

Наноэлектроника: возникновение тока, новая формулировка закона Ома и моды проводимости в концепции «снизу – вверх»

Ю.А. Кругляк¹, П.А. Кондратенко², Ю.М. Лопаткин³

¹ Одесский государственный экологический университет, ул. Львовская, 15, 65016 Одесса, Украина

² Национальный авиационный университет, пр. Космонавта Комарова, 1, 03058 Киев, Украина

³ Сумский государственный университет, ул. Римского-Корсакова, 2, 40007 Сумы, Украина

(Получено 12.02.2013; опубликовано online 28.03.2013)

В рамках концепции «снизу-вверх» теоретической и прикладной наноэлектроники рассматриваются общие вопросы электронной проводимости, причины возникновения тока и роль электрохимических потенциалов и фермиевских функций в этом процессе, модель упругого резистора, баллистический и диффузионный транспорт, моды проводимости, проводники *n*- и *p*-типа и графен и дается новая формулировка закона Ома.

Ключевые слова: Наноэлектроника, Молекулярная электроника, Снизу-вверх, Электрический ток, Электрохимический потенциал, Функция Ферми, Упругий резистор, Моды проводимости, Закон Ома, Проводники *n*-типа, Проводники *p*-типа, Графен.

PACS numbers: 71.15.Mb, 71.20. – b, 73.22.Pr ,
73.23.Ad, 84.32.Ff, 85.35. – p

1. ВВЕДЕНИЕ

Бурное развитие наноэлектроники в последние 10-15 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 50 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. Со 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 законов квантовой механики и неравновесной статистической термодинамики, что неизбежно ведет к пересмотру физического образования уже на начальных университетских курсах.

Согласно закону Ома, сопротивление R и проводимость G проводника длиной L с площадью поперечного сечения A даются выражениями:

$$R \equiv V / I = \rho L / A \quad \text{и} \quad 1 / R \equiv G = \sigma A / L \quad (1)$$

где удельное сопротивление ρ и обратная ему удельная проводимость σ 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 $1 \mu\text{m}$ 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 $\sim 40 \text{ nm}$, 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 в привычной записи? 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 [5].

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 «свер-

ху-вниз», от массивных проводников до 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» [6-8], которая не только оказалась комплементарной концепции «сверху-вниз», но и привела к переосмыслению 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 [9-11] 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» [12] привело к созданию единой картины транспортных явлений в электронных устройствах как наноразмерных, так и макроразмерных.

В настоящей работе в рамках концепции «снизу-вверх» рассматриваются причины возникновения тока и роль электрохимических потенциалов и фермиевских функций в этом процессе. Далее рассматривается модель «упругого резистора» и дается новая формулировка закона Ома. В рамках концепции «снизу-вверх» будут также рассмотрены общие вопросы электронной проводимости, в том числе на примере графена.

2. ПРИЧИНА ВОЗНИКНОВЕНИЯ ТОКА

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

$$\mathbf{j} = \sigma \mathbf{E}, \quad (2)$$

другими словами, причиной возникновения тока обычно считают электрическое поле. Ответ, в лучшем случае, не полный. Еще до подключения про-

водника 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 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 вычлнить отдельно внутренние и внешние электрические поля. Приходится считаться с этим уроком, преподнесенным нам современной экспериментальной наноэлектроникой, и заново задать вопрос 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 на единицу энергии $D(E)$, занятых электронами и свободных и электрохимического потенциала μ_0 (рис. 1). Простоты ради, что никак не скажется на окончательных выводах, будем пользоваться точечной моделью проводника (канала переноса электронов), которая предполагает неизменность плотности состояний $D(E)$ при движении вдоль проводника. Если система, включающая источник электрод (S/Source), проводник M и стоковый электрод (D/Drain), находится в равновесии (закорочена), то электрохимический потенциал μ_0 везде одинаков и все состояния с $E < \mu_0$ заполнены электронами, а состояния с $E > \mu_0$ пустые (рис. 1).

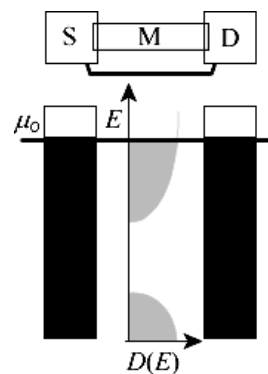


Рис. 1 – Первым шагом в объяснении работы любого электронного устройства должно быть задание плотности состояний $D(E)$ как функции энергии E внутри проводника M и определение равновесного значения электрохимического потенциала μ_0 , отделяющего заполненные электронами состояния от пустых состояний

При включении в цепь источника напряжения (рис. 2) 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

$$\mu_1 - \mu_2 = qV. \quad (3)$$

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 $\mu_1 - \mu_2$ и находящиеся достаточно близко к значениям μ_1 и μ_2 дают вклад в поток электронов, тогда как все состояния значительно выше μ_1 и ниже μ_2 не играют никакой роли. Причина этого кроется в следующем.

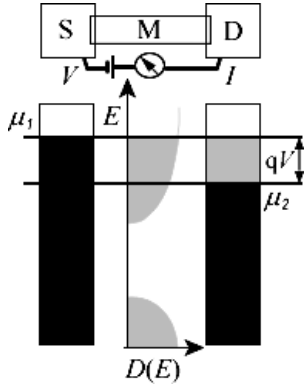


Рис. 2 – При подаче напряжения V на клеммы проводника потенциал анода D понижается на величину qV , создавая разность электрохимических потенциалов на концах проводника $\mu_1 - \mu_2 = qV$

Каждый контакт стремится привести токовый канал в равновесие с собой путем заполнения электронами всех состояний канала с энергией, меньшей электрохимического потенциала μ_1 , и опорожнения состояний канала с энергией, большей потенциала μ_2 . Рассмотрим токовый канал с состояниями, энергия которых меньше μ_1 , но больше μ_2 . Контакт 1 стремится заполнить эти состояния, поскольку их энергия меньше μ_1 , а контакт 2 стремится опорожнить эти состояния, поскольку их энергия больше μ_2 , что и приводит к непрерывному движению электронов от контакта 1 к контакту 2.

Рассмотрим теперь состояния канала с энергией, большей μ_1 и μ_2 . Оба контакта стремятся опорожнить эти состояния, но они и так пусты и не дают вклада в электрический ток. Аналогична ситуация с состояниями, энергия которых одновременно меньше обоих потенциалов μ_1 и μ_2 . 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 проводимости, 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.

3. РОЛЬ ФЕРМИЕВСКИХ ФУНКЦИЙ

Итак, утверждалось, что в равновесии все состояния с энергией $E < \mu_0$ заполнены электронами, а состояния с энергией $E > \mu_0$ пустые. 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 $E = \mu_0$, где k – постоянная Больцмана, T – абсолютная температура. Математически этот переход описывается функцией Ферми

$$f(E) = \frac{1}{\exp\left(\frac{E - \mu}{kT}\right) + 1} \quad (4)$$

График функции Ферми показан на рис. 3 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 $D(E)$.

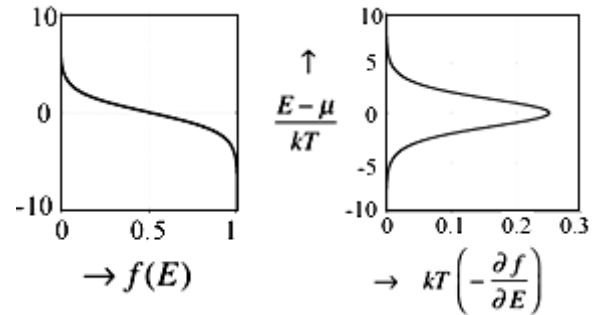


Рис. 3 – Графики функции Ферми и нормированной функции теплового уширения

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 пусты ($f = 0$), а переход от $f = 1$ к $f = 0$ происходит в узком интервале энергии $\sim \pm 2 kT$, охватывающем значение $E = \mu_0$.

