



БИБЛИОТЕЧКА • КВАНТ •

выпуск 33

М. Е. ЛЕВИНШТЕЙН
Г. С. СИМИН

ЗНАКОМСТВО С ПОЛУПРОВОДНИКАМИ





БИБЛИОТЕЧКА • КВАНТ •

выпуск 33

М. Е. ЛЕВИНШТЕЙН

Г. С. СИМИН

ЗНАКОМСТВО С ПОЛУПРОВОДНИКАМИ



МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1984



Scan AAW

22.37
Л 35
УДК 539.2

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:

Академик **И. К. Кикоин** (председатель), академик **А. Н. Колмогоров** (заместитель председателя), профессор **Л. Г. Асламазов** (ученый секретарь), член-корреспондент АН СССР **А. А. Абрикосов**, академик **Б. К. Вайнштейн**, заслуженный учитель РСФСР **Б. В. Воздвиженский**, академик **П. Л. Капица**, профессор **С. П. Капица**, академик **С. П. Новиков**, академик **Ю. А. Осипьян**, академик АПН СССР **В. Г. Разумовский**, академик **Р. З. Сагдеев**, профессор **Я. А. Смородинский**, академик **С. Л. Соболев**, член-корреспондент АН СССР **Д. К. Фаддеев**, член-корреспондент АН СССР **И. С. Шкловский**.

Левинштейн М. Е., Симин Г. С.

Л 35 Знакомство с полупроводниками/Под ред. **Л. Г. Асламазова**.—М.: Наука. Главная редакция физико-математической литературы, 1984. — 240 с. — (Библиотечка «Квант». Вып. 33). — 40 к.

В книге в увлекательной форме рассказывается о физике полупроводников, об истории открытия и изучения полупроводников от первых опытов Гильберта и Кавендиша до наших дней, о многих интересных эффектах в полупроводниках и о полупроводниковых приборах.

Для школьников, студентов, преподавателей, лиц, занимающихся самообразованием.

Л $\frac{1704060000-043}{053(02)-84}$ 192-84

ББК 22.37
539.1

Л $\frac{1704060000-043}{053(02)-84}$ 192-84

© Издательство «Наука».
Главная редакция
физико-математической
литературы, 1984

ОГЛАВЛЕНИЕ

От авторов	6
Введение. Главные герои	7
Часть I. НА ПЛЕЧАХ ГИГАНТОВ	23
Глава 1. <i>De profundis</i>	25
Глава 2. <i>Per aspera ad astra</i>	45
Часть II. СЕКРЕТ ЗОЛОТОЙ СЕРЕДИНЫ	64
Глава 3. Принцы на горошине	64
Что произойдет, если кристалл нагреть?	65
Что произойдет, если блюдо с водой поставить на шкаф?	65
Вот что произойдет, если кристалл нагреть!	67
Тепловая генерация электронов и дырок	67
Рекомбинация электронов и дырок	69
Собственные полупроводники	70
Примесные полупроводники	74
Донорная примесь (74). Примесное истощение (77). Снова собственная проводимость (77). Зависимость концентрации электронов от температуры (77). Дырки в электронном полупроводнике (79). Акцепторная примесь (80). Знакомая песня с одним новым словом (81).	
Ах, как кружится голова...	84
Компенсация	85
Глава 4. Компенсаторы, ступеньки, убийцы	87
Компенсаторы	88
Ступеньки	90
Равновесные и неравновесные условия (94).	
«Убийцы»	95
Избыточные носители. Фототок (96). Измерение времени жизни неравновесных носителей (98). Герои гибнут в пучине Мальстрема (100).	
Глава 5. Сквозь игольное ушко	102
Тепловое движение	102
Движение в электрическом поле	103
Подвижность (104). Механизмы рассеяния (103). Эффективная масса (111). Горячие электроны (112).	
Проводимость	114
Закон Ома в дифференциальной форме (114). Такие простые зависимости (115). Измерение проводимости (117).	

Диффузия	120
Кoeffициент диффузии (123). Диффузионный ток (125). Соотношение Эйнштейна (126). Скорость диффузии (127). Задача о случайных блужданиях (128). Диффузия неосновных носителей (130).	
Часть III. ПРОФЕССИИ ПРИНЦЕВ	134
Глава 6. Терморезисторы	135
Просто термосопротивление	136
Низкие температуры (136). Высокие температуры (139). Не очень высокие и не очень низкие температуры (139). Температурный коэффициент сопротивления (140). Измерение температуры (142). Температурная компенсация (143). Термистор — переменное сопротивление (144). Болометры (145).	
Секреты буквы S	146
S-транная вольтамперная характеристика (147). Волшебная прямая (151). S-транные прыжки (153). Сбор урожая (158).	
Глава 7. Тензорезисторы	160
Зачем нужно измерять деформации	161
Как можно измерить деформацию	163
Механические тензометры (163). Оптические тензометры (164). Электрические тензометры (164). Емкостный тензодатчик (164). Индуктивный тензодатчик (164).	
Проволочные тензодатчики сопротивления	165
Полупроводниковые тензодатчики сопротивления — тензорезисторы	169
Что изменяется при деформации кристалла (170). Главное свойство кристаллов (171). От физического исследования к прибору (174). Что умеют тензорезисторы (178).	
Глава 8. Фоторезисторы. Внутренний фотоэффект	179
Как трудно быть точным	180
Как важно быть точным	180
Внутренний фотоэффект	183
Кoeffициент поглощения (184). Спектральная зависимость коэффициента поглощения (185). Собственное поглощение (186). Поглощение на свободных носителях (187). Примесное поглощение (189). Квантовая эффективность (192).	
Фотопроводимость	194
Скорость генерации (195). Скорость рекомбинации (195). Стационарная фотопроводимость (196). Как устанавливается фотопроводимость (197). Как спадает фотопроводимость (199). Спектральная зависимость фотопроводимости (200).	
Фоторезисторы	203
Что умеют фоторезисторы (205). Основные параметры фоторезисторов (208).	
Глава 9. Эффект Холла	212
Почему возникает эффект Холла	214
Напряженность холловского поля	215
Что измеряется экспериментально	216
Применения эффекта Холла	217
Измерение магнитного поля (218). Измерение тока (219). Измерение мощности сверхвысокочастотных электромагнитных волн (220).	

Глава 10. Эффект Ганна	222
В чем состоит эффект Ганна	223
Что происходит в образце	224
Удивительная зависимость скорости электронов от электрического поля	226
Два важных обстоятельства (227). Спасительные импульсы (227). А все-таки они горячие! (228). Почему уменьшается подвижность (230).	
В сущности, все это довольно просто	232
Флуктуация поля на омическом участке зависимости $\bar{v}(E)$ рассасывается (232). Флуктуация поля на падающем участке зависимости $\bar{v}(E)$ нарастает и превращается в домен (233).	
Применения эффекта Ганна	235
Заключение	238

Знакомство с полупроводниками — дело безусловно нужное и полезное. Предмет знакомства вхож буквально в каждый дом. И, кроме того, обладает такими громадными связями в самых разных сферах — научных, промышленных и военных, которые человеку непосвященному просто трудно себе представить.

Однако мы давно убедились на собственном опыте, что завязывать и поддерживать знакомство только потому, что оно — нужное, во-первых, неэтично, а во-вторых, скучно.

Полупроводники — не только очень полезный, но и исключительно интересный предмет. Главным образом потому мы и взялись способствовать знакомству с ним читателя.

Создать обстановку, способствующую знакомству, нам помогли наши коллеги и друзья: А. Г. Бовтрук, М. И. Дьяконов, А. В. Зданович, Б. И. и Д. Б. Шкловские, В. Я. Френкель. Мы глубоко им за это признательны. Мы очень благодарны М. А. Симиной, чьи забавные и интересные иллюстрации очень оживили, как нам кажется, атмосферу знакомства с полупроводниками.

ГЛАВНЫЕ ГЕРОИ

Ба! знакомые все лица...
А. С. Грибоедов. «Горе от ума»

Имена главных героев нашей книги — электрон и дырка.

Электрон — герой далеко не одного только «полупроводникового романа». Мы встречаемся с электроном в космических лучах и в ускорителях элементарных частиц, в металлах и в вакуумных электронных приборах (например, в телевизионной трубке или радиолампе). История того, как был открыт электрон и как удалось убедиться в том, что во всех этих и многочисленных других случаях мы имеем дело с одной и той же элементарной частицей, подробно рассказана в замечательной книге М. П. Бронштейна «Атомы и электроны» (М.: Наука, 1980. Библиотечка «Квант», вып. 1).

Дырка ограничивает свою деятельность значительно более узкими рамками. Чаще всего она встречается в полупроводниках.

Дырка заметно моложе электрона. Термин «электрон» для частицы, несущей элементарный заряд отрицательного электричества $1,6 \cdot 10^{-19}$ Кл, предложил в 1891 г. ирландский физик Джонстон Стоней. Термин «дырка» для «частицы», способной создавать ток в полупроводниках и несущей элементарный положительный заряд, равный по абсолютной величине заряду электрона, предложил в 1933 г. советский физик Яков Ильич Френкель. Слово «частица» применительно к дырке не случайно поставлено в кавычки. Дело в том, что дырка, вообще-то говоря, частицей не является. Она обладает многими чертами, которые роднят ее с героем чудесной повести Юрия Тынянова «Подпоручик Киже». Этот

подпоручик имел воинское звание, супругу, детей, награды, командовал воинскими соединениями, но... никакого подпоручика Кижэ в природе никогда не существовало.

Дырка имеет заряд и массу. Часто говорят о концентрации дырок в полупроводниковом материале, о траектории дырки в электрическом и магнитном поле, о взаимодействии дырок и электронов, но... никакой такой частицы в природе не существует. Понятие «дырка» означает просто, что в том месте кристаллической решетки, где полагалось бы находиться электрон, он отсутствует.

Чтобы понять, почему пустое место, дырка, получила почетное наименование частицы, как она приобрела массу и заряд, нам придется внимательно присмотреться к сцене, на которой действуют наши герои — электрон и дырка: к полупроводникам.

Вот как определяет полупроводники новое издание Большой Советской Энциклопедии: «Полупроводники — широкий класс веществ, характеризующийся значениями электропроводности σ , промежуточными между электропроводностью металлов ($\sigma \sim 10^8\text{—}10^6 \text{ Ом}^{-1}\cdot\text{м}^{-1}$) и хороших диэлектриков ($\sigma \leq 10^{-8}\text{—}10^{-10} \text{ Ом}^{-1}\cdot\text{м}^{-1}$), электропроводность указана при комнатной температуре. Характерной особенностью полупроводников, отличающей их от металлов, является возрастание электропроводности σ с ростом температуры...»

Взглянем, однако, на рис. 1. На этом рисунке показана зависимость проводимости кремния, содержащего одну десятиллионную долю процента примеси фосфора, от температуры. (Если бы речь шла не о полупроводнике, то о таком малом количестве примеси не стоило бы и упоминать, но нам еще предстоит убедиться в том, что в полупроводниках с примесями шутки плохи. Даже ничтожное количество примесей способно очень сильно сказаться на свойствах полупроводника.) Кремний, как известно, типичный полупроводник. Подавляющее большинство полупроводниковых приборов изготавливается именно из кремния. Однако из рис. 1 видно, что если величина σ полностью отвечает приведенному выше определению, то зависимость σ от температуры носит прямо противоположный характер. Как и у ме-

таллов, проводимость уменьшается с ростом температуры.

На рис. 2 показаны две кривые. Кривая 1 представляет зависимость проводимости от температуры для полупроводникового соединения InSb (антимонид индия), содержащего в виде примеси одну тысячную долю процента теллура. Кривая 2 — такую

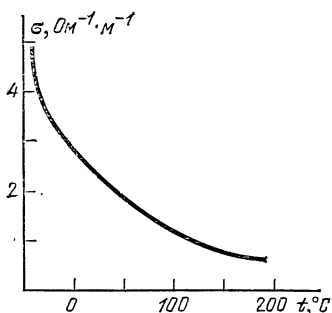


Рис. 1. Зависимость проводимости кремния от температуры (кремний содержит одну десятимиллионную долю процента примеси фосфора).

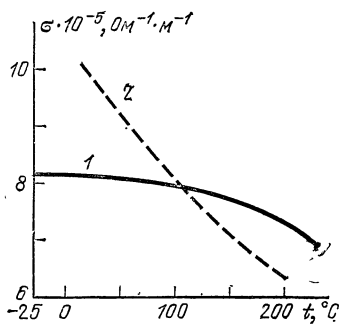


Рис. 2. Зависимость проводимости от температуры.

Кривая 1 — для полупроводникового соединения InSb (антимонид индия), содержащего одну тысячную долю процента примеси теллура. Кривая 2 — для ртути.

же зависимость для ртути. Ртуть — металл. Антимонид индия — типичный полупроводник. (Из него изготавливают приемники, чувствительные к инфракрасным лучам. Такие фотоприемники лежат в основе приборов ночного видения.) Между тем из рис. 2 видно, что не только у ртути — металла, но и у InSb — полупроводника проводимость падает с ростом температуры. Кроме того, и величины проводимости у металла и у полупроводника не так уж сильно отличаются даже при комнатной температуре ($\sigma \sim 10^6$ Ом⁻¹·м⁻¹). А в области чуть более высоких температур проводимость полупроводника даже больше, чем проводимость металла!

Как же так? Неужели такое авторитетное издание, как Большая Советская Энциклопедия, помещает на своих страницах непродуманное определение? Конечно, нет. Дело в том, что сам предмет, подлежащий определению — *полупроводники*, слишком сложен, чтобы описать его по одним лишь внешним

признакам. Такая ситуация встречается нередко и отнюдь не только в физике. Например, каждый из нас, конечно, знает, что такое человек. Однако...

Около 2500 лет назад ученики попросили знаменитого греческого философа Платона дать определение тому, что такое человек. Платон ответил блестящим и, на первый взгляд, безупречно точным афоризмом: «Двуногое без перьев». На следующий

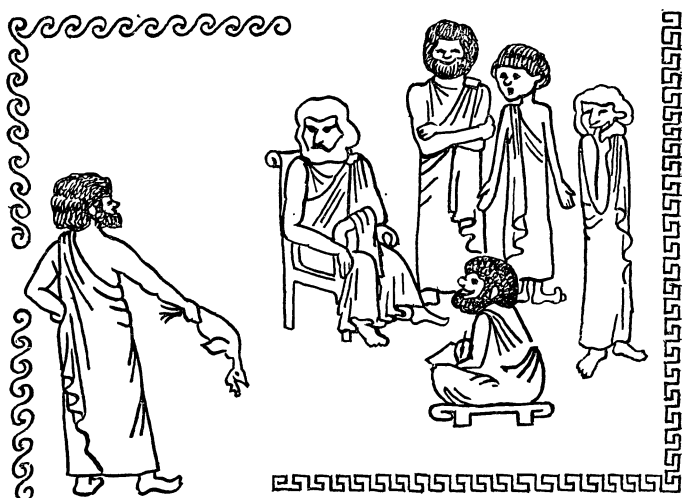


Рис. 3. Человек по Платону.

