллистический: } L\approx\lambda; T<1.\end{gathered} \tag{88}$$

Примечательно, что сравнительно простая транспортная модель ЛДЛ применима ко всем трем режимам.

1.3.4. Режим линейного отклика

Для тока в модели ЛДЛ получены два выражения:

$$I = \frac{2q}{h}\int \gamma(E)\pi\frac{D(E)}{2}\left(f_1 - f_2\right)dE,$$
$$I = \frac{2q}{h}\int T(E)M(E)\left(f_1 - f_2\right)dE. \quad (89)$$

В них не учитывается важный в прикладном отношении низковольтный режим. Если подать на проводник большое напряжение, то вследствие появления множества неупругих столкновений в режиме горячих электронов нарушатся принятые нами допущения, которые привели нас к модели независимых мод проводимости. Мы далее продолжим строить модель ЛДЛ для режима линейного отклика, что позволит упростить уравнение (89).

В режиме линейного отклика, иначе линейного транспорта или почти равновесного транспорта (near-equilibrium transport), фермиевские функции контактов (20) и (21) и равновесная функция Ферми

$$f_0\left(E\right) = \frac{1}{e^{(E-E_{F0})/kT}+1} \quad (90)$$

находятся в соотношении

$$f_0(E) \approx f_1(E) > f_2(E) \approx f_0(E), \quad (91)$$

при этом подаваемое на концы проводника напряжение

$$V = \Delta E_F/q = (E_{F1} - E_{F2})/q \quad (92)$$

незначительно. Пользуясь малостью ΔE_F, фермиевскую функцию контакта 2 можно разложить в ряд Тейлора и ограничиться первой степенью, а именно:

$$f_2 = f_1 + \frac{\partial f_1}{\partial E_F}\Delta E_F = f_1 + \frac{\partial f_1}{\partial E_F}qV, \quad (93)$$

а учитывая (30) и (91), окончательно для разности фермиевских функций контактов в режиме линейного отклика имеем:

$$f_1 - f_2 = (-\frac{\partial f_0}{\partial E})qV. \quad (94)$$

Окончательно для тока (89) и проводимости в режиме линейного отклика модели ЛДЛ имеем:

$$I = \left[\frac{2q^2}{h}\int T(E)M(E)\left(-\frac{\partial f_0}{\partial E}\right)dE\right]V = GV, \quad (95)$$

$$G = \frac{2q^2}{h}\int T(E)M(E)\left(-\frac{\partial f_0}{\partial E}\right)dE, \quad (96)$$