Действительно, на рис. 3 показана производная от функции Ферми, помноженная на kT с тем, чтобы сделать ее безразмерной

$$F_T(E, \mu) = kT \left(-\frac{\partial f}{\partial E} \right) \quad (5)$$

Подставляя выражение для функции (4), видно, что

$$F_T(E, \mu) = \frac{e^x}{(e^x + 1)^2} \quad (6)$$

где $x \equiv (E - \mu) / kT$. Из (6) сразу видно, что

$$F_T(E, \mu) = F_T(E - \mu) = F_T(\mu - E), \quad (7)$$

а из уравнений (6) и (4) следует, что

$$F_T = f(1 - f). \quad (8)$$

Интегрирование функции (8) во всем интервале изменения энергии дает площадь, равную kT

$$\int_{-\infty}^{+\infty} dE F_T(E, \mu) = kT \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f}{\partial E} \right) = kT [f]_{-\infty}^{+\infty} = kT(1 - 0) = kT, \quad (9)$$

Так что функцию F_T можно приблизительно представить себе в виде прямоугольного «импульса», центрированного относительно значения $E = \mu$ с высотой, равной $1/4$ и шириной $\sim 4 kT$.

4. ВЫХОД ИЗ РАВНОВЕСИЯ

Когда рассматриваемая система (рис. 1) находится в равновесии, 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. 2) 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 фермиевские функции (рис. 4)

$$f_1(E) = \frac{1}{\exp\left(\frac{E - \mu_1}{kT}\right) + 1} \quad (10)$$

и

$$f_2(E) = \frac{1}{\exp\left(\frac{E - \mu_2}{kT}\right) + 1}. \quad (11)$$

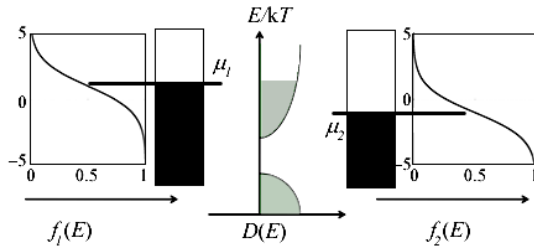


Рис. 4 – При выходе из равновесия электроны в контактах занимают доступные им состояния в соответствии с фермиевскими распределениями и значениями электрохимических потенциалов

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 макроразмерных. Однако, для наноразмерных проводников ток при любых значениях энергии электронных состояний в проводнике пропорционален разности $I(E) \sim f_1(E) - f_2(E)$ фермиевских распределений в обоих контактах. Эта разность зануляется, если энергия E больше μ_1 и μ_2 , поскольку в этом случае обе функции Ферми равны нулю. Эта разность также зануляется, если энергия E меньше μ_1 и μ_2 , поскольку в этом случае обе фермиевские функции равны единице. Ток возникает лишь в окне $\mu_1 - \mu_2$, если в этом окне есть хотя бы одно электронное состояние проводника.

5. ЛИНЕЙНЫЙ ОТКЛИК

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

Построим функцию разности двух фермиевских функций, нормированную к приложенному напряжению

$$F(E) = \frac{f_1(E) - f_2(E)}{qV/kT}, \quad (12)$$

где

$$\begin{aligned} \mu_1 &= \mu_0 + (qV/2) \\ \mu_2 &= \mu_0 - (qV/2) \end{aligned} \quad (13)$$

Функция разности $F(E)$ сужается по мере того, как напряжение V , помноженное на заряд электрона, становится все меньше величины kT (рис. 5). Отметим также, что по мере того, как kT начинает превосходить энергию qV , функция $F(E)$ все больше

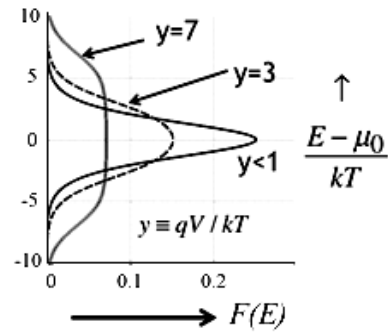


Рис. 5 – График функции разности $F(E)$ в зависимости от значения $(E - \mu_0) / kT$ для различных значений $qV / kT \equiv y$

приближается к функции теплового уширения (5)

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

так что из уравнения (12) следует, что

$$f_1(E) - f_2(E) = \frac{qV}{kT} F_T(E, \mu_0) = \left(-\frac{\partial f_0}{\partial E}\right) qV \quad (14)$$

если приложенное напряжение, помноженное на заряд электрона, $\mu_1 - \mu_2 = qV$ становится намного меньшим kT .

Нам потребуется также следующее выражение

$$f(E) - f_0(E) = -\frac{\partial f_0}{\partial E} (\mu - \mu_0) \quad (15)$$

которое, как и уравнение (14), можно получить следующим образом.

Для функции Ферми

$$f(x) \equiv \frac{1}{e^x + 1}, \quad x \equiv \frac{E - \mu}{kT} \quad (16)$$

имеем

$$\frac{\partial f}{\partial E} = \frac{df}{dx} \frac{\partial x}{\partial E} = \frac{df}{dx} \frac{1}{kT}$$

$$\frac{\partial f}{\partial \mu} = \frac{df}{dx} \frac{\partial x}{\partial \mu} = -\frac{df}{dx} \frac{1}{kT} \quad , \quad (17)$$

$$\frac{\partial f}{\partial T} = \frac{df}{dx} \frac{\partial x}{\partial T} = -\frac{df}{dx} \frac{E - \mu}{kT^2}$$

откуда

$$\frac{\partial f}{\partial \mu} = -\frac{\partial f}{\partial E}$$

$$\frac{\partial f}{\partial T} = -\frac{E - \mu}{T} \frac{\partial f}{\partial E} \quad . \quad (18)$$

Уравнение (15) получается из разложения функции Ферми вблизи точки равновесия в ряд Тейлора

$$f(E, \mu) \cong f(E, \mu_0) + \left(\frac{\partial f}{\partial \mu} \right)_{\mu=\mu_0} (\mu - \mu_0). \quad (19)$$

Из уравнения (18) следует

$$\left(\frac{\partial f}{\partial \mu} \right)_{\mu=\mu_0} = \left(-\frac{\partial f}{\partial E} \right)_{\mu=\mu_0}. \quad (20)$$

Пусть $f(E)$ соответствует $f(E, \mu)$, а $f_0(E)$ соответствует $f(E, \mu_0)$, тогда

$$f(E) \approx f_0(E) + (\mu - \mu_0) \left(-\frac{\partial f}{\partial E} \right), \quad (21)$$

что после перегруппировки дает искомое уравнение (15), которое справедливо при $\mu - \mu_0 \ll kT$.

Подведем предварительные 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 нескольких kT , охватывающих равновесный электрохимический потенциал μ_0 , определяемый функцией F_T (ф-ла 5, рис. 3), 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 μ_0 , что в корне отличает одно вещество от другого.

Настоящий ответ нельзя назвать новым, и он хорошо известен специалистам в области наноэлектроники. Тем не менее обсуждение проводимости и поныне обычно начинается с теории Друде [13], которая сыграла важнейшую историческую роль в понимании природы тока. К сожалению, подход Друде породил два недоразумения, которые следовало бы преодолеть и прежде всего в преподавании физики, а именно:

- ток порождается электрическим полем и
- 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 окне вокруг равновесного электрохимического потенциала μ_0 . Проводимость канала высокая или низкая зависит только от наличия электронных состояний в этом окне. К этому выводу обычно приходят на основе транспортного уравнения Больцмана [14] или формализма Кубо [15], тогда как используемая нами концепция «снизу – вверх» сразу дает физически корректную картину возникновения тока.

6. МОДЕЛЬ УПРУГОГО РЕЗИСТОРА

Итак, ток порождается «подготовкой» двух контактов 1 и 2 с фермиевскими функциями $f_1(E)$ и $f_2(E)$. Отрицательному контакту 1 соответствует большее значение электрохимического потенциала, а положительному – меньшее значение. Отрицательный контакт стремится передать электроны в канал проводимости, а положительный контакт стремится извлечь электроны из канала проводимости. Это справедливо для любых проводников – и наноразмерных и макро-размерных.

Модель упругого резистора служит полезной идеализацией, обеспечивающей физически корректное объяснение работы наноразмерных проводников и открывающей возможности для новой интерпретации работы макро-размерных устройств. Концепция «упругого резистора» была предложена Рольфом Ландауэром в 1957 году [9-11] задолго до ее экспериментального подтверждения в нанотранзисторах [1]. Концепция «упругого резистора», строго говоря, является идеализацией, но она надежно подтверждена многочисленными экспериментальными данными для ультрамалых нанотранзисторов [3]. Развитие концепции упругого резистора [6-8, 12] привело к созданию единой картины транспортных явлений в электронных устройствах любой размерности.