день не менее знаменитый Диоген принес на занятия платоновской Академии ощипанного петуха и, демонстрируя его ученикам, пояснил: «Человек по Платону» (рис. 3) *).

Тех, кого интересует вопрос, как определить, что такое человек, мы отсылаем к интересной книге французского писателя Веркора «Люди или животные», а сами возвратимся к полупроводникам.

Если судить по внешним признакам — проводимости и ее зависимости от температуры, понять, что такое полупроводник, затруднительно; нам ничего не

*) Было ли это происшествие на самом деле — неизвестно, но поскольку эта история впервые была рассказана всего через 100 лет после смерти Платона, то уже в силу давности к ней следует отнести с надлежащим почтением.

остается, как обратиться к рассмотрению внутренних свойств кристаллов.

Рис. 4 многим из читателей может показаться знакомым. Он с небольшими изменениями заимствован из школьного учебника физики (9 класс) и представляет собой упрощенное изображение кристаллической решетки кремния (Si). Каждый атом кремния имеет на внешней электронной оболочке

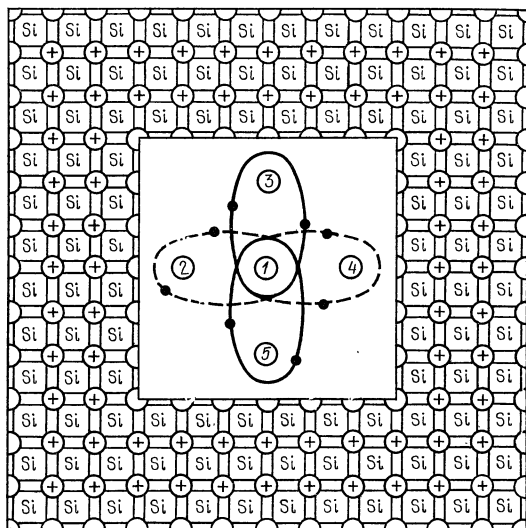


Рис. 4. Схематическое изображение кристаллической решетки кремния (Si). Черточки, связывающие друг с другом атомы Si, изображают электронные связи.

четыре валентных электрона, которые он может использовать для связи с соседними атомами. Из рис. 4 видно, что именно так и происходит: каждый атом связан с четырьмя ближайшими соседями электронными связями.

В центре рис. 4 электронные связи выбранного атома 1 с его четырьмя ближайшими соседями изображены несколько более подробно. Вокруг атома по замкнутым (а на самом деле имеющим сложную форму) орбитам вращаются восемь электронов.

Поля, удерживающие электроны на орбите, очень велики. Оценим, например, поле, действующее на

валентный электрон со стороны атома кремния. Порядковый номер кремния в таблице Менделеева — 14. Это означает, что ядро атома кремния содержит четырнадцать протонов. В целом атом нейтрален: число протонов равно числу электронов, а заряд протона равен по величине и противоположен по знаку заряду электрона. Валентных электронов, находящихся на внешней электронной оболочке кремния, как мы знаем, четыре. Остальные десять электронов размещаются на внутренних электронных оболочках и частично экранируют действие электрического поля ядра на валентные электроны. Таким образом, действие ядра на каждый валентный электрон эквивалентно притяжению со стороны четырех протонов.

Примем для оценки, что электрическое поле, создаваемое ядром и частично экранирующими его электронами внутренних оболочек, действует на валентный электрон как поле точечного заряда. Будем считать, что среднее расстояние, на котором валентный электрон находится от атома решетки, равно расстоянию между атомами (рис. 4). Тогда, исходя из закона Кулона, мы можем найти среднюю величину напряженности поля, удерживающего валентный электрон на орбите:

$$E \approx \frac{4e}{4\pi\epsilon_0 a_0^2},$$

где e — заряд электрона, $\epsilon_0 = 8,85 \cdot 10^{-12}$ Ф/м — электрическая постоянная, a_0 — расстояние между атомами кристаллической решетки. В большинстве твердых тел величина a_0 равняется нескольким долям нанометра ($1 \text{ нм} = 10^{-9} \text{ м}$). В кремнии $a_0 \approx 0,35 \text{ нм}$. Подставляя в формулу известные нам величины e , ϵ_0 и a_0 , получим $E \approx 2 \cdot 10^{10} \text{ В/м}$.

Очень важно то, что на примере кремния мы в действительности сумели оценить величину поля, действующего в пространстве между атомами практически любого кристалла, так как величина a_0 (ее называют *постоянная решетки*) приблизительно одна и та же для всех кристаллов; число валентных электронов у различных атомов также отличается только в несколько раз.

Чтобы лучше представить себе величину напряженности поля $E = 10^{10} \text{ В/м}$, скажем, что поле на-

пряженностью $E \approx 10$ В/м может раскалить и расплавить железный гвоздь (вычислите плотность тока в железной проволоке при $E = 10$ В/м), и напомним, что сам мечущий молнии Зевс-громовержец имеет в своем распоряжении всего лишь поля, приблизительно равные $3 \cdot 10^6$ В/м. Когда напряженность электрического поля между тучами и Землей или между двумя тучами достигает этого значения (оно называется электрической прочностью воздуха) — гремит гром и блещут молнии. И все-таки поля, при которых разгрызаются эти грандиозные явления природы, в 10 000 раз меньше, чем поля, которые удерживают электроны на их орбитах.

Ну а теперь займемся одним из самых важных и интересных для нас вопросом. Если к такому кристаллу, какой показан на рис. 4, приложить внешнее электрическое поле, возникнет ли электрический ток?

Ответ: нет, конечно, не возникнет. Такой кристалл представляет собой идеальный диэлектрик, так как даже очень сильное по обычным понятиям поле ($\sim 10^6$ В/м) лишь чуть-чуть деформирует электронные орбиты, но разорвать их окажется не в состоянии. Свободных носителей заряда — электронов (и дырок, но о них речь чуть впереди) в кристалле не окажется, следовательно, не будет и тока.

Заметим, что создать такую идеальную решетку, какая показана на рис. 4, совсем не просто. Чтобы решетка была действительно идеальной, необходимо исключить наличие каких-либо искажений, несовершенств, дефектов, которые могут разорвать электронные связи. Кроме того, в кристалле не должно быть абсолютно никаких примесей. И наконец, кристалл должен быть охлажден до температуры абсолютного нуля. В противном случае тепловые колебания решетки могут разорвать электронные связи и создать свободные носители тока. Все эти факторы: температура, примеси, дефекты — имеют большое значение, и мы будем в дальнейшем обсуждать их роль очень подробно. А пока отметим, что при температуре абсолютного нуля кристалл без примесей и дефектов, подобный показанному на рис. 4, действительно представляет собой идеальный диэлектрик.

На рис. 5 показана кристаллическая решетка другого типа. Упрощенное изображение кристалла меди

демонстрирует основные особенности металлических кристаллов. Как и ионы кремния, ионы меди образуют правильную кристаллическую решетку. Однако между кристаллическими решетками кремния и меди имеется принципиальное различие. Каждый атом меди при образовании решетки лишился одного валентного электрона и эти электроны (черные точки на рис. 5) не принадлежат теперь никакому конкретному иону меди. Они, как

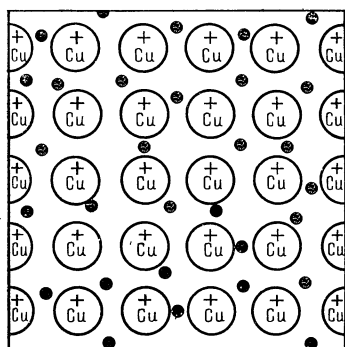


Рис. 5. Схематическое изображение кристаллической решетки меди (Cu). Правильная решетка ионов меди погружена в «газ» из свободных электронов, не связанных жесткими связями с отдельными ионами.

принято говорить, обобществлены кристаллом и могут свободно передвигаться под действием внешнего электрического поля. Даже при температуре абсолютного нуля электроны остаются свободными и металлы сохраняют способность проводить ток.

Итак, мы рассмотрели два типа кристаллов. Одни из них при температуре, близкой к абсолютному нулю, не проводят ток (диэлектрики). Другие — прекрасно проводят ток (металлы). Теперь, казалось бы, са-

мое время обратиться к рассмотрению кристаллов третьего, самого интересного для нас типа, — полупроводниковых кристаллов.

Но... напомним, что на рис. 4 изображена кристаллическая решетка самого типичного, самого распространенного и самого широко используемого полупроводника — кремния! И вблизи абсолютного нуля температуры такой кристалл, как мы отметили, является идеальным диэлектриком.

Этот момент кажется авторам самым подходящим для чистосердечного признания: никакой принципиальной, качественной разницы между диэлектриками и полупроводниками нет.

Модель кристалла, показанная на рис. 4, в равной мере относится и к типичным полупроводникам, и к типичным диэлектрикам. Разница между полу-

проводниками и диэлектриками скорее количественная. Она определяется величиной энергии, которая необходима, чтобы разорвать электронные связи между атомами. Чем больше величина энергии, которую нужно затратить, чтобы выбить электрон с его орбиты и превратить в свободный, способный двигаться под действием электрического поля, тем больше оснований считать, что мы имеем дело с диэлектриком.

Прежде всего оценим, какую приблизительно энергию нужно затратить, чтобы превратить связанный электрон в свободный. Мы рассчитали, что напряженность поля, удерживающего электрон на орбите, $E \sim 10^{10}$ В/м. Сила, действующая на электрон, $F = eE$. Чтобы удалить электрон с орбиты, необходимо, противодействуя этой силе F , «оттащить» электрон от притягивающего его иона на расстояние порядка постоянной решетки. Таким образом, энергия, необходимая для того, чтобы разорвать электронную связь и превратить электрон в свободный, по порядку величины равна $\mathcal{E}_g = eEa_0$. Индекс « g » у величины $\mathcal{E}_g$ происходит от английского слова gap — щель, зазор, интервал. В данном случае имеется в виду интервал энергий. Подставляя уже известные величины e , E и a_0 в эту формулу, получим $\mathcal{E}_g \approx 8 \cdot 10^{-19}$ Дж = 5 эВ *).

В действительности величина $\mathcal{E}_g$ зависит от структуры кристаллической решетки, от свойств входящих в нее атомов и у различных материалов лежит в пределах от десятых долей до нескольких десятков электронвольт. (Таким образом, наш грубый расчет дал возможность правильно оценить порядок необходимой энергии.)

Как всегда, когда речь идет о количественном, а не о качественном критерии, нельзя указать на четкую границу и сказать, например, что вещества, у которых значение $\mathcal{E}_g$ меньше определенной величины, называются полупроводниками, а те, у которых $\mathcal{E}_g$ больше — диэлектриками. Хотя попытки установить

*) Единица энергии 1 электронвольт (1 эВ) равняется $1,6 \cdot 10^{-19}$ Дж. Энергию в 1 эВ приобретает электрон (с зарядом $e = 1,6 \cdot 10^{-19}$ Кл), пройдя в электрическом поле разность потенциалов в 1 В. Когда речь идет о свойствах полупроводников, измерять энергию в электронвольтах оказывается очень удобным, и мы будем широко пользоваться этой единицей энергии.

такую классификацию в истории полупроводников известны, их все же следует скорее отнести к схоластическим упражнениям на тему «сколько волос образуют прическу» или «сколько песчинок образуют кучу». С другой стороны, если об одном человеке говорят: «какие у него прекрасные густые волосы», а о другом — «экая у него, прости господи, лысина», разногласия тоже возникают довольно редко. Так вот, помня об условности называемых чисел и о том, что не только от величины $\mathcal{E}_g$ зависит наличие полупроводниковых свойств у материала, принято все же ориентировочно считать, что кристаллы, у которых величина $\mathcal{E}_g$ лежит в пределах от десятых долей до 2—3 эВ, принадлежат к полупроводникам. Кристаллы с $\mathcal{E}_g \geq 3$ эВ в большинстве случаев будут вести себя как диэлектрики.

Величина $\mathcal{E}_g$ является одной из самых основных характеристик полупроводника. И в гл. 3 мы подробно обсудим, к каким важным следствиям приводит даже небольшое различие в значениях $\mathcal{E}_g$ у разных материалов.

А сейчас посмотрим, что произойдет, если, затратив энергию $\mathcal{E}_g$, нам удастся разорвать одну из электронных связей и создать свободный электрон.

Прежде всего, как это можно сделать? Один из возможных путей — осветить кристалл светом с подходящей длиной волны. Как известно, свет представляет собой поток световых квантов (фотонов), каждый из которых обладает строго определенной энергией $\mathcal{E}_\phi = h\nu = hc/\lambda$. В этой формуле h — постоянная Планка, равная $6,62 \cdot 10^{-34}$ Дж·с, ν — частота, λ — длина волны, c — скорость света, равная 3×10^8 м/с. Если энергия фотона $\mathcal{E}_\phi$ больше, чем энергия связи $\mathcal{E}_g$, световой квант способен выбить связанный электрон с его орбиты и превратить в свободный *).

*) Энергию фотона $\mathcal{E}_\phi$ удобно измерять в электронвольтах, а длину волны света λ — в микрометрах. Поэтому для удобства расчетов формулу для энергии фотона часто записывают в виде: $\mathcal{E}_\phi \approx 1,24/\lambda$. Получите эту формулу сами. Используйте ее для того, чтобы выяснить, какой свет: инфракрасный, видимый или ультрафиолетовый — необходим для образования свободного электрона в Ge ($\mathcal{E}_g = 0,72$ эВ), SiC ($\mathcal{E}_g = 2,3$ эВ), C (алмаз) ($\mathcal{E}_g = 5,6$ эВ).

Что при этом произойдет, показано на рис. 6. Одна из связей (между атомами 19 и 20) разорвана. Выбитый со своей орбиты электрон находится в данный момент в квадрате, образованном атомами 1, 2, 6 и 7.

Из рис. 6 сразу видно, что говоря об электро-не, выбитом из системы межатомных связей, как о сво-бодном, лучше бы всегда, по крайней мере в уме, ставить слово «свобод-ный» в кавычки. В самом деле, по-настоящему, сво-бодный электрон — это электрон в вакууме. «Сво-бодный» электрон в кри-сталле находится в слож-ном электрическом поле, образованном ионами ре-шетки и валентными элек-тронами соседних атомов. Под воздействием внеш-него электрического поля E свободный электрон в вакууме движется с по-стоянным ускорением $a = eE/m_0$ (m_0 — масса свободного электрона). «Свободный» электрон в кристалле под воздействи-ем внешнего поля может

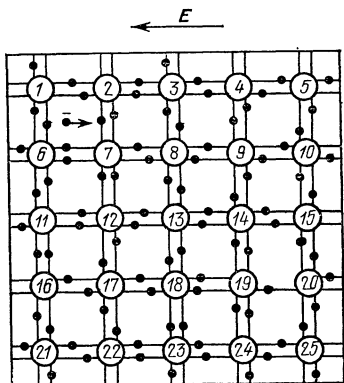


Рис. 6. Фотон выбил электрон с траектории, связывавшей атомы 19 и 20. Возникли пустая связь, дырка, и свободный электрон (он находится между атомами 1, 2, 6 и 7). В электрическом поле E электрон движется направо, а дырка — налево.

двигаться свободно только в течение очень короткого времени. Затем... электрон столкнется с колеблющимся под действием температуры атомом решетки, либо натолкнется на атом примеси или на дефект и т. д. Так что, употребляя термины «свободный электрон в кристалле», «свободный носитель заряда», следует помнить, что прилагательное «свобод-ный» означает не тождество со свободным электро-ном в вакууме, а способность направленно двигаться под действием внешнего электрического поля, обус-лавливая электрический ток.

А теперь наступило, наконец, время от главного героя вернуться к главной героине нашего романа, так невежливо и надолго нами оставленной, — дырке. На рис. 6 одна из связей между атомами 19 и 20

разорвана: один из электронов, осуществлявших связь, выбит с орбиты. Проследим за судьбой этой разорванной связи, сравнивая ее с поведением выбитого с орбиты свободного электрона. Если к кристаллу не приложено внешнее поле, электрон хаотически перемещается между атомами решетки под действием тепловых колебаний. Что в это время происходит с разорванной связью? Любой из электронов, связывающих атомы 19 и 20 с соседними атомами, может попасть на траекторию, с которой был выбит электрон, и восстановить полноценную связь между атомами 19 и 20. Если это произошло с электроном, связывавшим, например, атомы 14 и 19, то разорванная связь (будем для краткости называть ее дыркой) переместится в положение между ними. Следующим шагом может быть перемещение дырки в позицию между атомами 9 и 14 или 19 и 18, или 14 и 15; подобно свободному электрону, разорванная связь, дырка, может хаотически перемещаться между атомами решетки.

Если к кристаллу приложено внешнее электрическое поле, наряду с хаотическим движением свободного электрона возникает и направленное движение его против поля — электрический ток. Напомним, что электрон, как отрицательно заряженная частица, движется от «минуса» к «плюсу», т. е. *против* вектора напряженности электрического поля, который всегда направлен от «плюса» к «минусу». Что в такой ситуации происходит с дыркой? Снова возвратимся к рис. 6. И теперь, конечно, на траекторию между атомами 19 и 20 может, вообще говоря, переместиться любой электрон из связывающих атомы 19 и 20 с любым из их соседей. Но теперь, при наличии внешнего поля, наибольшую вероятность заместить разорванную связь имеют электроны, связывающие атомы 18 и 19. Именно эти электроны внешнее поле «тянет» на место разорванной связи. На хаотическое перемещение дырки накладывается направленное перемещение вдоль поля. Конечно, это не значит, что замещение обязательно произойдет электроном, связывающим атомы 18 и 19, а на следующем этапе дырка обязательно окажется между атомами 17 и 18. Внешнее электрическое поле не прекращает хаотического движения дырки так же, как оно не прекращает хаотического движения электрона. В обоих случаях

наличие поля приводит только к появлению в хаотическом движении частиц определенной направленности.

Еще раз подчеркнем: реально перемещаются электроны. Никакой реальной частицы «дырки» не существует (например, в отличие от электрона, дырку нельзя извлечь из кристалла и изучать ее поведение в вакууме). Дырка есть просто отсутствие электрона на траектории, связывающей друг с другом атомы кристалла. Но гораздо удобнее следить не за перемещением электронов с одной связи на другую, а за перемещением самой пустой связи — дырки.

Представим себе следующую простую аналогию (рис. 7). Мы смотрим сверху, с балкона старинного здания университета в Эстонии, на университетскую площадь, сплошь заполненную студентами. На площади негде яблоку упасть: люди стоят плечом к плечу. Все студенты пришли на площадь в форме, которая в Эстонии включает обязательную форменную фуражку, кроме одного разгильдяя, забывшего свою

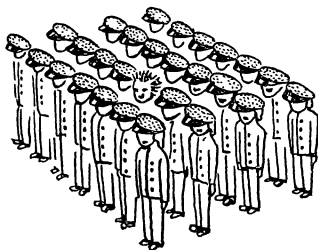
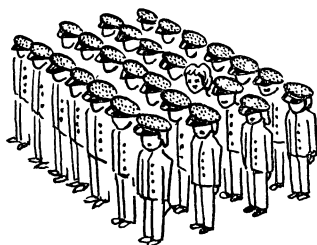
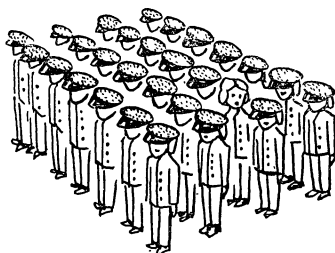
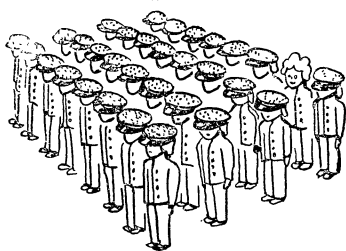


Рис. 7. Иллюстрация к понятию «дырка»: гораздо удобнее следить за одной непокрытой головой, чем за перемещением многих фуражек.

фуражку дома. Сверху мы ясно видим, как в море одинаковых форменных фуражек торчит непокрытая голова. Внезапно разгильдяй замечает, что он — единственный, кто явился на площадь с непокрытой головой. Пользуясь некоторой вольностью, допустимой среди студентов, он срывает фуражку с головы ближайшего зазевавшегося соседа и надевает ее. И... дает начало веселой игре. Каждый, кто оказался без фуражки, старается сорвать ее с головы кого-нибудь из соседей и надеть на себя. Сверху мы видим, как непокрытая голова — «дырка» — хаотически мечется по толпе студентов. И хотя мы ясно понимаем, что все студенты стоят неподвижно, и что реально перемещаются только фуражки, нам гораздо удобнее следить за одной непокрытой головой, чем за перемещением сотен фуражек. Допустив некоторую долю фантазии, мы можем продолжить эту аналогию. На балконе появляется ректор. И, о чудо! Под его строгим взглядом «дырка», не прекращая хаотических метаний, начинает заметно «уплывать» в задние ряды. На хаотическое движение наложилось направленное — вдоль непреклонного ректорского взгляда.

Итак, если каким-либо образом (например, с помощью кванта света) выбить один из электронов с орбиты межатомных связей, то у кристалла появляется не одна, а две возможности проводить ток. Первая — движение свободного электрона под действием внешнего электрического поля. Вторая — направленное перемещение электронов с соседних орбит на место освободившейся пустой связи. Этот второй механизм электропроводности удобнее описывать как движение фиктивной частицы, дырки, в направлении, противоположном движению электронов.

Обсудим теперь некоторые свойства дырок. Первый и самый главный вопрос: должны ли мы считать дырку заряженной частицей? На первый взгляд, если думать о дырке просто как о месте, где нет электрона, то ответ кажется очевидным: нет, не должны. Однако этот «первый взгляд» слишком поверхностный. Когда к кристаллу приложено электрическое поле, наличие свободной связи создает дополнительный механизм электропроводности. Электроны с соседних связей, последовательно перескакивая на место свободной, двигаются против поля и переносят ток. Если мы хотим описать этот механизм электропроводности,

используя понятие дырки, мы должны считать, что дырка (которая движется навстречу электронам — по полю) имеет *положительный* заряд, в точности равный по абсолютной величине заряду электрона.

Второй вопрос: следует ли ожидать, что электрон и дырка будут иметь одинаковые свойства? Ну, например, двигаться в одном и том же электрическом поле с одинаковой скоростью? Очевидно, для этого нет никаких оснований. В самом деле, просто взглянув на рис. 6, можно убедиться, что движение свободного электрона в межатомном пространстве кристалла и движение электрона с одной межатомной связи на другую, описываемое как движение дырки, протекают в совершенно различных условиях. Забегая вперед, скажем, что двигаться со связи на связь для электрона оказывается труднее — в том же внешнем электрическом поле дырка, как правило, движется медленнее, чем электрон.

И, наконец, последний вопрос. В кристалле, который мы рассматривали, электроны и дырки создаются всегда парами. Число электронов равно числу дырок. Пусть свет, который вызывает появление *электронно-дырочных пар*, падает на кристалл непрерывно. Каждый фотон, разрывая одну связь, порождает к жизни электрон и дырку. Подумаем над вопросом: неужели даже слабый свет, падающий на кристалл достаточно долго, разорвет все связи и превратит диэлектрик в металл? Так случилось бы, если бы, раз появившись, свободный электрон и дырка жили бы вечно. Но век электрона и дырки, созданных светом, короток. В разных полупроводниках этот век имеет разную продолжительность; он может лежать в пределах от 10^{-10} до 10^{-2} с. Как же погибают электрон и дырка? Хаотически блуждая по кристаллу, свободный электрон и дырка могут оказаться рядом. Свободный электрон занимает свое место на свободной траектории межатомной связи и... свободный электрон и дырка одновременно исчезают. Такой процесс называется *рекомбинацией* *).

При рекомбинации, конечно, выделяется энергия $\mathcal{E}_g$, которая была затрачена на создание электронно-дырочной пары. Иногда рекомбинация сопровождается

*) Рекомбинация в переводе с латыни означает воссоединение.

рождением фотона, и тогда рождается точно такой же квант света, который некогда погиб, породив электрон и дырку. Но довольно часто электрон и дырка рекомбинируют, не порождая фотона. Тогда энергия $\mathcal{E}_g$ отдается решетке, нагревая ее.

Ну что ж, первое знакомство с главными героями книги — электроном и дыркой — можно считать состоявшимся. Мы присутствовали при рождении героев, имели возможность чуть-чуть последить за их жизнью и движением, и даже «тайны гроба роковые» в некоторой степени приоткрылись для нас. Однако это очень поверхностное знакомство, и нам предстоит еще углублять и углублять его, прежде чем мы поймем, на что способны наши удивительные герои. Пока же, следуя старинному и мудрому правилу, познакомимся с их родословной: с историей полупроводников.

НА ПЛЕЧАХ ГИГАНТОВ

Если мы видели дальше других, то это потому, что стояли на плечах гигантов.

И. Ньютон

Не слушай учения тех мыслителей, доводы которых не подтверждаются опытом.

Леонардо да Винчи

«Полупроводники — сравнительно молодая наука — ей нет еще 30 лет. Она носит на себе черты нового времени», — эти слова написаны академиком Абрамом Федоровичем Иоффе в 1956 г.

«Английский физик Кавендиш опытами доказал, что вода проводит электричество в 400 миллионов раз хуже металла; невзирая на сие, она еще не совсем худой проводник электричества. Тела, кои в рассуждении способности проводить электричество, занимающие как бы среднее место между проводниками и непроводниками, обыкновенно называются полупроводниками», — эти слова написаны в 1826 г. автором учебника «Начальные основания опытной физики» Иваном Двигубским.

Последняя цитата замечательна во многих отношениях. Во-первых, из нее ясно, что уже 150 лет назад существовали вещества, которые «обыкновенно назывались» полупроводниками. Во-вторых, признаком таких веществ служило то самое качество, с которого начинается определение слова «полупроводники» в Большой Советской Энциклопедии. Взгляните на с. 8 нашей книги: «Полупроводники — класс веществ, характеризующийся значениями электропроводности, промежуточными между электропроводностью металлов и хороших диэлектриков...» Но самая интригующая часть цитаты — та, которая на первый взгляд выглядит наиболее естественно и невинно: «...Кавендиш опытами доказал, что вода проводит электричество в 400 миллионов раз хуже металла...»

Какими опытами? Каким способом живший в XVIII веке знаменитый физик и химик Генри

Кавендиш сравнил электропроводности металлов, диэлектриков и воды?

Немного воображения, и... мы легко перенесемся в старинный родовой замок герцогов Девонширских. В старинном кабинете Генри Кавендиш, одетый в камзол, в шелковых чулках и башмаках с пряжками сидит в высоком кресле работы Чиппендейла за старинным столом (рис. 8). На столе — старинный

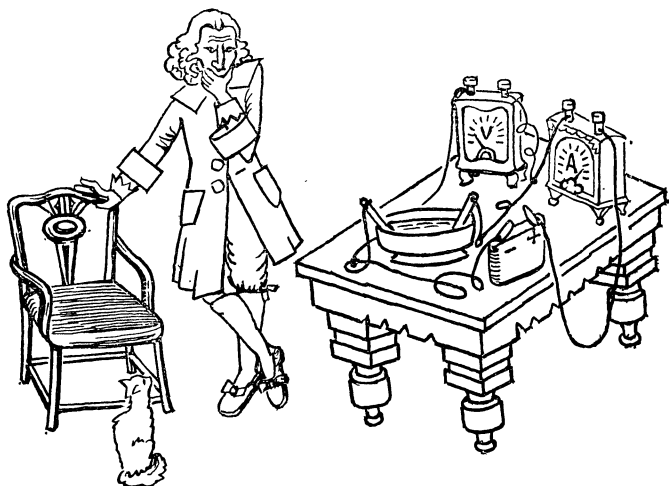


Рис. 8. Лорд Кавендиш обдумывает эксперимент.

гальванометр, старинный вольтметр и старинный аккумулятор. С их помощью лорд Кавендиш измеряет электропроводность воды.

Если поставить под этой картиной дату — 1776 г., когда Кавендиш сообщил о проведенных им опытах по измерению электрического сопротивления, и показать ее сведущему в истории физики человеку, то последствия предвидеть трудно.

Хорошо, если это окажется человек с крепкими нервами и чувством юмора. Тогда он просто расхохочется. И, может быть, вспомнит пародию Виктора Ардова «Исторические романы»: ...царь Иван Васильевич взялся за трубку старинного резного телефона с двуглавым орлом на деревянном коробе. — Боярышня, дайте мне царевича Ивана...»

Для слабонервного историка последствия могут быть печальными...

Первый аккумулятор был изобретен итальянским физиком Алессандро Вольта через 24 года — в 1800 г. (Помните, как это делается в мультфильмах? Фьють... — и аккумулятор со старинного стола исчезает.)

Первый гальванометр был изготовлен немецким физиком Вильгельмом Вебером в 1846 г., через 36 лет после смерти Кавендиша. (Фьють... со старинного стола исчезают вольтметр и гальванометр.)

Так каким же способом лорд Кавендиш «опытами показал»? Каким образом в 1776 г. можно было узнать о том, что существуют металлы, диэлектрики и полупроводники? Как появилась возможность изучать электрические свойства веществ?

Глава 1

DE PROFUNDIS *)

Отыщи всему начало и ты многое
поймешь.

Козьма Прутков

Многое твердо должно здесь быть
установлено прежде,
Нежели сможешь постичь ты пра-
вильно сущность предмета:
Надо к нему подходить и окольной
и длинной дорогой.

Лукреций Кар

Физика полупроводников развилась из хрупкой, почти незаметной веточки на мощном и древнем древе науки об электрических явлениях.

Янтарь, натертый шерстью, приобретает способность притягивать пылинки и соломинки, — традиция приписывает честь этого открытия греческому философу Фалесу из г. Милета, жившему в 640—550 гг. до н. э. Фалес считал, что в янтаре есть душа, от которой исходит притяжение, ибо только одушевленные предметы могут вызывать движение.