В модели упругого резистора электроны проскакивают канал проводимости от истокового контакта S к стоковому D упруго, без потери или приобретения энергии (рис. 6).

В упругом резисторе ток в промежутке с энергией от E до $E + dE$ отделен от каналов с другими значениями энергии, что позволяет нам записать для тока в дифференциальной форме

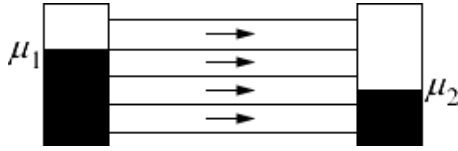


Рис. 6 – В упругом резисторе электроны движутся баллистически по каналам с постоянной энергией.

$$dI = dEG(E)(f_1(E) - f_2(E)), \quad (22)$$

а после интегрирования получить выражение для полного тока. Используя далее выражение (14), получим выражение для низковольтной проводимости (линейный отклик)

$$G \equiv \frac{I}{V} = \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) G(E), \quad (23)$$

в котором отрицательную производную $(-\partial f_0 / \partial E)$ можно представить себе в виде прямоугольного импульса, площадь которого равна единице при ширине $\sim \pm 2 kT$ (рис. 3). Согласно (23), функция проводимости $G(E)$ для упругого резистора, будучи усредненной по интервалу $\sim \pm 2 kT$, включает значение электрохимического потенциала μ_0 дает экспериментально измеряемую проводимость G . При низких температурах можно просто использовать значение $G(E)$ при $E = \mu_0$.

Такой энергетический подход к проводимости в модели упругого резистора дает существенное упрощение в понимании причин тока, хотя и звучит парадоксально, поскольку мы традиционно связываем ток I , текущий по проводнику с сопротивлением R , с джоулевым теплом I^2R . Как можно говорить о сопротивлении, если движущиеся по проводнику электроны не теряют энергии?

Ответ на этот вопрос следующий: поскольку электроны не теряют энергию при движении по упругому резистору, потеря энергии происходит на границах проводника с истоковым и стоковым контактами, на которых и происходит диссипация джоулева тепла. Иначе говоря, упругий резистор, характеризуемый сопротивлением R канала проводимости, диссипирует джоулево тепло I^2R за пределами канала проводимости. На это указывают различные многочисленные экспериментальные измерения, косвенные и прямые, на наноразмерных проводниках [3, 4], не говоря уже о том, что диссипация тепла, будь она шла бы на одиночной молекуле или на нанопроводнике привела бы к их сгоранию, и общее мнение сейчас таково, что сгорания в реальных экспериментах не происходит, поскольку подавляюще большая часть тепла генерируется на контактах, которые достаточно массивны и тем самым с достаточной легкостью диссипируют тепло.

В понятие упругого резистора не вкладывается обязательного следования электроном по прямолинейной траектории от истока к стоку, допускается и «диффузионное» движение с переменным вектором импульса, но без изменения энергии.

Модель упругого резистора вводится не только в качестве полезного понятия для объяснения работы наноразмерных устройств, но также и потому, что

эта модель позволяет объяснить такие транспортные свойства как проводимость в макроразмерных проводниках. Именно эта модель делает концепцию «снизу-вверх» [8, 12] столь эффективной в объяснении транспортных явлений в целом. Еще вернемся к этому утверждению позже, а сейчас получим выражение для проводимости упругого резистора.

7. ПРОВОДИМОСТЬ УПРУГОГО РЕЗИСТОРА

Следуя духу концепции «снизу-вверх», рассмотрим простейший упругий резистор с одним каналом энергии ε , с которой электрон проскакивает от истока к стоку (рис. 7).

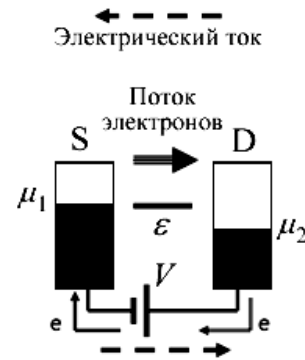


Рис. 7 – Одноуровневая модель упругого резистора с энергией ε в канале

Напомним, что в результате приписывания электрону отрицательного заряда, что уже изменить невозможно, контакт с большим напряжением D имеет меньший электрохимический потенциал, а движение электрона по каналу происходит от большего значения электрохимического потенциала к меньшему, так что направление тока противоположно реальному движению электронов от истока S к стоку D . Фактически, всегда имеется в виду, что речь идет о токе реальных электронов, а не о токе в привычном звучании.

Результирующий одноканальный ток будет

$$I = \frac{q}{t} (f_1(\varepsilon) - f_2(\varepsilon)) \quad (24)$$

где t есть время, затрачиваемое электроном на прощолк от истока S к стоку D . Теперь можно выражение (24) обобщить на произвольный упругий резистор (рис. 6) с произвольной плотностью состояний $D(E)$, имея в виду, что все энергетические каналы в упругом резисторе проводят независимо друг от друга и в параллельном режиме. Выпишем вначале выражение для тока в канале с энергией от E до $E + dE$

$$dI = dE \frac{D(E)q}{2t} (f_1(E) - f_2(E)), \quad (25)$$

в котором учтено, что в таком канале имеется $D(E)dE$ состояний, из которых только половина дает вклад в ток от истока S к стоку D . Интегрируя, получим выражение для тока по упругому резистору

$$I = \frac{1}{q} \int_{-\infty}^{+\infty} dE G(E) (f_1(E) - f_2(E)), \quad (26)$$

где

$$G(E) = \frac{q^2 D(E)}{2t(E)}. \quad (27)$$

Если разность $\mu_1 - \mu_2 = qV$ за счет напряжения на контактах V намного меньше, чем kT , вправе воспользоваться уравнением (14) и написать

$$I = V \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) G(E), \quad (28)$$

что и приводит к уравнению (23). В общем случае

$$G = \frac{q^2 D}{2t}, \quad (29)$$

однако, нужно помнить, что плотность состояний D и время пролета t в общем случае зависят от энергии и должны быть усреднены в пределах $\pm 2kT$, включающих электрохимический потенциал μ_0 . Это выражение представляется правильным и на интуитивном уровне. Оно утверждает, что проводимость пропорциональна произведению двух факторов, а именно: наличию состояний (D) и легкости, с которой электрон преодолевает расстояние от истока к стоку ($1/t$). Это ключевой результат для дальнейшего обсуждения. Сейчас же мы перейдем к более подробному рассмотрению диссипации тепла упругим резистором.

8. ДИССИПАЦИЯ ТЕПЛА УПРУГИМ РЕЗИСТОРОМ

Сопротивление R упругого резистора определяет каналом, а соответствующее тепло P^2R выделяется вне канала проводимости. Рассмотрим подробнее эту ситуацию на примере одноуровневой модели упругого резистора с энергией в канале ε (рис. 7). Каждый раз когда электрон из истока баллистически проскакивает канал, он оказывается в состоянии «горячего электрона» на стоке с энергией ε , превышающей электрохимический потенциал стока μ_2 (рис. 8).

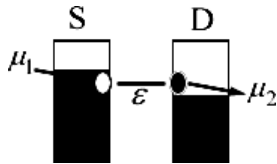


Рис. 8 – Мгновенная картина после проскока электрона с истока на сток с энергией в канале ε , превышающей электрохимический потенциал стока μ_2

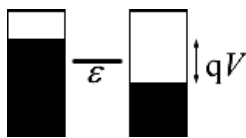


Рис. 9 – После баллистического пролета электрона с истока на сток по каналу с энергией ε , на истоке выделяется энергия $\mu_1 - \varepsilon$, а на стоке – энергия $\varepsilon - \mu_2$, и контакты снова находятся в равновесии

Диссипационные процессы в стоковом контакте быстро рассеивают избыток энергии:

$\varepsilon - \mu_2$. Подобным же образом на истоковом контакте «дырка» с энергией ε , меньшей электрохимического потенциала истока μ_1 заполняется электроном, а избыток энергии;

$\mu_1 - \varepsilon$ диссипирует на истоке (рис. 9).

Суммарная энергия, диссипирующая на контактах, равна $\mu_1 - \mu_2 = qV$. Если N электронов проскакивают с истока на сток за время t , то на контактах диссипирует мощность

$$P = V \cdot I = qV \cdot N / t. \quad (30)$$

Таким образом, тепло, выделяемое при прохождении тока по упругому резистору, рассеивается на контактах, чему уже достаточно имеется экспериментальных подтверждений на нано- и мезоразмерных проводниках [16]. Привлекательность модели упругого резистора состоит в том, что механические и термодинамические процессы пространственно разделены.

9. МОДЕЛЬ УПРУГОГО РЕЗИСТОРА И МАКРОПРОВОДНИКИ

Естественно задаться вопросом о применимости и пользе модели упругого резистора для объяснения и понимания физики явления протекания тока в обычных макропроводниках, движение электронов в которых очевидно неупругое. В макропроводниках неупругие процессы совмещены с упругими и беспорядочно распределены по всему проводнику (рис. 10). Разобьем такой макропроводник с беспорядочно распределенными неупругими столкновениями на последовательность упругих резисторов (рис.11) длины L , намного меньшей длины реального макропроводника и с перепадом напряжения между соседними упругими резисторами, составляющим лишь незначительные доли от разности потенциалов на концах реального проводника $\mu_1 - \mu_2 = qV$. В разбиении длины L должны быть меньше длины L_{in} , представляющей собой среднюю длину пробега электрона до очередного неупругого столкновения. Кроме условия

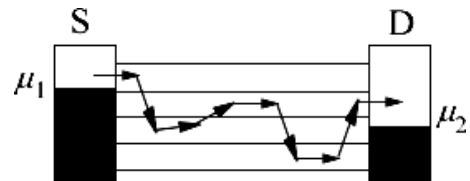


Рис. 10 – В макропроводниках неупругие процессы совмещены с упругими

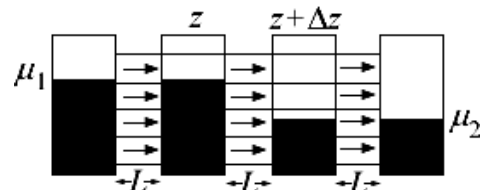


Рис. 11 – Гипотетическое разбиение реального макропроводника на серию упругих резисторов

$L < L_{in}$ должно при разбиении соблюдаться также требование падения напряжения между соседними упругими резисторами $\Delta V < kT/q$.

Разбиение крупного проводника на упругие резисторы требует некоторой осторожности. Как будет показано ниже, стандартное выражение для закона Ома следует заменить на

$$R = \rho(L + \lambda) / A, \quad (31)$$

где A – площадь поперечного сечения проводника и в котором дополнительное сопротивление $\rho\lambda / A$ не зависит от длины проводника и может быть интерпретировано как пограничное сопротивление, возникающее на границе канал/контакт. В выражении (31) λ есть длина, близкая к средней длине свободного пробега, так что такая модификация закона Ома существенна лишь для баллистических проводников ($L \sim \lambda$) и не существенна для крупных проводников ($L \gg \lambda$). Однако, концептуально это дополнительное сопротивление окажется исключительно важным, если пользоваться гипотетической структурой на рис. 11 для объяснения реальной ситуации на рис. 10. Структура на рис. 11 имеет много пограничных поверхностей раздела, отсутствующих в реальной ситуации (рис. 10), так что следует избавиться от этих виртуальных границ. Например, если каждая секция длины L на рис. 11 характеризуется сопротивлением (31), то правильным выражением для проводника на рис. 10, например, длины $3L$ будет

$$R = \rho(3L + \lambda) / A, \quad (32)$$

а не

$$R = \rho(3L + 3\lambda) / A. \quad (33)$$

Таким образом, для получения корректного выражения для проводимости длинного проводника в рамках модели упругого резистора следует аккуратно отделить пограничное сопротивление от сопротивления проводника, зависящего от его длины.

10. БАЛЛИСТИЧЕСКИЙ И ДИФFUЗИОННЫЙ ТРАНСПОРТ

Как мы видели выше, проводимость упругого резистора дается выражением (29) $G = q^2 D / 2t$.

Покажем, что время пролета t через резистор длины L при диффузионном режиме с длиной свободного пробега λ связано с временем пролета при баллистическом режиме t_B соотношением

$$t = t_B \left(1 + \frac{L}{\lambda} \right). \quad (34)$$

Подставим (34) в (29), учтем, что

$$G_B \equiv q^2 D / 2t_B \quad (35)$$

и окончательно для проводимости в диффузионном режиме получим

$$G = G_B \lambda / (L + \lambda). \quad (36)$$

Обращая проводимость (36), для закона Ома в новой формулировке получим

$$R = \frac{\rho}{A} (L + \lambda), \quad (37)$$

где

$$\frac{\rho}{A} = \frac{1}{\sigma A} = \frac{1}{G_B \lambda}. \quad (38)$$

До сих пор речь шла о трехмерном резисторе с площадью поперечного сечения A (рис. 12).

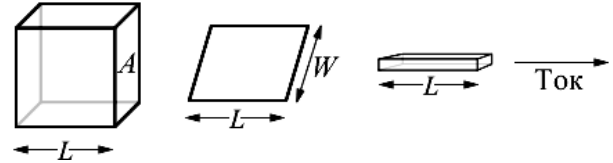


Рис. 12 – Проводники размерности 3d, 2d и 1d.

Разнообразные эксперименты выполняются на двухмерных проводниках шириной W с одномерным поперечным сечением. Для таких 2d-резисторов соответствующие выражения для закона Ома, очевидно, имеют вид

$$R = \frac{\rho}{W} (L + \lambda) \quad (39)$$

где

$$\frac{\rho}{W} = \frac{1}{\sigma W} = \frac{1}{G_B \lambda}. \quad (40)$$

Наконец, для одномерных проводников имеем

$$R = \rho (L + \lambda), \quad (41)$$

где

$$\rho = \frac{1}{\sigma} = \frac{1}{G_B \lambda}. \quad (42)$$

Запишем закон Ома компактно для проводников всех трех размерностей

$$R = \rho (L + \lambda) \left\{ 1, \frac{1}{W}, \frac{1}{A} \right\}, \quad (43)$$

где

$$\frac{1}{\rho} = \sigma = G_B \lambda \left\{ 1, \frac{1}{W}, \frac{1}{A} \right\}. \quad (44)$$

Выражение в фигурных скобках соответствует 1d-, 2d- и 3d-проводникам. Заметим, что удельные сопротивление и проводимость имеют разную размерность в зависимости от размерности проводника, а проводимость и длина по-прежнему измеряются в сименсах и метрах.

Стандартный закон Ома утверждает, что сопротивление стремится к нулю при уменьшении длины проводника до нуля. Никто, конечно, не ожидает, что сопротивление обратится в нуль, но обычное об-

щее мнение таково, что сопротивление стремится к некому пограничному сопротивлению, которое можно сделать произвольно малым по мере совершенствования измерительной техники. Что сейчас надежно установлено экспериментально, так это то, что при наиболее тщательно подготовленных контактах наименьшее наблюдаемое сопротивление связано с каналом проводимости и не зависит от контактов [2]. Модифицированный закон Ома отражает это обстоятельство: даже при приближении длины проводника к нулю остается остаточное сопротивление, связанное с эффективной длиной λ . Уместно, однако, спросить себя, какой смысл имеет говорить о ненулевой длине λ при нулевой длине проводника. Ответ состоит в том, что для наноразмерных проводников ни удельное сопротивление ρ , ни длина λ не имеют смысла порознь, существенно только их произведение.

11. БАЛЛИСТИЧЕСКИЙ И ДИФFUЗИОННЫЙ ТРАНСПОРТ

Рассмотрим как плотность состояний D и время пролета t в выражении для проводимости (29) соотносится с размерами канала в крупных проводниках. Что касается плотности состояний, то это аддитивное свойство. В два раза больший канал имеет в два раза больше электронных состояний, так что плотность состояний для крупных проводников должна быть пропорциональна объему проводника $A \cdot L$.

Что касается времени пролета t , обычно рассматривают два транспортных режима:

баллистический с $t \sim L$ и диффузионный с $t \sim L^2$. Баллистическая проводимость пропорциональна площади поперечного сечения A проводника и, согласно (29), не зависит от длины проводника. Такое «неомическое» поведение действительно наблюдается в наноразмерных проводниках [17]. Что же касается проводников с диффузионным транспортным режимом, то они демонстрируют нормальное «омическое» поведение проводимости $G \sim A / L$.