Янтарь у греков назывался «электрон», откуда и произошло название явления — электричество. Греки еще не заметили, что кроме янтара электризоваться при трении могут и другие тела. Впрочем, философ

*) Из глубины (лат.).

Теофраст, живший на 300 лет позже Фалеса, упоминал в своем сочинении «О камнях» еще один электризующийся камень: линкурион. Но что это за камень, нам неизвестно.

В Средиземном море водится рыба электрический скат, которую древние греки называли «наркэ» (парализующая). Сейчас известно, что напряжение, вырабатываемое этой рыбой, достигает 200 В. Римский врач Скрибоний, живший в начале новой эры, с помощью наркэ лечил головную боль, подагру и другие болезни, не подозревая, конечно, об электрической природе своего метода. Электричество, вырабатываемое скатом, было первым природным электрическим явлением, которое использовал человек, думая, однако, что целительное действие производит особый яд.

Что касается янтаря, то его способность притягивать пылинки вызывала не столько интерес, сколько раздражение: дорогие бусы античных дам постоянно были облеплены пылинками и соломинками...

Годы складывались в столетия. Возникали, расцветали и уходили в небытие целые цивилизации. Медленно, неуверенно, спотыкаясь и обращаясь вспять, человечество училось узнавать законы природы, осознавать и использовать их.

Во II веке до н. э. Архимед сформулировал законы плавания тел, создал основы механики. В III веке н. э. в Китае появился первый компас. В XI веке выдающийся ученый средневекового Востока Мухаммед ибн Ахмед аль-Бируни измерил плотность многих веществ и определил радиус Земли. В XV веке Иоганн Гутенберг изобрел книгопечатание, а в XVI веке Николай Коперник и Джордано Бруно доказали, что Земля вращается...

В области исследования электрических явлений за 2000 лет не было сделано ни одного открытия.

В последние годы XVI века лейб-медик английской королевы Елизаветы Тюдор, доктор медицины и магистр искусств Вильям Гильберт, известный своими познаниями в химии, физике и астрономии, занялся изучением магнитных явлений. Результатом его опытов и рассуждений явилось сочинение «О магните, магнитных телах и о большом магните — Земле», вышедшее в свет в 1600 г. и состоящее из шести книг. Во второй книге этого сочинения, вызвавшего

Наиболее выдающийся ученый Англии эпохи Возрождения Уильям Гильберт родился в 1544 г. в городе Колчестере, где его отец был главным городским судьей. Уильям учился в Кембридже и Оксфорде, а затем занялся врачебной практикой. В двадцать лет он стал бакалавром и магистром искусств, в двадцать пять — доктором медицины.

Гильберт был не только искусным врачом и преуспевающим лейб-медиком королевы Елизаветы Тюдор. Он был смелым и талантливым ученым, одним из тех, кто решительно поддержал учение Коперника, одним из основожателей нового естествознания, где главным критерием истины стал опыт.

Гильберт был первым естествоиспытателем, научно подошедшим к изучению электрических и магнитных явлений. Он впервые понял, что источником притяжения стрелки компаса является сама Земля, впервые ввел понятия магнитной оси, магнитных меридианов и параллелей, почти за 300 лет до работ П. Кюри установил, что магнитные свойства железа исчезают при нагревании.

Работы Гильберта ознаменовали собой рождение науки об электричестве. Гильберт установил, что не только янтарь, но и многие другие вещества притягиваются при их натирании; сделал первые попытки объяснить электрические явления.

В 1603 г. Гильберт умер, заразившись чумой. Созданные им приборы и большую библиотеку он завещал Коллегии врачей, однако большая часть материалов сгорела во время лондонского пожара 1666 г.

Самая крупная работа Гильберта «О магните, магнитных телах и большой магните — Земле» (М.: Изд-во АН СССР, 1956) сохранилась и стала одним из наиболее интересных памятников науки XVII века.



Уильям Гильберт
(1544—1603)

большой интерес, описывались опыты в области электричества. Пользуясь версором — прообразом будущего электроскопа (рис. 9), Гильберт показал, что не только янтарь, но и алмаз, сапфир, аметист, горный хрусталь, сера и другие тела, которые он назвал «электрическими», в результате натирания приобретают способность притягивать стрелку вер-

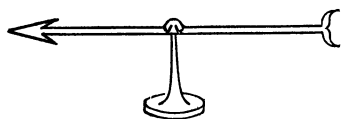


Рис. 9. Прообраз будущего электроскопа — версор — был изобретен, по-видимому, итальянским ученым Джироламо Фракастори в середине XVI века. Легкая стрелка, закрепленная на оси, поворачивается под действием даже очень не-
большой силы.

сора. Он обнаружил также, что жемчуг, мрамор, кость, металлы не способны электризоваться. Замечательной заслугой Гильберта следует считать его настойчивую пропаганду преимуществ экспериментального исследования перед софистическими упражнениями и многословным переливанием из «пустого в порожнее». «Если кто и не

признает возможным согласиться с высказанными мнениями и некоторыми парадоксами, — писал Гильберт, — то пусть усмотрит все-таки великое обилие опытов и открытий, кои достигнуты и доказаны многими трудами нашими и издержками. Ими будьте довольны и пользуйтесь, если можете, к лучшему употреблению».

Опыты Гильберта были повторены и продолжены в различных странах.

Важное изобретение было сделано в 1672 г. в Германии бургомистром г. Магдебурга Отто Герике. Он изготовил шар из серы диаметром примерно 15 см. Шар был насажен на металлическую ось, вращался с помощью рукоятки и натирался ладонями руки, давая электризацию значительно более сильную, чем в опытах Гильберта. Это была первая электрическая машина.

В начале XVIII века член Лондонского Королевского общества Стефан Грей, повторяя опыты Гильберта, заметил, что пушинки притягиваются не только к натертой стеклянной трубке, но и к шарику из слоновой кости, который был привязан к трубке пеньковой веревкой. Это явление заинтересовало Грея, и он стал увеличивать длину веревки. При

длине 26 футов (примерно 8,5 м) шар продолжал притягивать пушинки. Грею пришлось влезть на крышу — шар притягивал так, как будто был наэлектризован сам. Когда пеньковая веревочка оборвалась, Грей заменил ее шелковой нитью. Притяжение исчезло. Шелк, понял Грей, в отличие от пеньки, не «проводит» электричество. Когда вместо шелковой нити Грей взял металлическую проволоку, электризация стеклянной трубки вновь передалась шарик.

Доклад Стефана Грея Королевскому обществу вызвал бурный интерес и стимулировал новые исследования. Французский ученый Шарль Франсуа дю Фэй повторил опыты Грея и пришел к заключению, что не только вещества, указанные Гильбертом, но и все другие, кроме металлов и влажных тел, электризуются при трении (рис. 10). В 1739 г. профессор физики и теологии Оксфордского университета Ж. Т. Дезагюлье, повторив опыты Грея, ввел термины «проводник» и «изолятор», отразив этим способность тел передавать или не передавать электризацию на расстояние.

Электрические явления вызвали интерес не только у ученых. Благодаря электрической машине Герике и другим подобным машинам, электричество стало легко доступным. Сеансы демонстрации электрических явлений проводились повсюду: на площадях, в домах зажиточных горожан, при королевских дворах. Бродячие фокусники с радостью обнаружили, что на электричестве можно неплохо заработать: публика с любопытством наблюдала, как вспыхивает в чашке винный спирт от искры, выскочившей из пальцев «наэлектризованной» дамы. Электричество стало столь популярным, что считалось полезным электризовать воду, предназначенную для питья, для чего в банку с водой опускали конец проволоки от электрической машины.

Искры, свечение и треск, сопровождавшие работу мощных электрических машин, заставили наиболее пытливых ученых с сомнением поглядывать на небеса. Уж не являются ли молнии, «огненные стрелы Зевса Кронида», гигантскими искрами чудовищной природной электрической машины? Уже знакомый нам французский ученый Шарль дю Фэй высказал по этому поводу чрезвычайно смелую для того времени (1735 г.) идею: «Возможно, что в конце

концов удастся найти средства для получения электричества в больших масштабах и, следовательно, усилить мощь электрического огня, который во многих опытах представляется как бы одной природой с громом и молнией».

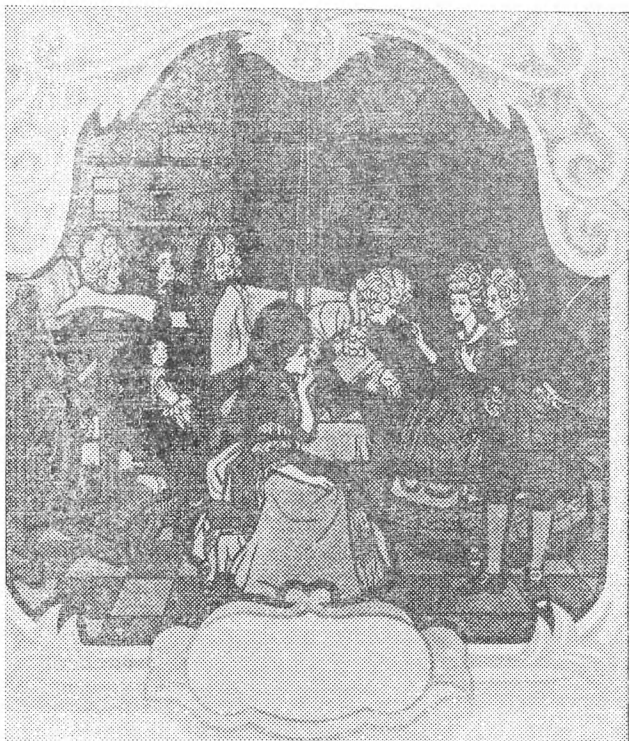


Рис. 10. Франсуа дю Фэй демонстрирует явление электризации (XVIII век). Пятки дю Фэя, обтянутые шелковыми чулками, электризуются вращающимся стеклянным шаром. Между носом ученого и пальчиком дамы проскакивает искра.

Перенесемся теперь за много тысяч километров от Европы, через океан, в Северную Америку. Блестящий политик и дипломат, один из лидеров борьбы за независимость Штатов от Англии, один из авторов первой Конституции США и петиции об уничтожении рабства, один из самых популярных американских писателей своего времени, Бенджамин

Франклин в возрасте сорока лет увлекся исследованием электрических явлений. Франклину принадлежит много важных открытий в области электричества, но главной его заслугой, безусловно, является установление электрической природы молнии.

Как и дю Фэй, Франклин заметил сходство между искрой и молнией, между потрескиванием электрической машины и раскатами грома. Решающий опыт Франклина, позволивший установить природу молнии, был гениально прост. В то время, когда низкие тучи закрыли небо и гроза вот-вот готова была разразиться, Франклин запустил бумажного змея.

Змей был снабжен металлическим острием, а к концу бечевки, у земли, был привязан металлический ключ. Бечевка намокла от дождя, и когда к ключу подносили металлический предмет, из него с треском выскакивала искра. Природное электричество, скопившееся в тучах, вызывало точно такие же явления, как и уже хорошо знакомые исследователям электрические машины.

Практичный американец предложил использовать грозоотводы — высоко расположенные, изолированные от зданий и заземленные металлические острия для защиты домов от удара молний. Сейчас, когда грозоотвод стал стандартной принадлежностью любого сооружения, трудно представить себе, какой накал страстей вызвало это предложение Франклина и в какие яростные стычки выливались споры между сторонниками и противниками грозоотвода.

Франклин внес также важный вклад в исследование первого электрического конденсатора — знаменитой «лейденской банки».

В 1745—1746 гг. настоятель одного из соборов Померании Юрген фон Клейст и профессор философии из Лейдена Мушенброк почти одновременно сделали замечательное открытие. Вот что писал Мушенброк одному из своих корреспондентов: «Хочу сообщить новый, но ужасный опыт, который советую Вам никак не повторять. Я делал некоторые исследования над электрической силой и для этой цели повесил на двух шнурках из голубого шелка железный ствол, получавший электричество через сообщение от стеклянного шара, который быстро вращался и натирался руками. На другом конце висела медная проволока, конец которой был погружен в круглый

стеклянный сосуд, отчасти наполненный водою. В правой руке я держал этот сосуд, другою же рукою пробовал извлечь искры из наэлектризованного ствола. Вдруг моя правая рука была поражена с такой силой, что все тело содрогнулось как от удара молнии... я думал, что пришел конец *)).

Совету Мушенброка «никак не повторять» этот «ужасный опыт» никто не последовал. Напротив, физиологическое действие электричества вызвало интерес необычайный и побудило к дальнейшим опытам людей самых разных вкусов и различного общественного положения.

Его величество король Франции Людовик XV пропускать через цепь солдат разряд лейденской банки и забавлялся, наблюдая гримасы бедняг. А один из вождей французской революции врач Жан Поль Марат написал сочинение «Насколько и в каких условиях можно рассчитывать на электричество как на положительное в лечении болезней», за что в 1783 г. получил золотую медаль Руанской академии. Но наибольшее значение открытие лейденской банки имело, конечно, для физиков: теперь в их распоряжении был мощный и надежный источник электричества.

В 1761 г. другу Франклина Эбенезеру Кеннерсли удалось нагреть до красного каления проводник, через который прошел разряд лейденской банки. Еще через несколько лет Джозеф Пристли высказал мысль о том, что плавление различных металлов под действием искры может быть использовано для сравнения той «легкости», с которой они проводят электричество. Идея Пристли была подтверждена опытами Джамбатисты Беккариа в Италии и Джона Кантона в Англии. И, наконец, в 1776 г. один из гениальнейших и удивительнейших ученых XVIII века Генри Кавендиш проводит (ровно за 50 лет до открытия закона Ома!) систематические измерения относительного удельного электрического сопротивления различных веществ.

Теперь, снова обратившись к рис. 8, мы можем убрать со старинного стола аккумулятор, вольтметр и гальванометр и поместить на их место те приборы, которые были использованы Кавендишем в его заме-

*) Подробно об истории открытия конденсатора — лейденской банки — можно прочитать в статье: Кикоин А. К. — Квант, 1971, № 9.

Бенджамин Франклин родился 17 января 1706 г. в американском городе Бостоне. Он был восьмым ребенком в большой семье эмигранта из Англии Жозефа Франклина — ремесленника, занимавшегося изготовлением мыла и свечей. Всего в семье Жозефа было 17 детей, поэтому уже с 10 лет Бенджамин стал сам зарабатывать свой хлеб.

Необычайные и разносторонние способности Франклина проявились с ранних лет. Он настойчиво занимался самообразованием, тратя на книги каждый свободный цент. Поработав в нескольких мастерских в Бостоне и Нью-Йорке, Франклин переехал в Филадельфию и организовал там собственную типографию. Сделавшись человеком состоятельным, Франклин увлекся общественной деятельностью и не оставлял ее до последних дней жизни.

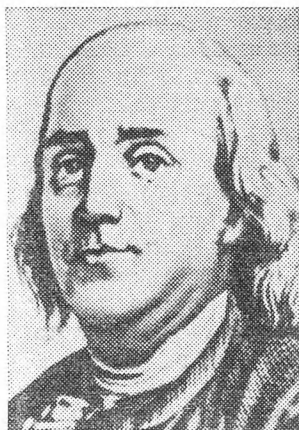
Природный талант Франклина проявлялся и в его научных занятиях. В 1746 г., случайно попав на сеанс «чудес физического кабинета» доктора Спенса, Франклин впервые познакомился с электрической машиной и увлекся электрическими опытами.