Различие в двух транспортных режимах можно пояснить следующим образом. В баллистическом режиме время пролета от истока к стоку

$$t_B = \frac{L}{\bar{u}}, \quad (45)$$

где

$$\bar{u} = \langle |v_z| \rangle \quad (46)$$

есть среднее значение скорости электронов вдоль оси z , направления движения электронов от истока к стоку.

В случае диффузионного режима время t зависит от длины проводника квадратично

$$t = \frac{L}{\bar{u}} + \frac{L^2}{2\bar{D}}, \quad (47)$$

где величина \bar{D} в рамках теории случайных блужданий [18] есть коэффициент диффузии

$$\bar{D} = \langle v_z^2 \tau \rangle \quad (48)$$

где τ есть время свободного пробега.

Используя определение (45), перепишем уравнение (47) в виде

$$t = t_B \left(1 + \frac{L\bar{u}}{2\bar{D}} \right), \quad (49)$$

которое вместе с уравнением (34) для длины λ дает значение

$$\lambda = \frac{2\bar{D}}{\bar{u}}. \quad (50)$$

Для вычисления констант \bar{u} по уравнению (46) и \bar{D} по уравнению (48) требуется усреднить скорость электронов в направлении их движения (ось z) по всем угловым переменным в зависимости от размерности проводника $d = \{1, 2, 3\}$. Простые вычисления дают для

$$\text{1d-проводника } \langle |v_z| \rangle = v \text{ и } \langle |v_z|^2 \rangle = v^2; \quad (51)$$

$$\text{2d-проводника } \langle |v_z| \rangle = 2v/\pi \text{ и } \langle |v_z|^2 \rangle = v^2/2; \quad (52)$$

$$\text{3d-проводника } \langle |v_z| \rangle = v/2 \text{ и } \langle |v_z|^2 \rangle = v^2/3. \quad (53)$$

так что в итоге

$$\bar{u} = \langle |v_z| \rangle = v(E) \left\{ 1, \frac{2}{\pi}, \frac{1}{2} \right\}, \quad (54)$$

$$\bar{D} = \langle v_z^2 \tau \rangle = v^2 \tau(E) \left\{ 1, \frac{1}{2}, \frac{1}{3} \right\}, \quad (55)$$

или окончательно для длины среднего свободного пробега λ имеем

$$\lambda = \frac{2\bar{D}}{\bar{u}} = v\tau \left\{ 2, \frac{\pi}{2}, \frac{4}{3} \right\}. \quad (56)$$

Подчеркнем, что длина λ включает численный множитель, зависящий от размерности проводника, сравнительно с ее стандартным значением $\lambda = v \cdot \tau$. Можно ли просто использовать это стандартное значение λ . Можно, но тогда в новой формулировке закона Ома (43) нужно не просто заменять L на $L + \lambda$, а на L прибавить λ , помноженное на численный фактор, зависящий от размерности проводника, либо пользоваться определением длины λ по (56). Любопытно, что этот фактор даже для одномерного проводника не равен единице, а равен двум. Величина τ есть среднее время свободного пролета до очередного столкновения. Поскольку рассеяние предполагается изотропным, только половина актов столкновения ведет электрон от истока к стоку. С учетом обратного рассеяния значение длины λ для 1d-проводника оказывается равным $2v\tau$.

Теперь получим уравнение для баллистической проводимости. Из уравнений (35) и (45) имеем

$$G_B \equiv \frac{q^2 D \bar{\mu}}{2L}, \quad (57)$$

а после подстановки уравнения (54) получим

$$G_B \equiv \frac{q^2 D v}{2L} \left\{ 1, \frac{2}{\pi}, \frac{1}{2} \right\}. \quad (58)$$

Окончательно, подставляя в (44) уравнения (56) и (58) и учитывая (55) для удельной проводимости получим

$$\sigma = q^2 \bar{D} \frac{D}{L} \left\{ 1, \frac{1}{W}, \frac{1}{A} \right\}. \quad (59)$$

Итак, на основе общего выражения для проводимости (29) и времен пробега (45) и (47) получены выражения для проводимости в баллистическом режиме (59) и в диффузионном режиме (36).

12. МОДЫ ПРОВОДИМОСТИ

Из уравнения (58) видно, что баллистическая проводимость пропорциональна плотности состояний на единицу длины проводника D/L . Поскольку плотность состояний пропорциональна объему, то можно ожидать, что баллистическая проводимость пропорциональна площади поперечного сечения A $3d$ -проводника или ширине W $2d$ -проводника.

Многочисленные эксперименты показывают [17], что для наноразмерных проводников баллистическая проводимость не изменяется линейно с их площадью поперечного сечения, а кратна кванту проводимости

$$G_B \equiv \frac{q^2}{h} M. \quad (60)$$

Другими словами, реальный проводник можно рассматривать как M независимых мод проводимости, дающих в сумме баллистическую проводимость. Учитывая (58), для числа мод проводимости получаем

$$M \equiv \frac{h D v}{2L} \left\{ 1, \frac{2}{\pi}, \frac{1}{2} \right\}, \quad (61)$$

а из уравнений (44) и (60) удельная проводимость выражается через число мод проводимости M и среднюю длину свободного пробега λ

$$\sigma = \frac{q^2}{h} M \lambda \left\{ 1, \frac{1}{W}, \frac{1}{A} \right\}. \quad (62)$$

Более подробно концепцию мод проводимости рассмотрим ниже.

13. ФУНДАМЕНТАЛЬНОЕ СООТНОШЕНИЕ

Стандартное выражение для удельной проводимости дается формулой Друде [13], связывающей проводимость σ с электронной плотностью n , эффективной массой m и временем свободного пробега τ

$$\sigma \equiv \frac{1}{\rho} = \frac{q^2 n \tau}{m}, \quad (63)$$

или используя понятие мобильности

$$\bar{\mu} = \frac{q \tau}{m}, \quad (64)$$

имеем

$$\sigma = q n \bar{\mu}. \quad (65)$$

С другой стороны, для удельной проводимости в рамках концепции «снизу-вверх» получаются два эквивалентных выражения, одно из которых выражает проводимость через произведение плотности состояний D и коэффициента диффузии \bar{D} (59), а другое – через произведение числа мод M в канале проводимости и средней длины пробега λ (62).

Как и проводимость

$$G \equiv \frac{I}{V} = \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) G(E),$$

удельную проводимость по уравнениям (59) и (62) нужно усреднить по энергии в области нескольких kT , включающей $E = \mu_0$ с помощью функции теплового уширения

$$\sigma = \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) \sigma(E). \quad (66)$$

Уравнение (59) хорошо известно, выводится в стандартных учебниках по физике твердого тела [13], чего нельзя сказать об эквивалентном ему уравнению (62), вывод которого обычно требует использования статистической термодинамики необратимых процессов, например, формализма Кубо [14, 15].

По поводу модели Друде хотелось бы подчеркнуть следующее. Применимость модели Друде весьма ограничена, в то время как уравнения для проводимости (59) и (62) имеют самое общее значение. Например, эти уравнения применимы к графену [19, 20] с непараболическим поведением зон и «безмассовыми» электронами – свойствами, описание которых невозможно в модели Друде. Один из уроков, преподнесенных наноэлектроникой, – широкая применимость уравнений для проводимости (59) и (62).

Фундаментальное различие между уравнениями (59) и (62) и теорией Друде состоит в том, что усреднение (66) делает проводимость свойством фермиевской поверхности: проводимость определяется уровнями энергии, близкими к $E = \mu_0$. А согласно уравнениям (63)-(65) теории Друде проводимость зависит от общей электронной плотности, суммированной по всему спектру энергий, что и приводит к ограниченной применимости модели Друде. Проводимость веществ меняется в очень широких пределах несмотря на то, что число электронов приблизительно одинаково. Низкая проводимость стекла объясняется не тем, что в стекле мало так называемых «свободных» электронов, а потому, что для стекла характерна очень низкая плотность состояний и числа мод вбли-

зи $E = \mu_0$. Понятие же «свободных» электронов относится к интуитивным понятиям.

Для произвольных проводников, как с кристаллической структурой, так и аморфных, так и для молекулярных проводников, следуя [12], покажем, что независимо от функциональной зависимости $E(p)$ плотность состояний $D(E)$, скорость $v(E)$ и импульс $p(E)$ связаны с числом электронных состояний $N(E)$ с энергией, меньшей значения E , соотношением

$$D(E)v(E)p(E) = N(E) \cdot d, \quad (67)$$

где d – размерность проводника. Используя (67) для вычисления проводимости (59) с учетом коэффициента диффузии (55)

$$\bar{D} = \langle v_z^2 \tau \rangle.$$

для $3d$ -проводника получим

$$\sigma(E) = q^2 \frac{N(E) \tau(E)}{A \cdot L m(E)}. \quad (68)$$

где масса определена как

$$m(E) = \frac{p(E)}{v(E)}. \quad (69)$$

Легко убедиться, что фундаментальное соотношение (67) справедливо и для параболической зависимости $E(p)$ и для линейной, как в графене [20]. Для параболической зависимости масса носителя тока от энергии не зависит, что в общем случае не так.

Формула (68) выглядит как выражение (63) теории Друде, если $N/A \cdot L$ считать электронной плотностью n . При низких температурах это действительно так, поскольку усреднение (66) при $E = \mu_0$ дает

$$\sigma = \left(q^2 \frac{N}{A \cdot L m} \right)_{E=\mu_0} = q^2 n \tau / m, \quad (70)$$

так как $N(E)$ при $E = \mu_0$ есть полное число электронов (рис. 13). При ненулевой температуре ситуация тем более сложнее, если плотность состояний непараболическая. Отметим, что ключевым моментом в сведении общего выражения для проводимости (59) к выражению (68), похожему на формулу Друде (63), есть фундаментальное выражение (67), связывающее плотность состояний $D(E)$, скорость $v(E)$ и импульс $p(E)$ для данного значения энергии с полным числом состояний $N(E)$, полученных интегрирование плотности

$$N(E) = \int_{-\infty}^E dE D(E). \quad (71)$$

Каким образом полное число состояний $N(E)$ по (71) может быть однозначно связано с плотностью состояний $D(E)$, скоростью $v(E)$ и импульсом $p(E)$ для конкретного значения энергии? Ответ состоит в том, что уравнение (67) удовлетворяется только в том случае, когда уровни энергии вычисляются

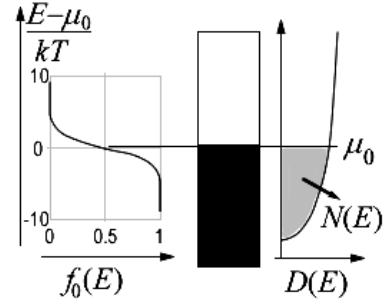


Рис. 13 – Равновесная функция Ферми $f_0(E)$. Плотность состояний $D(E)$ и полное число электронов $N(E)$

однозначно из выражения для $E(p)$. Оно может не выполняться в области энергий перекрывающихся зон или, например, для аморфных тел, когда не известна зависимость $E(p)$. В этих случаях уравнения (59) и (62) не эквивалентны уравнению (68) и пользоваться можно только первыми.

Теперь посмотрим, как одиночные зоны, описываемые различными соотношениями $E(p)$ приводят к фундаментальному уравнению (67) и таким образом откроется возможность установления связи между выражениями для проводимости (59) и (62) и формулами Друде (63-65). Это приведет также к новой интерпретации мод $M(E)$, введенных выше, и к объяснению их целочисленности.

14. ДИСПЕРСИИ $E(p)$ ДЛЯ КРИСТАЛЛИЧЕСКИХ ТВЕРДЫХ ТЕЛ

Пусть стандартно связь энергии с импульсом параболическая (рис. 14)

$$E(p) = E_c + \frac{p^2}{2m}, \quad (72)$$

где под m подразумевается эффективная масса. Будем пользоваться соотношениями $E(p)$ вместо $E(k)$, хотя всегда можно перейти к волновому вектору $k = p/\hbar$. Дисперсия (72) широко используется для различных веществ – и для металлов и для полупроводников. Но это не единственная возможность. Для графена [19, 20], использование которого в нанoeлектронике, как ожидается, приведет к следующему шагу в миниатюризации, имеет место линейная зависимость от импульса (рис. 15)

$$E = E_c + v_0 p, \quad (73)$$

где v_0 – константа, равная приблизительно 1/300 от скорости света. Здесь и ранее под импульсом p подразумевается его абсолютное значение. Другими словами, подразумевается, что зависимость $E(p)$ изотропная.

Для изотропных $E(p)$ скорость сонаправлена с импульсом, а величина ее равна

$$v \equiv \frac{dE}{dp}. \quad (74)$$

15. К ПОДСЧЕТУ ЧИСЛА СОСТОЯНИЙ

На длине резистора L должно укладываться целое число де-бройлевских волн с длиной $\lambda = h/p$. $Lp/h =$ целому числу или $p =$ целое число (h/L) .

Это означает, что разрешенные состояния однородно распределены для данного значения p и каждое из состояний занимает промежуток

$$\Delta p = \frac{h}{L}. \tag{75}$$

Определим функцию $N(p)$ как полное число состояний со значениями импульса, меньшими, чем заданное значение p . Для одномерных проводников $1d$ (рис. 16) эта функция равна отношению доступной длины $2p$ (от $-p$ до $+p$) к промежутку Δp

$$N(p) = \frac{2p}{h/L} = 2L \left(\frac{p}{h} \right). \tag{76}$$

Для $2d$ -проводников (рис. 17) нужно разделить площадь поперечного сечения πp^2 на промежутки по длине h/L и по площади поперечного сечения h/W , так что окончательно

$$N(p) = \frac{\pi p^2}{(h/L)(h/W)} = \pi W \cdot L \left(\frac{p}{h} \right)^2. \tag{77}$$

Для $3d$ -проводников объем сферы радиуса p делится на произведение промежутков $(h/L) \cdot (h/W_1) \cdot (h/W_2)$, где площадь поперечного сечения $A = W_1 \cdot W_2$, так что окончательно

$$N(p) = \frac{(4\pi/3)p^3}{(h/L)(h^2/A)} = \frac{4\pi}{3} A \cdot L \left(\frac{p}{h} \right)^3, \tag{78}$$

или собирая все вместе для $d = \{1, 2, 3\}$ имеем

$$N(p) = \left\{ 2 \frac{L}{h/p}, \pi \frac{L \cdot W}{(h/p)^2}, \frac{4\pi}{3} \frac{L \cdot A}{(h/p)^3} \right\}. \tag{79}$$

Задаввшись законом дисперсии $E(p)$, теперь можно посчитать зависимость числа состояний $N(E)$ с энергией, меньшей заданного значения E .

16. ПЛОТНОСТЬ СОСТОЯНИЙ $D(E)$

Полученная таким образом функция числа состояний $N(E)$ должна быть равна плотности состояний $D(E)$, проинтегрированной вплоть до энергии состояний E

$$N(E) = \int_{-\infty}^E dE D(E),$$

так что плотность состояний

$$D(E) = \frac{dN}{dE}, \tag{80}$$

а с использованием уравнения (79) получается

$$D(E) = \frac{dN}{dp} \frac{dp}{dE} = \frac{dp}{dE} \frac{p^{d-1} d}{h^d} \left\{ 2L, \pi LW, \frac{4}{3} \pi LA \right\}. \tag{81}$$

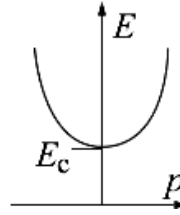


Рис. 14 – Параболическая дисперсия

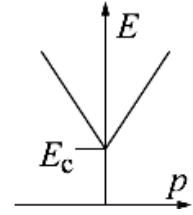


Рис. 15 – Линейная дисперсия

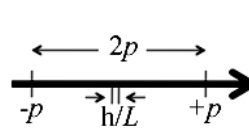


Рис. 16 – К подсчету числа состояний для $1d$ -проводника

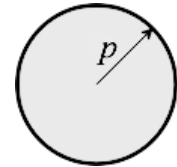


Рис. 17 – К подсчету числа состояний для $2d$ -проводника

Воспользовавшись (74) и (79), окончательно получим искомое фундаментальное уравнение (67), независимое от закона дисперсии.