В науке об электричестве Франклину принадлежат несколько замечательных открытий: он объяснил действие лейденской банки, сконструировал плоский конденсатор со стеклянной изоляцией, открыл электрическую природу молнии, изобрел первый грозоотвод.

Франклин был избран членом Лондонского Королевского общества и удостоен его высшей награды — медали Коплея. В 1789 г. Франклин был избран почетным членом Российской академии наук.

Америка XVIII века не знала человека, столь всесторонне одаренного и деятельного как Б. Франклин. Писателя и издателя, ученого, общественного деятеля и дипломата, его называли «первым среди цивилизованных американцев».

Франклин умер 17 апреля 1790 г. в возрасте 84 лет. Траур по выдающемуся ученому продолжался в США 30 дней. Решением Всемирного Совета Мира имя Бенджамина Франклина, «которого гений был красой человеческой природы и различные стремления которого послужили на пользу науке, свободе и отечеству» *) включено в список наиболее выдающихся представителей Человечества **).



Бенджамин Франклин
(1706—1790)

*) Из постановления Конгресса США о 30-дневном трауре по случаю смерти Франклина.

**) О научной деятельности Б. Франклина можно более подробно прочитать в статье: *Капица П.* — *Квант*, 1981, № 7.

чательных опытах: лейденские банки и электрические машины для их заряда (рис. 11).

«Он брал две проволоки разных металлов одинаковой длины и сечения и разряжал через них батарею лейденских банок. Он повторял этот эксперимент, изменяя заряд батареи, до тех пор, пока одна из проволок не расплавлялась. Полагая, что проволока, оставшаяся целой, обладает большей электрической проводимостью, он получил ряд металлов, расположенных в порядке уменьшения их проводимости».

Не правда ли, как просто?! И это — как раз тот самый случай, когда добавление «гениально просто» не будет выглядеть преувеличением!

Итак, проследив путь, которым 180 лет шли естествоиспытатели, изучавшие электричество, мы теперь представляем, как могло появиться понятие о проводниках и изоляторах. И «полупроводниках» — веществах, о которых пока что ничего определенного сказать нельзя. Они слишком хорошо проводят электричество, чтобы использовать их в качестве изоляторов, и слишком плохо, чтобы служить проводниками. Не правда ли, трудно ожидать, что именно эта группа веществ почему-либо особенно интересует исследователей?

Наступление XIX века отмечено в истории электричества замечательным открытием итальянского физика Алессандро Вольта. В 1800 г. Вольта сообщил, что столб, составленный из чередующихся серебряных и цинковых дисков, переложенных влажными картонными кружками, способен служить источником постоянного электрического тока в течение многих часов. Это открытие позволило наблюдать множество явлений, которые было очень трудно заметить за время молниеносного проскакивания искры, созданной лейденской банкой или электрической машиной.

В 1819 г. датский физик Ханс Кристиан Эрстед, используя в качестве источника тока «вольты столб», обнаружил, что магнитная стрелка отклоняется, если рядом с ней находится проводник, по которому течет ток. Открытие Эрстеда, а затем блистательные опыты французского ученого Андре Мари Ампера*), продемонстрировавшие тесную связь между

*) О жизни Эрстеда и Ампера можно прочитать в книге: Замечательные ученые. — М.: Наука, 1981. Библиотечка «Квант», вып. 9.

Знаменитый английский химик и физик Генри Кавендиш родился во Франции 10 октября 1731 г. Это был удивительный человек необычной судьбы. Кавендиш был старшим сыном лорда Чарльза Кавендиша и принадлежал к очень знатной, но не блестящей состоятельной семье.

В возрасте 40 лет Генри Кавендиш неожиданно получил огромное наследство и стал одним из богатейших людей Англии. Основную часть средств Кавендиш расходовал на организацию научных опытов и помощь молодым ученым. В обществе Генри Кавендиш практически не бывал, с родственниками почти не общался. Не желая отвлекаться от размышлений и экономя время для научной работы, этот необычный человек объяснялся с домашними знаками.

Кавендиш был гениальным ученым, предвосхитившим многие выдающиеся открытия в физике и химии. Он прославился исследованиями свойств газов, впервые получил водород действием соляной кислоты на цинк (и железо), доказал, что вода имеет сложный состав, определил содержание кислорода в воздухе. Кавендиш на 12 лет раньше Кулона открыл закон взаимодействия электрических зарядов, но нигде не опубликовал своих результатов. За 65 лет до Фарадея он исследовал влияние диэлектрика на емкость конденсаторов, изготовил уникальную батарею конденсаторов для измерения емкостей, измерил относительные удельные сопротивления многих веществ задолго до открытия закона Ома.

Кавендиш рассчитывал и изготавливал метеорологические приборы, приборы для анализа химического состава газов, создавал приборы для мореходной астрономии и др. Девизом Кавендиша было: «Все определяется мерой, количеством и весом». Удивительна та точность, с которой он ставил свои опыты в те далекие времена, когда практически не было измерительных приборов. Закон взаимодействия электрических зарядов был проверен им с точностью до 0,025%, содержание кислорода в воздухе найдено равным 20,83% (современное значение 20,93%), плотность Земли — $5,45 \text{ г/см}^3$ (по современным данным $5,52 \text{ г/см}^3$)!

Умер Г. Кавендиш в 1810 г. в Лондоне.

Кавендиш не публиковал многих своих работ, особенно в области электричества. Более 100 лет они пролежали в виде рукописей, пока в 1879 г. Максвелл не изучил и не опубликовал его труды, после чего они получили широкое и заслуженное признание *).



Генри Кавендиш
(1731—1810)

*) Подробнее о Генри Кавендише можно прочитать в статье: *Филонович С.* — Квант, 1981, № 10,

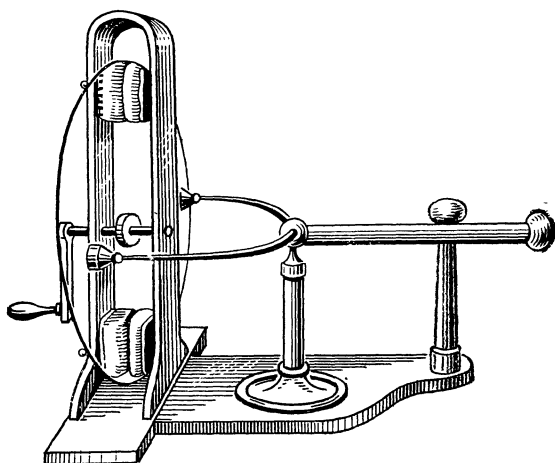
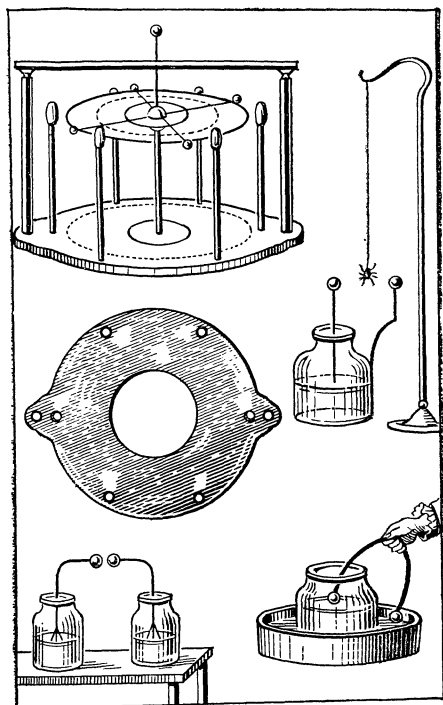


Рис. 11. Образцы лейденских банок и электрических машин.

электрическими и магнитными явлениями, побудили к постановке тысячи экспериментов. Один из них, не вызвавший поначалу особенной сенсации, с точки зрения истории полупроводников особенно интересен.

В 1821 г. немецкий физик Томас Зеебек, увлеченный опытами Эрстеда и Ампера, также занялся целью исследовать «магнитную атмосферу электрического тока». Схема опыта Зеебека показана на рис. 12.

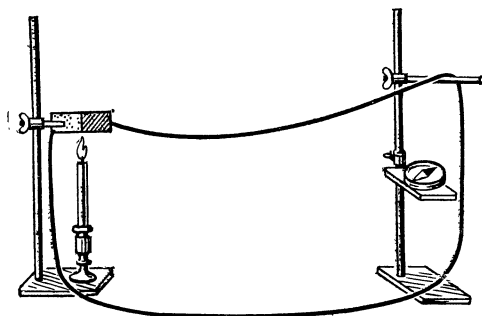


Рис. 12. Схема опыта Зеебека.

Зеебек припаивал друг к другу два разнородных металла, соединял их медным проводником и помещал внутрь петли, образованной проводником, магнитную стрелку. Нагревая место спая с помощью свечи, Зеебек обнаружил, что магнитная стрелка, расположенная вблизи проводника, отклоняется. Нагревание спая двух разнородных металлов порождает электрический ток! С необыкновенной тщательностью и трудолюбием Зеебек исследовал открытый им эффект, экспериментируя с сотнями различных веществ. При этом обнаружилось, что, когда одним из элементов спая служили теллур, сульфид свинца или некоторые другие материалы (которые через 100 лет будут названы ПОЛУПРОВОДНИКАМИ), отклонение стрелки резко возрастало. В последующие годы источники тока, основанные на эффекте Зеебека, широко использовались физиками-экспериментаторами. Для того чтобы увеличить силу тока, соединяли много десятков или даже сотен спаев (рис. 13), составляя конструкции,

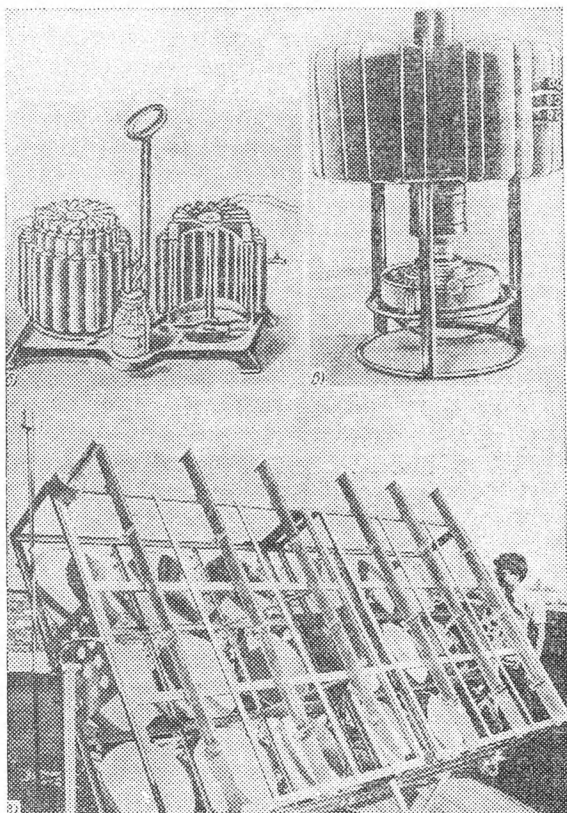
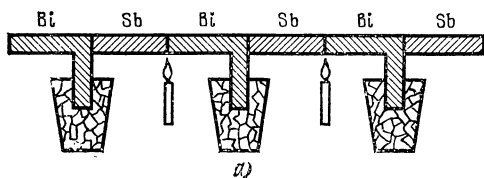


Рис. 13. Термобатареи.

а) Первая термобатарея, предложенная Эрстедом и Фурье. Горячие спаи висмута (Bi) и сурьмы (Sb) нагреваются свечами. Холодные спаи опущены в сосуды со льдом; *б)* батарея Ноэ (70-е годы XIX века). Горячие спаи нагреваются бунзеновской горелкой, охлаждение холодных спаев естественное. К. п. д. $\sim 0,1\%$; *в)* термобатарея ТКГ-9 (50-е годы XX века). Горячие спаи нагреваются керосиновой лампой, охлаждение холодных спаев естественное. Термобатарея (термоэлектрогенератор) всего 30 лет назад служила источником питания для бытовых радиоприемников в неэлектрифицированных районах нашей страны. К. п. д. $\sim 3\%$; *г)* современная солнечная батарея. Вогнутые отражатели концентрируют солнечное излучение на элементы, основу которых составляет контакт арсенида галлия с тройным полупроводниковым соединением — GaAlAs. К. п. д. $\geq 20\%$.

которые назывались «термобатареями» *). Естественно, казалось бы, ожидать, что в первую очередь в качестве элементов этих батарей будут использоваться вещества, которые обеспечивают наибольшую силу тока. В свою очередь, это могло бы повести к углубленному исследованию свойств таких веществ и, кто знает, может быть история полупроводников, электроники и многих других дисциплин сложилась бы совсем по-другому. Однако, как говорит замечательная греческая пословица: «Прошлого не могут изменить даже боги», — увы, этого не произошло.

В термоэлектрических источниках ценилась прежде всего их стабильность и надежность. Гораздо проще было удвоить при необходимости число металлических слоев, чем разбираться, почему в одном случае использование сульфида свинца приводит к замечательным результатам, а в другом — результаты оказываются никуда не годными. Во второй части книги мы познакомимся с поразительной чувствительностью полупроводников к наличию примесей и загрязнений и тогда, конечно, без труда сможем себе представить, какой «стабильности», «воспроизводимости» и «чистоты» можно было ожидать от полупроводниковых материалов 150 лет назад!

Первый сигнал о существовании класса веществ с необычными свойствами остался незамеченным.

Второй сигнал был обнаружен великим английским физиком Майклом Фарадеем. Учитель Фарадея, знаменитый химик Хэмфри Дэви, еще в 1821 г. установил, что проводимость металлов уменьшается с ростом температуры. В 1833 г. Фарадей, продолжая опыты Дэви по изучению зависимости электропроводности различных веществ от температуры, столкнулся с необычной ситуацией. Исследуя температурную зависимость электропроводности сернистого серебра, которое

*) Зеебек назвал открытый им эффект «термомагнетизмом». Теперь эффект Зеебека и ряд других связанных с ним эффектов называют термоэлектрическими эффектами. В 1940 г. в Ленинградском Физико-техническом институте под руководством Юрия Петровича Маслакова была собрана и испытана первая экспериментальная полупроводниковая термобатарея. Материалом, из которого изготавливались термоэлементы, служил сульфид свинца.

все в то время считали металлом, Фарадей установил, что электропроводность его не падает, а возрастает с ростом температуры. Вот как описывает Фарадей результаты этого опыта в своих знаменитых «Экспериментальных исследованиях по электричеству»: «Я не знаю ни одного вещества, которое, подобно сернистому серебру, может в горячем состоянии сравниться с металлами в отношении проводимости электричества, и у которого, наряду с этим, в отличие от металлов, проводимость при охлаждении уменьшается, тогда как у металлов она, наоборот, увеличивается. Однако если поискать, то, вероятно, можно будет найти немало таких веществ».