17. ФОРМУЛА ДРУДЕ

Как уже было показано, используя (67) для вычисления проводимости (59) для $3d$ -проводника получаем выражение (68), в котором зависящая от энергии масса определена уравнением (69). Также было показано, что уравнение (68) сводится к формуле Друде (63) при температурах, близких к нулю. Теперь рассмотрим проводники n -типа и p -типа отдельно при температурах, отличных от нуля.

17.1 Проводники n -типа

Используя уравнение (68) и предполагая в нем независимость от энергии массы m и времени τ , получим

$$\sigma = \frac{q^2 \tau}{m} \frac{1}{A \cdot L} \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) N(E). \tag{82}$$

Интегрируя по частям, имеем

$$\int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) N(E) = \left[-N(E) f_0(E) \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} dE \frac{dN}{dE} f_0(E) \tag{83}$$

$$= [0 - 0] + \int_{-\infty}^{+\infty} dE D(E) f_0(E) = \text{полному числу электронов,}$$

поскольку произведение $dE \cdot D(E) f_0(E)$ есть число электронов в области энергий от E до $E + dE$. Таким образом, уравнение (82) сводится к формуле Друде

$$\sigma = \frac{q^2 \tau}{m} \frac{N}{A \cdot L}, \tag{84}$$

имея в виду, что $N / A \cdot L = n$.

17.2 Проводники *p*-типа

Любопытная ситуация имеет место для *p*-проводников с ниспадающей дисперсией, например,

$$E(p) = E_c - \frac{p^2}{2m}. \quad (85)$$

Вместо числа состояний по (71) мы теперь имеем (рис. 18)

$$N(E) = \int_E^{+\infty} dE' D(E') \quad (86)$$

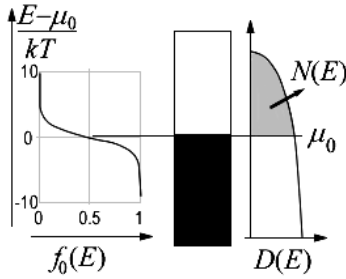


Рис. 18 – Равновесная функция Ферми $f_0(E)$, плотность состояний $D(E)$ и число состояний $N(E)$ для *p*-проводника с дисперсией (85)

что дает

$$D(E) = -\frac{dN}{dE}. \quad (87)$$

Так получается потому, что функция $N(E)$ определяется через функцию $N(p)$, которая дает полное число состояний с импульсом, меньшим заданного значения p , что соответствует энергиям, большим чем заданное значение E согласно дисперсионному соотношению (85).

Если, как и прежде, проинтегрировать по частям

$$\int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) N(E) = \left[-N(E) f_0(E) \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} dE \frac{dN}{dE} f_0(E), \quad (88)$$

теперь первое слагаемое не зануляется, поскольку и $N(E)$ и $f_0(E)$ в нижнем пределе не нули.

Обойти эту ситуацию можно следующим образом: вместо производной от f_0 брать производную от $(1 - f_0)$

$$\int_{-\infty}^{+\infty} dE \left(-\frac{\partial(1-f_0)}{\partial E} \right) N(E) = \left[-N(E)(1-f_0(E)) \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} dE \frac{dN}{dE} (1-f_0(E)) \quad (89)$$

$$= [0-0] + \int_{-\infty}^{+\infty} dE D(E)(1-f_0(E)) = \text{полному числу «дырок»}.$$

Другими словами, для проводников *p*-типа можно пользоваться формулой Друде

$$\sigma = q^2 n \tau / m, \quad (90)$$

если под значением n понимать число «дырок»: большему значению n соответствует меньшее число электронов.

17.3 Графен

Как подсчитывать значение n в случае, когда зоны распространяются в обоих направлениях как в графене с дисперсией $E = \pm v_0 p$ [19, 20] (рис. 7 слева). Нельзя не признать изобретательным разбить зоны в графене на зону *n*-типа и зону *p*-типа (рис. 19 справа) так, что а затем пользоваться формулами Друде.

Хотелось бы подчеркнуть, что нет нужды для подобной изобретательности, поскольку уравнения (59) и (62) применимы во всех случаях и корректно отражают физику проводимости.

$$D(E) = D_n(E) + D_p(E), \quad (91)$$

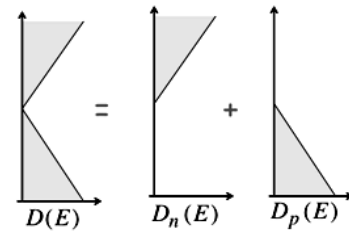


Рис. 19 – Искусственное разбиение зонной структуры графена на зоны *n*- и *p*-типа

18. ПРОПОРЦИОНАЛЬНА ЛИ ПРОВОДИМОСТЬ ЭЛЕКТРОННОЙ ПЛОТНОСТИ?

Экспериментальные измерения проводимости часто ведутся в зависимости от электронной плотности, которые, согласно теории Друде, связаны линейно, так что отклонения от линейности трактуют как проявления зависимости времени свободного пробега от энергии. При этом не учитывают, что для непараболических дисперсий масса носителя тока, сама определяемая как p/v , может зависеть от энергии и тем самым приводить к нелинейности проводимости от плотности электронов.

Сначала определим электронную плотность из уравнения (79)

$$n(p) = \left\{ 2 \frac{p}{h}, \pi \frac{p^2}{h^2}, \frac{4\pi}{3} \frac{p^3}{h^3} \right\}, \quad (92)$$

где плотность n равна N/L , $N/W \cdot L$ и $N/A \cdot L$ для $d = 1, 2$ и 3 . Перепишем (92) в виде

$$n(p) = K p^d, \quad (93)$$

где коэффициент пропорциональности $K = \{2/h, \pi/h^2, 4\pi/3h^3\}$. Теперь для проводимости (70) с учетом (69) имеем

$$\sigma = q^2 \frac{n(p)\tau(p)}{m(p)} = q^2 K p^{d-1} v(p) \tau(p). \quad (94)$$

Если известна или выбрана зависимость скорости и времени свободного пробега от энергии, а стало быть и от импульса, в уравнениях (93) и (94) можно избавиться от зависимости от импульса и таким образом установить связь между проводимостью σ и электронной плотностью n .

Например, в случае графена $E = \pm v_0 p$, скорость dE/dp постоянна и равна v_0 , не зависит от импульса. Полагая время свободного пробега от энергии не зависящим, проводимость от плотности электронов из уравнений (93) и (94) с учетом уравнения (56) для длины свободного пробега

$$\lambda = \frac{2\bar{D}}{u} = v\tau \left\{ 2, \frac{\pi}{2}, \frac{4}{3} \right\}$$

получается следующей

$$\sigma = \frac{q^2}{h} \lambda \sqrt{\frac{4n}{\pi}} \quad (95)$$

или с учетом g -фактора (для графена $g = 4$)

$$\sigma = \frac{q^2}{h} \lambda \sqrt{\frac{4gn}{\pi}} \quad (96)$$

Таким образом, проводимость в графене получается пропорциональной $\sim \sqrt{n}$, а не как обычно предполагается $\sim n$, и с временем свободного пробега, не зависящем от энергии. Вычисления по формуле (96) с $\lambda = 2 \mu\text{м}$ и $\lambda = 300 \text{ нм}$ (рис. 20) согласуются с экспериментальными данными [21].

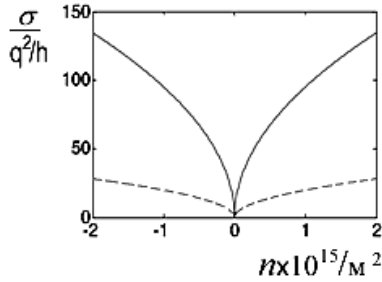


Рис. 20 – Проводимость графена по уравнению (96) как функция электронной плотности для значений $\lambda = 2 \mu\text{м}$ (сплошная) и $\lambda = 300 \text{ нм}$ (пунктиром) согласуется с экспериментальными данными (рис. 1 в [21])

19. КВАНТОВАНИЕ ПРОВОДИМОСТИ И МОДЫ ПРОВОДИМОСТИ

Баллистическая проводимость квантуется

$$G_B \equiv \frac{q^2}{h} M.$$

где для низкоразмерных проводников при низких температурах число M – целое. Выше было получено выражение для числа мод

$$M \equiv \frac{\hbar D v}{2L} \left\{ 1, \frac{2}{\pi}, \frac{1}{2} \right\} \quad (97)$$

через произведение плотности состояний D и скорости электронов v , и совершенно не очевидна целочисленность выражения (97). Пользуясь выражениями для дисперсии $E(p)$ оказывается возможным дать иную интерпретацию $M(p)$, свидетельствующую

о целочисленной природе величины M . Используя (67), перепишем (97) в виде

$$M = \frac{\hbar N}{2Lp} \left\{ 1, \frac{4}{\pi}, \frac{3}{2} \right\}, \quad (98)$$

где $N(p)$ есть полное число состояний с импульсом, меньшим заданного значения p . С использованием (79) преобразуем (98) в

$$M(p) = \left\{ 1, 2 \frac{W}{\hbar p}, \pi \frac{A}{(\hbar p)^2} \right\}. \quad (99)$$

Так же как число состояний $N(p)$ дает нам число де-бройлевских длин волн, укладывающихся в проводнике, так и $M(p)$ дает нам число мод, укладывающихся в поперечном сечении проводника, причем это число не зависит от закона дисперсии, поскольку при выводе (99) нигде не использовался какой-либо конкретный закон дисперсии.

При оценке чисел $N(p)$ и $M(p)$ в конкретной задаче на практике получаются, конечно, дробные числа. Однако, по физическому смыслу эти числа должны быть целыми. В больших проводниках при низких температурах квантование $M(p)$ смазывается, однако, в мезо- и наноразмерных проводниках наблюдается целочисленная природа числа мод $M(p)$ и квантование проводимости. Поэтому (99) правильнее переписать в виде

$$M(p) = \text{Int} \left\{ 1, 2 \frac{W}{\hbar p}, \pi \frac{A}{(\hbar p)^2} \right\}, \quad (100)$$

где под $\text{Int}\{x\}$ подразумевается наибольшее целое число, меньшее значения x .

В одномерных проводниках число мод совпадает с g -фактором, равным числу долин, умноженному на вырождение по спину, равное 2. Сопротивление баллистических проводников $\sim M \hbar / q^2$, так что сопротивление баллистического $1d$ -проводника приблизительно равно 25 К Ω , деленных на g , что и наблюдается экспериментально [1]: большинство металлов и полупроводников типа GaAs имеют $g = 2$ и баллистическое сопротивление $1d$ -образцов $\sim 12.5 \text{ К}\Omega$, а углеродные нанотрубки двухдолинные с $g = 4$ и их баллистическое сопротивление $\sim 6.25 \text{ К}\Omega$.

Настоящая работа явилась результатом посещения проф. Кругляком Ю.А. курса лекций «Fundamentals of Nanoelectronics, Part I: Basic Concepts», прочитанного он-лайн в январе – феврале 2012 года проф. С. Датта (Supriyo Datta) в рамках инициативы Purdue University / nanoHUB-U [www.nanohub.org/u].

БЛАГОДАРНОСТЬ

Авторы выражают благодарность проф. Supriyo Datta за проявленный интерес к работе и поддержку её публикации.

Nanoelectronics: Current Generation, New Formulation of Ohm's Law and Conduction Modes by "Bottom-Up" Approach

Yu.A. Kruglyak¹, P.A. Kondratenko², Yu.M. Lopatkin³

¹ *Odessa State Environmental University str. Lviv, 15, 65016 Odessa, Ukraine*

² *National Aviation University, 1, Komarov ave. , 03058 Kyiv, Ukraine*

³ *Sumy State University, 2, Rimskiy-Korsakov str., 40007 Sumy, Ukraine*

General questions of electronic conductivity, current generation with the use of electrochemical potentials and Fermi functions, elastic resistor model, ballistic and diffusion transport, conductivity modes, *n*- and *p*-conductors and graphene, new formulation of the Ohm's law are discussed in the frame of the «bottom-up» approach of modern nanoelectronics.

Keywords: Nanoelectronics, Molecular electronics, Bottom-up, Electric current, Electrochemical potential, Fermi function, Elastic resistor, Conductivity modes, Ohm's law, *n*-type conductors, *p*-type conductors, Graphene.

Наноелектроніка: виникнення струму, нове формулювання закону Ома і моди провідності в концепції «знизу-вгору»

Ю.О. Кругляк¹, П.О. Кондратенко², Ю.М. Лопаткін³

¹ *Одеський державний екологічний університет, вул. Львівська, 15, 65016 Одеса, Україна*

² *Національний авіаційний університет, пр. Космонавта Комарова, 1, 03058 Київ, Україна*

³ *Сумський державний університет, вул. Римського-Корсакова, 2, 40007 Суми, Україна*

В рамках концепції «знизу-вгору» теоретичної і прикладної наноелектроніки розглядаються загальні питання електронної провідності, причини виникнення струму та роль електрохімічних потенціалів і фермієвських функцій в цьому процесі, модель пружного резистора, балістичний і дифузійний транспорт, моди провідності, провідники *n*- і *p*-типу та графен, дається нове формулювання закону Ома.

Ключові слова: Наноелектроніка, Молекулярна електроніка, Знизу-нагору, Електричний струм, Електрохімічний потенціал, Функція Фермі, Пружний резистор, Моди провідності, Закон Ома, Провідники *n*-типу, Провідники *p*-типу, Графен.

СПИСОК ЛІТЕРАТУРЫ

- V.V. Mitin, V.A. Kochelap, M.A. Strosio, *Introduction to Nanoelectronics: Science, Nanotechnology, Engineering, and Applications* (Cambridge: Cambridge University Press: 2012).
- Hoefflinger Bernd (Editor), *Chips 2020: A Guide to the Future of Nanoelectronics (Frontiers Collection)*, (Berlin: Springer-Verlag: 2012).
- Дж.М. Мартинес-Дуарт, Р.Дж. Мартин-Палма, Ф. Гулло-Руеда, *Нанотехнологии для микро- и оптоэлектроники* (Москва: Техносфера: 2007).
- В.П. Драгунов, И.Г. Неизвестный, В. А. Гридчин, *Основы наноелектроники* (Москва: Логос: 2006).
- R.H.M. Smit, Y. Noat, C. Untiedt, N.D. Lang, M.C. van Hemert, J.M. van Ruitenbeek, *Nature* **419**, N 3, 906 (2002).
- Datta Supriyo, *Electronic Transport in Mesoscopic Systems* (Cambridge: Cambridge University Press: 2001).
- Datta Supriyo, *Quantum Transport: Atom to Transistor* (Cambridge: Cambridge University Press: 2005).
- Electronics from the Bottom Up: A New Approach to Nanoelectronic Devices and Materials* [Електронний ресурс]// Режим доступа: www.nanohub.org/topics/ElectronicsFromTheBottomUp <05.01.2013>
- Landauer Rolf, *IBM J. Res. Dev.* **1**, 223 (1957).
- Landauer Rolf, *Philos. Mag.* **21**, 863 (1970).
- Landauer Rolf, *J. Math. Phys.* **37**, 5259 (1996).
- Datta Supriyo, *Lessons from Nanoelectronics: A New Perspective on Transport* (Hackensack, New Jersey: World Scientific Publishing Company: 2012).
- Н. Ашкрофт, Н. Мермин, *Физика твердого тела, том 1* (Москва: Мир: 1979).
- F.W. Sears, G.L. Salinger, *Thermodynamics, Kinetic Theory, and Statistical Thermodynamics* (Boston: Addison-Wesley: 1975).
- R. Kubo, *J. Phys. Soc. Japan* **12**, 570 (1957).
- M. Lundstrom, Guo Jing, *Nanoscale Transistors: Physics, Modeling, and Simulation* (Berlin: Springer: 2006).
- Yu.V. Nazarov, Ya.M. Blanter, *Quantum Transport. Introduction to nanoscience* (Cambridge: Cambridge University Press: 2009).
- Berg Howard C., *Random walks in biology* (Princeton: Princeton University Press: 1993).
- М.В. Стріха, *Сенсорна електроніка і мікросистемні технології* **1(7)**, N 3, 5 (2010).
- Ю.А. Кругляк, Н.Е. Кругляк, *Вісник Одеського держ. екологічного ун-ту*, В. 13, 207 (2012).
- K.I. Bolotin, K.J. Sikes, J. Hone, P. Kim, H.L., *Phys. Rev. Lett.* **101**, 096802 (2008).