В течение последующих пяти лет Фарадей обнаружил еще несколько «таких веществ» (фторид свинца, PbF_2 , сульфид ртути, HgS , и ряд других), также обладавших необычной зависимостью проводимости от температуры.

К 1838 г., когда Фарадей завершил измерения температурной зависимости электропроводности, он был уже всемирно известным ученым, прославившимся открытием законов электромагнитной индукции и электролиза, членом Лондонского Королевского общества, директором лабораторий Британского Королевского института, почетным профессором химии. Все свои результаты Фарадей регулярно докладывал и публиковал. Таким образом, и результаты, полученные Фарадеем, и его предсказание о возможном существовании целого класса веществ с необычными электрическими свойствами были доступны и известны его современникам. И что же, многие ли исследователи приняли предложение Фарадея «поискать» и занялись изучением нового класса веществ?

История не знает ни одного сколько-нибудь серьезного продолжения работ Фарадея в течение нескольких последующих десятилетий. Никто не последовал заманчивому призыву знаменитого ученого найти и исследовать новый класс веществ. Второй сигнал также остался незамеченным.

Что это — досадная случайность, обидный исторический парадокс?

Напротив, — проявление общего закона истории науки: бурное и плодотворное развитие любой области естествознания происходит только тогда, когда, как писал один из самых замечательных ученых

Майкл Фарадей родился 22 сентября 1791 г. в семье небогатого лондонского кузнеца. С 12 лет он был вынужден работать подмастерьем переплетчика и рассыльным в книжном магазине. Глубокие знания в области физики и химии Майкл приобрел путем упорного самообразования. Когда Фарадею было 19 лет, произошел случай, во многом определивший его судьбу: юноше предложили билеты на цикл лекций профессора Королевского института сэра Хэмфри Дэви. Майкл не только прослушал весь курс, но и законспектировал его. А затем написал сэру Дэви письмо, приложив к нему переплетенные им самим конспекты, и был принят ассистентом в Королевский институт.



Майкл Фарадей
(1791—1867)

Первые опыты по электричеству Фарадей провел в возрасте 22 лет вместе с сэром Дэви, изучая мощные импульсы, создаваемые электрическим скатом. И вплоть до последних дней жизни, занимаясь самыми различными проблемами, Фарадей продолжал работать над разгадкой тайны электричества.

Один из самых гениальных и знаменитых физиков всех времен Майкл Фарадей — автор многих замечательных работ в области физики и химии. Среди его открытий — закон электролиза, закон, устанавливающий связь между магнитными и электрическими явлениями, открытие диа- и парамагнетизма, открытие вращения плоскости поляризации света в магнитном поле.

Майкл Фарадей был первым, кто экспериментально обнаружил эффект резкого увеличения проводимости с температурой у некоторых веществ (сульфид серебра, фторид свинца) и понял, что в основе этого эффекта лежит глубокое физическое явление, указывающее, возможно, на существование нового класса веществ. Теперь мы называем эти вещества полупроводниками.

Скончался Фарадей 25 августа 1867 г. Ученые всего мира чтут память Майкла Фарадея не только как гениального физика и химика, замечательного экспериментатора и педагога, но и как человека, которому были присущи строжайшая научная щепетильность, добросовестность и скромность.

«...Внешние признаки явлений не должны связывать суждений ученого. Истина должна быть главной целью его исследований. Если к этому добавится трудолюбие, то он может надеяться приподнять завесу в храме природы...» (Майкл Фарадей)

Л и т е р а т у р а

Мак-Дональд Д. Фарадей, Максвелл и Кельвин. — М.: Атомиздат, 1967.

Фарадей М. История свечи. — М.: Наука, 1980. Библиотечка «Квант», вып. 2.

прошлого века Джеймс Клерк Максвелл *), «у практики существует потребность в знаниях».

Применительно к полупроводникам эта потребность сформировалась лишь через несколько десятилетий после работ Фарадея. Случай, который привел в действие «потребность в знаниях», произошел через 40 лет после публикаций Фарадея и явился следствием развития ... дальней телеграфной связи.

Модели телеграфных аппаратов, позволяющих передавать сообщения на довольно большие расстояния, стали появляться одна за другой сразу же после открытия Эрстедом магнитного действия электрического тока. Первый использованный на практике аппарат предложил вскоре после сообщения Эрстеда русский изобретатель П. Л. Шиллинг. В течение 30-х годов прошлого века электромагнитные телеграфы были построены в Германии, Англии, Америке.

В 1844 г. нью-йоркский художник Морзе, увлекшийся электротехникой, изобрел аппарат с движущейся бумажной лентой, на которой притягивающийся электромагнитом карандаш оставлял точки и тире. Это изобретение стало поворотным пунктом в истории почтовой связи: во всем мире телеграф начал решительно теснить почтовые дилижансы. К 70-м годам XIX века телеграф распространился повсеместно. Образовалась сеть телеграфных компаний, возникла специальность инженер-телеграфист, появились даже Общества телеграфистов.

В 1873 г. инженер-электрик из Лондона Уиллоби Смит занимался испытаниями подводного телеграфного кабеля. Для изоляции кабеля Смит решил использовать селен. Селен, открытый в начале века Берцелиусом, будучи расплавленным, а затем быстро охлажденным, застывает в стекловидную массу с очень большим сопротивлением. Эта масса и использовалась в качестве изоляции для кабеля. Наблюдательный помощник Смита Мей заметил, что на свету сопротивление селена становится значительно меньше, чем в темноте. Случайное наблюдение, о котором сообщил Смит, вызвало целый шквал экспериментов. В том же году открытие было подтверждено не менее

*) О Максвелле можно прочитать в статье: *Сморodinский Я. А.* — Квант, 1981, № 11.

чем десятью физиками; было установлено, что селен чувствителен даже к слабому свету Луны.

Как оказалось, способность селена преобразовывать свет в электрические сигналы отвечала насущнейшим практическим потребностям. Селеновые фотосопротивления почти немедленно стали использоваться в различных оптических приборах *). Возникла необходимость объяснить причину фоточувствительности селена, начались поиски новых светочувствительных материалов. На рубеже XX столетия физики начали специально изучать материалы, которые не были ни металлами, ни диэлектриками.

И вот тут вспомнили и о необычно большом термоэлектрическом эффекте в некоторых материалах, обнаруженном Зеебеком почти 70 лет назад, и об «аномальной» температурной проводимости, открытой Фарадеем. Обратили внимание на открытие Фердинанда Брауна, сделанное в 1874 г.: контакт сернистого свинца с металлом имеет малое сопротивление при одном направлении протекающего через него тока и очень большое — при противоположном.

Число ежегодных публикаций, посвященных исследованию необычных материалов, начало исчисляться десятками, число работ по исследованию приборов на основе таких материалов — сотнями. За полупроводники взялись всерьез.

Однако процесс исследования полупроводников меньше всего напоминал парадный марш стройных колонн армии-победительницы. Картина скорее напоминала ярмарку: от обилия самых разнообразных и противоречивых явлений рябило в глазах. Голоса различных исследователей никак не могли слиться

*) Грэхем Белл (который в 1876 г. получил патент на изобретение первого в мире телефона) изготовил на основе селена «фотофон» — прибор, передававший звук на большое расстояние при помощи светового луча.

Гибкое плоское зеркало освещалось сильным источником света так, что отраженный луч попадал на приемной станции на линзу, фокусировавшую свет на селеновую пластинку, соединенную с батареей и телефоном. Когда на передающей станции на заднюю поверхность зеркала направлялся звук человеческой речи — звук отчетливо воспроизводился на приемном телефоне. Фотофон просуществовал недолго, он не выдержал конкуренции с изобретенным тем же Беллом телефоном. Однако через 100 лет на совершенно ином техническом уровне идея передачи звука с помощью светового луча нашла воплощение в лазерных телефонных системах.

в согласный хор. Каждый выкликал свои результаты, и результаты эти отличались в такой степени, что с трудом открытые полупроводники сгоряча чуть было не закрыли. Об этом трудном времени в истории полупроводников мы расскажем в следующей главе.

А пока заметим, что несмотря на все эти трудности, уже в самом начале XX века сформировались правильные представления о природе электропроводности полупроводников. В значительной мере этот успех обязан работам замечательного немецкого физика И. Кенигсбергера.

В 1906 г. Кенигсбергер опубликовал работу, в которой писал: «При повышении температуры в оксидах и сульфидах число проводящих или свободных квантов электричества — электронов — увеличивается, пока не станет предельным, после чего их поведение уподобляется металлам, в которых при нормальной температуре почти все электроны свободны».

Прошло еще 8 лет, и в 1914 г. тот же Кенигсбергер опубликовал первый в истории обзор по свойствам полупроводников. В этом обзоре, подводя итоги многочисленным экспериментам последних лет, в которых видное место занимали и его собственные исследования, Кенигсбергер вводит, наконец, понятие «класса полупроводников»: «Полупроводниками... будут называться проводники с металлической проводимостью, сопротивление которых очень сильно изменяется с температурой...» Характеризуя количественно зависимость проводимости от температуры, Кенигсбергер предлагает зависимость, которая, как мы увидим во второй части книги, действительно правильно описывает одно из самых важных свойств полупроводников:

$$\sigma = Ae^{-q/T} \quad (1)$$

([σ — проводимость, T — температура, A и q — константы]).

Помимо характерной температурной зависимости проводимости, класс полупроводников с точки зрения Кенигсбергера характеризуется еще несколькими основными свойствами: величиной удельного сопротивления в пределах от 10^{-8} до 10 Ом·м; большими значениями термо-э.д.с. относительно металлов, эффектом выпрямления тока (или невыполнением закона Ома на контактах), чувствительностью к свету,

Слово «полупроводник» осталось тем же, что и во времена Кавендиша. Но как углубилось и прояснилось понятие, обозначаемое этим словом! Вместо одного достаточно произвольного признака: не слишком хорошей, не слишком плохой электропроводности — целая совокупность свойств, позволяющая выделить новый класс веществ. Объяснение одного из основных свойств — температурной зависимости проводимости, дано с точки зрения внутренних процессов, происходящих в кристалле. И, наконец, этот новый класс веществ сулит новые и заманчивые практические применения!

Глава 2

PER ASPERA AD ASTRA *)

Так как теперь это все
установлено твердо и точно,
И основания все подготовлены
нами как должно,
То остальному уже не трудно
найти объяснение.

Лукреций Кар

Бди!

Козьма Прутков

Предыдущая глава закончилась оптимистическим аккордом — полупроводники выделены в особый класс веществ. Они интенсивно изучаются, находят важные практические применения. Кажется бы, есть все основания ожидать еще более уверенного и мажорного продолжения...

В 1935 г., через 21 год после появления в печати первого обзора Кенигсбергера, другой ученый, один из самых в то время известных исследователей в области полупроводников, Б. Гудден опубликовал обширный обзор «Электропроводность электронных полупроводников». Вот выдержки из этого обзора:

«Полупроводников в кенигсбергеровом смысле не существует...»; «...продолжительные и обширные исследования ясно указывают, что в этом случае**) имеют место весьма запутанные соотношения...»

*) Через тернии — к звездам (лат.).

**) Речь идет о проводимости полупроводников и ее зависимости от температуры.

«Металлы, как, например, графит, КРЕМНИЙ, титан, цирконий и т. д. ни в коем случае не следует смешивать с электронными полупроводниками...»

Читая обзор Гуддена, легко понять, почему в 1956 г. А. Ф. Иоффе, прекрасно знавший историю полупроводников, имел все основания написать фразу, которой начинается эта часть книги: «Полупроводники — сравнительно молодая наука — ей нет еще 30 лет...»

Все ясно, — скажет торопливый читатель. — Старые применения полупроводников утратили ценность, а новые изобретения не появлялись. Интерес к полупроводникам постепенно угас...

Ничуть не бывало! С развитием радиосвязи широкое распространение получили сконструированные замечательным советским ученым Олегом Владимировичем Лосевым кристаллические полупроводниковые детекторы (начало 20-х годов). В 1926 г. американец Грондаль создал технически совершенный выпрямитель на основе полупроводникового соединения — закиси меди. Такие выпрямители нашли применение для зарядки электрических батарей, для питания телефонных систем на транспорте. А в 1932 г. на основе закиси меди был сконструирован твердотельный фотоэлемент — прибор, позволивший эффективно преобразовывать световую энергию в электрическую.

Ну что ж, — скажет читатель, — инженеры и изобретатели, экспериментируя, находят для полупроводников новые области применения и создают новые приборы. Но в теории царит путаница. Никто не может уяснить ни что такое полупроводники, ни чем обусловлены их замечательные свойства. Ведущие физики теряют интерес к проблеме полупроводников, и основные понятия и идеи остаются в таком же запутанном и темном состоянии, как и во времена Фарадея.

Совсем нет! Именно в 30-е годы идеи квантовой механики позволяют достичь замечательных успехов в объяснении электрических, оптических и тепловых свойств твердых тел. В том числе и полупроводников. В 1931 г. английский физик А. Вильсон, основываясь на идеях квантовой механики и используя замечательные работы своих непосредственных предшественников А. Зоммерфельда и Ф. Блоха, построил теорию электропроводности твердых тел. Теория Вильсона

позволила ответить на многие принципиальные вопросы, связанные с характером электропроводности. Была установлена связь между строением электронных оболочек атомов и тем, окажется ли кристалл, составленный из этих атомов, металлом, диэлектриком или полупроводником.

Работами немецкого физика В. Шоттки, английского — Н. Мотта, советского ученого Бориса Иосифовича Давыдова была построена теория процессов, происходящих в области контакта полупроводника с металлом. Эта теория позволила объяснить выпрямляющие свойства полупроводниковых точечных детекторов.

В 1933 г., как уже упоминалось, Яков Ильич Френкель сумел объяснить, почему ток в полупроводниках может переноситься не только отрицательно заряженными электронами, но и положительными дырками. Теория Френкеля разъяснила многие, до того непонятные, электрические свойства полупроводников.

Так в чем же дело? Какие основания были у Гуддена (и многих других исследователей) для путанных, противоречивых и пессимистических выводов?

Неужели только то обстоятельство, что полупроводники очень чувствительны к примесям и их трудно очистить...?

В основном, да. Тот же Гудден, объясняя истоки своего пессимизма, писал, что по его мнению, проводимость полупроводников полностью определяется содержащимися в нем примесями. И если бы удалось какой-либо полупроводник от примесей очистить (в возможности этого, впрочем, Гудден сильно сомневался), то этот «полупроводниковый» кристалл просто превратился бы в изолятор.

Но дело, конечно же, было не только в трудностях очистки полупроводников и сложностях исследования их «собственных» свойств. Замечательный советский ученый Абрам Федорович Иоффе еще в 1916 г. показал, что, по крайней мере для некоторых соединений, можно найти технологические приемы, позволяющие так очистить кристалл от примесей, чтобы стало возможным изучать собственную проводимость кристаллов.

Вот выдержки из лекции А. Ф. Иоффе по физике кристаллов, опубликованной в 1929 г.:

«Четыре образца алюминисолевых квасцов ($K_2SO_4 \times Al_2(SO_4)_3 \times 24H_2O$) давали значения проводимости ..., отличавшиеся друг от друга в 90 раз. Возможную причину этих колебаний можно было подозревать в наличии примесей, хотя кристаллы квасцов были обозначены как «purissimum» (чистейшие).

Мы попытались очистить кристаллы посредством продолжительного (пропускания) электрического тока, ибо мы ожидали, что ток унесет к электродам ионы примесей. После этого части кристалла, прилегающие к электродам, были удалены, а средняя часть применена для измерений электропроводности. Нам в самом деле удалось значительно понизить проводимость.

Проводимость кристалла квасцов, равная первоначально $10,8 \cdot 10^{-12} \text{ Ом}^{-1}$, понижалась сначала до $0,96 \cdot 10^{-12} \text{ Ом}^{-1}$ и, наконец, до $0,22 \cdot 10^{-12} \text{ Ом}^{-1}$. Тем не менее, этим путем не удалось достигнуть постоянной стандартной величины (проводимости) для всех кристаллов одного сорта.

Мы поэтому попытались очистить кристаллы «purissimum» посредством процесса многократной перекристаллизации. Мы пользовались наиболее чистой водой и растворяли в ней кристаллы, исследуемые предварительно в концентрированном пучке света на предмет возможных включений и трещин. Поверхностный слой растворявшихся кристаллов предварительно соскабливался ножом. Непрерывный и быстрый поток пересыщенного раствора обмывал растущие кристаллы, укрепленные на стеклянных трубках. Полученные в большом количестве кристаллы изучались в луче света. Первоначальный кристалл, начинавший кристаллизацию, и все включения удалялись, а совершенно однородные образцы направлялись для новой кристаллизации..., наконец, был получен определенный результат».

Каким был этот результат, показывает табл. I. Проводимость исходных образцов отличалась в 90 раз. Из таблицы видно, как уменьшался разброс в проводимости по мере роста числа кристаллизаций.

Т а б л и ц а I

Номер образца	1	2	3	4	5	6	7	Разброс в проводимости, %
Проводимость после первой кристаллизации	38	120	26	100	138	30	72	530%
После второй кристаллизации	21,0	20,6	21,8	21,6	21,8	22,8	22,2	9,7%
После третьей кристаллизации	22,8	21,6	20,4	22,8	22,8	21,6	21,3	10,5%

Процитированный отрывок из лекции А. Ф. Иоффе во многих отношениях примечателен. Прежде всего, следует обратить внимание на великолепный результат, достигнутый с помощью примененного А. Ф. Иоффе технологического процесса. Разброс в проводимости исходных (чистейших, «purissimum») образцов, составлявший 9000%, в результате первой кристаллизации снизился до 530% (в 17 раз), после второй — еще в 55 раз и со-

ставил всего около 10%! (Видно, что третья перекристаллизация уже не привела к улучшению результатов — достигнута предельная для данного метода степень однородности кристаллов.)

Далее в этом коротком отрывке перечислены многочисленные технологические приемы, которые полностью сохранили свое значение и в наши дни. Пропускание тока через кристалл в течение длительного времени — метод электролитической очистки, которым А. Ф. Иоффе попытался вначале очистить кристаллы, не привел в данном случае к желаемому результату. Дело в том, что этим методом можно очистить кристалл только от ионов примесей, которые слабо закреплены в решетке кристалла и могут двигаться под действием электрического поля.

Обратившись к методу многократной перекристаллизации, исследователь использовал приемы, имеющие с точки зрения современных понятий ясный физический смысл. Просвечивание кристаллов сильным пучком света и удаление участков, содержащих неоднородности, позволяло исключить области с особенно высоким содержанием дефектов. Как теперь известно, дефекты не только сами способны создавать донорные и акцепторные уровни, но и служат центрами, вокруг которых концентрируются примеси и загрязнения. Соскабливание поверхности ножом способствовало удалению наиболее загрязненных участков кристалла. Наконец, сама многократная перекристаллизация позволила получать совершенную кристаллическую структуру со все меньшим количеством примесей.

И, как следствие этой тщательной и кропотливой работы, — результат, о котором почти через 30 лет (в 1957 г.) А. Ф. Иоффе с вполне законной гордостью упомянул в своей замечательной книге «Физика полупроводников». «Путем многократной кристаллизации мы получили десятки чистых образцов квасцов и селитры, обладающих строго одинаковым значением электропроводности и одинаковым ходом ее при изменении температуры».

Так в чем же было дело?

Теперь, через 50 лет, оглядываясь назад, причины установить не так уж сложно.

Но, может быть, мы лучше поймем истоки пессимизма Гуддена, если проследим за двумя историями. Двумя из многих сотен и тысяч, составивших историю полупроводников.

Первая история о «старейшем» полупроводнике, открытом в 1833 г. Фарадеем, — сернистом серебре (Ag_2S).

Через 18 лет после открытия Фарадеем необычных свойств сернистого серебра в печати появилась статья физика Хитторфа, в которой утверждалось, что ничего необычного в температурной зависимости проводимости Ag_2S нет. И ни о каком особом классе веществ, упомянутом Фарадеем, говорить, стало быть, нет оснований. Дело в том, писал Хитторф, что Ag_2S является просто-напросто твердым электролитом.

Свойства электролитов — жидкостей, проводящих электричество, начали изучаться еще задолго до Фарадея. Учитель Фарадея Дэви в 1800—1807 гг. исследовал целый ряд таких жидкостей. Сам Фарадей посвятил много лет исследованию электролитов и установил известную связь между количеством прошедшего через электролит электричества и массой выделившегося на электродах вещества (первый закон Фарадея для электролиза). Было хорошо известно, что проводимость электролитов растет с температурой. Тот же Дэви и знаменитый шведский химик Берцелиус (1779—1848) совершенно правильно определили причину этого явления. Ток в электролитах переносится положительно и отрицательно заряженными частичками (ионами), движущимися к противоположным электродам. Чем выше температура, чем легче разделяется электролит на положительные и отрицательные ионы, тем выше концентрация ионов, а потому и больше величина проводимости.

Более того, было известно, что электролиты не обязательно должны быть жидкими. Дэви впервые выделил элемент калий, пропуская электрический ток через твердый электролит — едкое кали (KOH).

К тому времени, когда Хитторф опубликовал свою статью, авторитет Фарадея был непререкаем. И для того чтобы обвинить аккуратнейшего и щепетильнейшего Майкла Фарадея, искусного экспериментатора и опытейшего исследователя, в том, что он не заметил выделившегося на электродах серебра, нужно было иметь очень серьезные основания. Хитторф их имел: в его опытах серебро на электродах выделялось.

Прошло еще несколько десятилетий, и в распоряжении исследователей появился новый и очень эффективный метод исследования электропроводности — эффект Холла. С помощью этого эффекта можно определять подвижность носителей заряда. Подвижность ионов, как правило, в тысячи раз меньше, чем подвижность электронов или дырок. Чем выше подвижность носителей, тем легче обнаружить эффект Холла (почему — мы узнаем в третьей части книги, где подробно познакомимся с этим эффектом). В первые десятилетия после открытия эффекта Холла чувствительность метода была такова, что движение ионов с помощью этого эффекта обнаруживать не удавалось. В электролитах, где ток переносится ионами,

Яков Ильич Френкель родился в 1894 г. в Ростове-на-Дону. Его способности к точным наукам проявились очень рано. В 1912 г., учась в последнем классе гимназии, Я. И. Френкель начал работу «Электродинамика Солнечной системы», которую показал профессору Петербургского политехнического института А. Ф. Иоффе. Иоффе сразу оценил талант молодого человека и вскоре привлек его к работе в своем семинаре по новой (квантовой) физике. В 1918 г., когда был организован Физико-технический институт, А. Ф. Иоффе пригласил Якова Ильича работать в институте.



Яков Ильич Френкель
(1894—1952)

С 1921 г. и до последних дней жизни Я. И. Френкель — сотрудник, а затем — заведующий теоретическим отделом ФТИ. Одновременно он заведовал кафедрой теоретической физики в Ленинградском политехническом институте.

Круг интересов Я. И. Френкеля был необычайно широк. Он внес большой вклад в развитие физики твердого тела, физики жидкостей, физики атомного ядра, астрофизики, био- и геофизики. Я. И. Френкель является основоположником физики спекания — научной основы порошковой металлургии; международная премия его имени присуждается за лучшие работы в этой области.

Замечательные работы были выполнены Я. И. Френкелем в физике полупроводников. Совместно со своим учителем, А. Ф. Иоффе, Яков Ильич разработал туннельную теорию выпрямления на контакте металл — полупроводник. Эта теория объясняет работу туннельных диодов. Одна из самых плодотворных идей, выдвинутых в области физики полупроводников, — идея «квазичастицы» — экситона была предложена и детально разработана Я. И. Френкелем. До сих пор каждый год публикуются десятки и сотни научных работ, развивающих идеи Я. И. Френкеля.

Яков Ильич был и художественно одаренной натурой, он профессионально рисовал и прекрасно играл на скрипке.

Президент Академии наук СССР, академик А. П. Александров, писал о Я. И. Френкеле:

«Человек необычайно живого ума, Яков Ильич интересовался всем новым во всех областях физики. Он, как никто другой, умел в простейшем виде, без всякого формализма немедленно обсудить физическую сущность нового явления, как только о нем становилось известно».

Труды Я. И. Френкеля переведены на многие языки.

Л и т е р а т у р а

Френкель В. Я. Яков Ильич Френкель. — Л. — М.: Наука, 1966.
Сборник «Воспоминания о Я. И. Френкеле». — Л.: Наука, 1976.

эффект Холла не наблюдался. А вот в веществах с электронной проводимостью наблюдать эффект Холла было возможно.

В 1902 г. Стрейнци, исследовав Ag_2S , пришел к заключению, что Фарадей был прав, а Хитторф ошибался — проводимость Ag_2S — электронная.

Однако в 1920 г. Тубанд с сотрудниками не только подтвердили, что при прохождении тока через Ag_2S на электроде выделяется серебро, но и проверили количественно выполнение первого закона Фарадея. Оказалось, что весь ток переносится ионами серебра: с погрешностью $\sim 1\%$ количество выделившегося на электродах серебра было пропорционально протекающему заряду и коэффициент пропорциональности в точности соответствовал закону Фарадея. Итак, стало как будто бы ясно, что Хитторф был прав, а Фарадей и Стрейнци ошибались.

Однако результаты Тубанда угрожали поколебать законы электрохимии. Для того чтобы проводимость Ag_2S можно было объяснить движением ионов серебра, следовало приписать этим ионам подвижность, в 1000 раз превосходящую ту величину, которая следовала из самых оптимистических теоретических оценок.

Проблема была решена к ее столетнему юбилею. В 1933 г. К. Вагнер обнаружил источник ошибки в опытах Тубанда. Выделение серы и серебра на электродах происходило не вследствие электролиза, а вследствие побочной электрохимической реакции у электродов. Можно создать условия, при которых эта реакция не идет. В этом случае выделение серебра на электродах прекращается — однако это никак не сказывается на проводимости сернистого серебра. Итак, правота Майкла Фарадея была установлена. На этот раз окончательно.

На выяснение этой маленькой проблемы ушло, как мы видели, целое столетие.

В следующей истории речь пойдет о проводимости полупроводников и ее зависимости от температуры.

Одним из первых полупроводниковых материалов, подвергнутых тщательному и разностороннему изучению, была закись меди (Cu_2O). Интерес к этому материалу вполне понятен. Как уже упоминалось, на основе Cu_2O были созданы получившие широкое распространение выпрямители и фотоэлементы.

На рис. 14 показаны температурные зависимости проводимости закиси меди, полученные тремя различными группами исследователей. (Обратите внимание на масштаб по оси ординат. Сдвиг на одно деление соответствует изменению проводимости в 10 раз.)

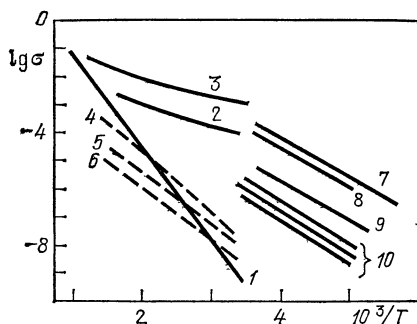


Рис. 14. Зависимость проводимости закиси меди (Cu_2O) от температуры (σ , $\text{Ом}^{-1}\cdot\text{см}^{-1}$).

Кривые 1—3 — данные В. П. Жузе и И. В. Курчатова; 4—6 — по данным Леблана и Закса; 7—10 — по данным Э. Энгельгарда.

Советские физики В. П. Жузе и И. В. Курчатов измерили проводимость соединения Cu_2O , очищенного до такой степени, что избыточный кислород, обычно присутствующий в нем, не обнаруживался. Для таких особо чистых образцов зависимость проводимости от температуры определялась кривой 1. Если в образцах содержалась примесь кислорода в концентрации 0,061% или 0,1 %, то зависимость $\sigma(T)$ определялась кривыми 2 и 3 соответственно.

Но вот Леблан и Закс получили для очищенного в максимально доступной им степени Cu_2O совершенно иные результаты (кривые 4, 5, 6). При температуре 700 К ($1000/T = 1,4$), как видно из рисунка, проводимость различных образцов отличалась по данным Леблана и Закса в 20 раз и была значительно меньше, чем значение, установленное советскими физиками. Из рисунка видно, что для образца Леблана и Закса с наименьшей проводимостью (кривая 6) величина σ в 200 раз меньше, чем значение, измеренное Жузе и Курчатовым! Что и говорить, разница заметная.

Зависимости 7—10 получены еще одним ученым, исследовавшим проводимость Si_2O , — Э. Энгельгардом. Стремясь «заморозить» связанный в Si_2O избыточный кислород, Энгельгард провел измерения при относительно низких температурах, надеясь в этих условиях получить надежные воспроизводимые результаты. Как видно из рисунка, он получил значения проводимости, отличающиеся для различных образцов более, чем в 1000 раз!

Рис. 14 взят из того же обзора Б. Гуддена, цитатами из которого открывается глава, но представлен в сильно упрощенном виде. В оригинале на рисунке показаны результаты не трех, а семи групп авторов. И, если не читать подробных комментариев Гуддена, рисунок выглядит как прямоугольник, заполненный линиями, идущими под самыми разными углами и на самых разных уровнях. Не правда ли ясно, что «... в этом случае имеют место весьма запутанные соотношения...»?

Для того чтобы при исследовании полупроводников получить надежные, воспроизводимые результаты, необходимо было научиться очищать материал с такой тщательностью, снижать концентрацию примесей до такого уровня, какой в те годы трудно было просто даже представить. Никогда раньше самые основные, самые фундаментальные свойства изучаемого материала не зависели в такой степени от ничтожных, невообразимо малых добавок постоянных веществ. Это обстоятельство надо было, прежде всего, ясно осознать. И понять, что никакого иного пути для изучения полупроводников, кроме разработки новых совершенных методов очистки и приготовления материалов, нет. Нужны были многолетние целенаправленные усилия физиков и технологов, чтобы разработать такие методы, физиков и химиков — чтобы разработать методики, с помощью которых можно было бы обнаруживать и анализировать ничтожные («следовые») количества примесей в материале, физиков и инженеров — чтобы разработать и создать необходимую измерительную аппаратуру.

А между тем, в 30-е годы нашего века общественная потребность в знаниях, о которой писал Максвелл, в части, касающейся полупроводников, проявлялась весьма скромно. Детище замечательных

американских изобретателей Томаса Эдисона и Ли де Фореста *) — электронная лампа стала материальной основой (или, как принято говорить, элементной базой) первой «электронной революции».

Да, именно в 30-е годы принцип работы полупроводникового точечного детектора получил, наконец, физическое объяснение. Но по иронии судьбы, именно в это время сами точечные детекторы перестали интересовать кого бы то ни было.

Электронные лампы обеспечивали качество и дальность радиоприема, не сравнимые со скромными возможностями лосевского детектора. Усилители на электронных лампах — надежные, долговечные, стабильные — опускались на дно океанов вместе с трансатлантическими телефонными кабелями и взлетали в небо с самолетными радиостанциями. Электронные лампы служили элементами генераторов, преобразователей, детекторов, усилителей. В 1927 г. американская фирма Bell Telephone продемонстрировала первую промышленную телевизионную установку, собранную на электронных лампах. Казалось, что электронные лампы на многие годы обеспечат прогресс только что зародившейся, но уже совершенно необходимой физикам и связистам, медикам и военным, навигаторам и криминалистам, людям десятков специальностей науки — радиоэлектронике. Скромные полупроводниковые приборы явно не годились в лидеры этой первой электронной революции.

Но... и это неизменно прослеживается в истории любой науки: как бы ни сложились внешние обстоятельства, каким бы немодным ни выглядело в определенное время то или иное научное направление, каким бы сложным и не сулящим быстрые выгоды и успехи оно ни казалось, — всегда находятся люди, продолжающие исследования в избранной ими области.

Так было и с полупроводниками. Практически в каждой технически развитой стране нашлись энтузиасты, которые терпеливо продолжали исследовать капризные, невоспроизводимые, неподдающиеся материалы и искать для них новые применения. В нашей стране работы по исследованию полупроводников

*) Ли де Форест (1873—1961) был не только замечательным изобретателем, но и талантливым физиком, учеником одного из самых крупных физиков прошлого столетия Уилларда Гиббса.

возглавил замечательный физик и блестящий организатор Абрам Федорович Иоффе.

Именно работы этих энтузиастов составили основу того поразительного взлета, который начался для полупроводников всего через десять лет после пессимистических выводов Гуддена и продолжается до настоящего времени.

30-е годы закончились катастрофой, коснувшейся всего человечества. Вторая мировая война не только привела к бесчисленным жертвам и невиданным разрушениям. Она на многое изменила взгляды. В том числе — и на роль и место науки в жизни общества. Уроки войны показали, как быстро самые абстрактные научные исследования могут привести к созданию нового оружия, к революционным изменениям в технологии, к коренным преобразованиям в промышленности. Стало ясно, что хотя крупные научные идеи и открытия по-прежнему связаны с деятельностью отдельных выдающихся ученых, для быстрого и эффективного воплощения этих идей в жизнь нужна координированная работа больших коллективов исследователей, конструкторов, технологов.

Между тем, потребность в знаниях о полупроводниках с течением времени становилась все более настоятельной. Уже в конце 30-х годов в связи с появлением радиолокации пришлось вспомнить о полупроводниковых детекторах. Оказалось, что германиевые и кремниевые точечные детекторы способны преобразовывать сигналы сверхвысоких частот, с которыми лампы того времени справиться не могли. В технике нашли важные применения полупроводниковые фотоэлементы, термобатареи, селеновые и меднозакисные выпрямители.

Но главное было даже не в этом. «Королева радиоэлектроники» — электронная лампа оказывалась явно не в состоянии справиться с усложняющимися задачами. Устройствам на электронных лампах нужны многие секунды, а иногда и минуты для того, чтобы прийти в рабочее состояние: лампы должны прогреться. Между тем, эти минуты могут решить судьбу самолета, с которым нельзя установить связь, или военного корабля, у которого не работает радиолокатор, автоматы управления стрельбой или усилитель ультразвукового локатора, предупреждающего о появлении подводной лодки.

Абрам Федорович Иоффе родился 29 октября 1880 г. в городе Ромны Полтавской губернии. Окончив Роменское реальное училище, А. Ф. Иоффе поступает в Петербургский технологический институт. Обучаясь в нем, Иоффе одновременно посещает курсы, основанные знаменитым физиологом Петром Францевичем Лесгафтом, занимается микробиологией, принимает деятельное и активное участие в революционном студенческом движении. Зарекомендовав себя во время студенческой практики знающим толковым и талантливым инженером, А. Ф. Иоффе по окончании института получает лестное и выгодное в материальном отношении предложение заняться инженерной деятельностью. А. Ф. Иоффе, однако, отказывается от него. Желая специализироваться в экспериментальной физике, Иоффе в 1902 г. уезжает в Мюнхен, в лабораторию Вильгельма Конрада Рентгена.



Абрам Федорович
Иоффе
(1880—1960)

В. Рентген, открывший названное его именем излучение, был одним из лучших экспериментаторов своего времени. В возглавляемой им лаборатории молодой ученый из России прошел замечательную школу экспериментального мастерства, преданности науке и строжайшей научной щепетильности.

В Мюнхене А. Ф. Иоффе провел первые экспериментальные исследования и в 1905 г. защитил докторскую диссертацию, удостоенную редкостного отличия — «с наивысшей похвалой». Ему было предложено место профессора, однако А. Ф. Иоффе вернулся на родину.

Степень доктора, полученная за границей, во внимание в России не принималась, и ученый должен был вновь защищать магистерскую (в 1913 г.) и докторскую (в 1915 г.) диссертации. Темой магистерской диссертации была экспериментальная работа по определению заряда электрона и работа, в которой Иоффе впервые показал, что вокруг потока электронов в вакууме (катодных лучей) существует магнитное поле. Темой докторской диссертации была начатая еще в Мюнхене работа по исследованию упругих свойств кварца. Замечательный математик В. А. Стеклов, выступая на защите, сказал, обращаясь к диссертанту: «Произведенная Вами при помощи самых простых средств экспериментальная работа может быть уподоблена по проявленной Вами систематической и неуклонной настойчивости работам Фарадея. Вместе с тем... в математической обработке наблюдений Ваша работа столь же безукоризненна, как работы английских физиков Максвелла, Томсона, Рэлея, Стокса и других, и я отдаю лишь должное, признавая Вашу диссертацию превосходной во всех отношениях».

С 1916 г. усилиями А. Ф. Иоффе в Петрограде организуется знаменитый впоследствии семинар по современной физике, который посещают будущие лауреаты Нобелевской премии П. Л. Капица, Н. Н. Семенов, а также П. И. Лукорский, Я. И. Френкель, Я. Г. Дорфман и многие другие.

После Октябрьской революции по инициативе А. Ф. Иоффе в Петрограде организуется Физико-технический институт, знаменитый «Физтех», который воспитал в своих стенах десятки замечательных ученых. Впоследствии усилиями ученых Физтеха во главе с Иоффе были организованы физико-технические институты в Харькове, Днепропетровске, Томске, Свердловске.

В 1920 г. Абрам Федорович Иоффе — ученый с мировой репутацией избирается академиком Академии наук.

А. Ф. Иоффе одним из первых в нашей стране предвидел замечательные перспективы полупроводниковой техники и технологий. С 1928 г. под его руководством в Физико-техническом институте ведутся работы по исследованию полупроводников, разработке термо- и фотоэлектрических приборов, полупроводниковых холодильников. В 1933 г. А. Ф. Иоффе издает первую на русском языке монографию, посвященную полупроводникам.

Абрам Федорович Иоффе был в дружеских отношениях со многими замечательными учеными самых разных стран: Рентгеном, Эренфестом, Эйнштейном, Бором, Дебаем, Лоренцем.

О встречах с ними он написал очень интересную книгу «Встречи с физиками».

А. Ф. Иоффе в течение ряда лет был вице-президентом АН СССР. Он был удостоен Ленинской и Государственной премий СССР, звания Героя Социалистического Труда и трех орденов Ленина, избран почетным членом многих академий и научных обществ мира.

Умер А. Ф. Иоффе 14 октября 1960 г., за две недели до своего 80-летия.

Л и т е р а т у р а

Иоффе А. Ф. Встречи с физиками. — М.: Физматгиз, 1962.

Сборник «Воспоминания об А. Ф. Иоффе». — Л.: Наука, 1973.

Срок службы электронной лампы принципиально ограничен. Раскаленный катод лампы, испускающий электроны, постепенно портится — и лампа выходит из строя. Конечно, пока электронная схема включает только несколько ламп, а срок службы каждой лампы достаточно велик, особых проблем не возникает. Но последствия отказа лампы в устройстве, содержащем тысячи и десятки тысяч элементов, или в условиях, когда заменить ее очень трудно (например, в усилителе межконтинентального телефонного кабеля), могут быть очень серьезными.

Необходимость создания нового элемента — с принципиально неограниченным сроком службы, не требующим прогрева, не боящимся механических воздействий, экономичного, легкого и дешевого, ощущалась все острее.

Наиболее дальновидные специалисты отдавали себе отчет в том, что таким элементом может скорее всего оказаться полупроводниковый триод — транзистор.

Собственно, попытки создать такой полупроводниковый прибор, аналогичный по функциям лампы-триода, предпринимались с 1925 г. Были испробованы самые разные полупроводниковые материалы, самые различные конфигурации электродов и способы введения «сетки» между ними.

В разных странах были выданы патенты на гипотетические твердотельные усилители, но... ни одной работающей конструкции не удалось создать в течение более, чем 20 лет!

Успех был достигнут организацией, которая не только привлекла к работе над созданием транзистора первоклассных физиков, будущих лауреатов Нобелевской премии, но и сосредоточила на этой проблеме скоординированные целенаправленные усилия многих десятков специалистов: металлургов, химиков, технологов, инженеров.

В 1948 г. фирма Bell Telephone впервые публично продемонстрировала твердотельный усилитель — точечный транзистор, созданный годом раньше сотрудниками фирмы Бардином, Браттейном и Шокли (рис. 15).

На рис. 16 показан точечный транзистор в том виде, в каком он был впервые предложен вниманию специалистов в 1948 г.

Эмиттерный и коллекторный $p-n$ -переходы создавались вплавлением в германиевый кристалл показанных на рисунке тонких проволочек. Область между проволочками служила базой транзистора. Обратите внимание на трогательную приверженность



Рис. 15. Создатели транзистора, будущие лауреаты Нобелевской премии по физике У. Шокли (слева), Дж. Бардин (в центре) и У. Браттейн (справа).

Дж. Бардин вместе с У. Шокли и У. Браттейном в 1956 г. получили Нобелевскую премию по физике за исследования полупроводников и открытие полупроводникового триода — транзистора.

человеческого воображения к привычным формам: первый транзистор повторяет по внешнему виду электронную лампу. Так, конструкции автомобилей в течение первых 15—20 лет повторяли по форме кареты и дилижансы. Конечно, конструкторы транзисторов расстались с традиционными формами гораздо быстрее. Уже через 2 года транзисторы начали выпускаться в специальных корпусах, ничем не напоминающих вакуумные лампы.

Появление транзистора «замкнуло» удивительно эффективно и плодотворно заработавшую «цепочку». Во всех без исключения промышленно развитых странах к работе над усовершенствованием транзистора приступили сотни и тысячи специалистов са-

мых различных профессий. Metallурги, химики и физико-химики углубленно исследовали возможности создания новых материалов. Технологи разрабатывали новые методы очистки материалов и создания совершенных монокристаллов. Физики, получив в свое

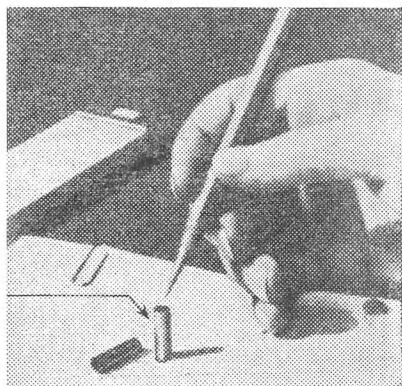
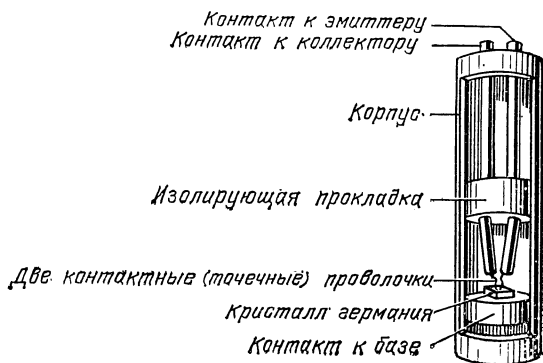


Рис. 16. Первый германиевый транзистор (на рисунке внизу на него указывает острое карандаша).

Для сравнения на рисунке показана самая маленькая из выпускавшихся в 1948 г. вакуумных электронных ламп (лежит на листе бумаги рядом с транзистором).

распоряжение чистые материалы с воспроизводимыми параметрами, не только подсказывали пути к улучшению транзисторов, но и обнаруживали новые физические эффекты. Некоторые из этих эффектов могли быть положены в основу новых приборов. Инженеры, физики и конструкторы, разрабатывавшие эти

новые приборы, предъявляли новые требования к материалам и степени их очистки. Металлурги, химики и технологи...

Волшебная карусель, запущенная изобретением транзистора, раскручивается все быстрее и быстрее и в наши дни вращается просто с фантастической скоростью.

Исследованы сотни полупроводниковых материалов и каждый год синтезируются все новые и новые полупроводники. Уровень чистоты полупроводниковых материалов и методы контроля доведены до удивительного совершенства. В некоторых случаях удается обнаруживать один атом примеси на десять тысяч миллиардов атомов полупроводника. Разработаны методы получения монокристаллов чудовищной величины — технологи умеют выращивать монокристаллы кремния объемом в 6000 см³. Придуманы способы, позволяющие под контролем ЭВМ наращивать буквально атом за атомом, один атомный слой на другой. Придуманы тысячи различных полупроводниковых приборов, основывающихся на сотнях открытых физиками эффектов. Ежегодно в мире выпускаются миллиарды полупроводниковых приборов и миллионы интегральных схем, каждая из которых содержит десятки и сотни тысяч отдельных диодов и транзисторов.

Полупроводники и приборы на их основе сделались неотъемлемой частью современной цивилизации, во многом определяющей ее развитие.

Оглянемся и еще раз вспомним пройденный путь. 60-е годы XVIII века: трудами Джозефа Пристли, Джамбатисты Беккариа, Джона Кантона и, наконец, Генри Кавендиша установлено, что существует группа веществ, которые проводят электрический ток гораздо лучше, чем диэлектрики, но хуже, чем металлы.

1833 г.: Фарадей обнаруживает необычное поведение сопротивления некоторых веществ при изменении температуры и высказывает гипотезу о том, что это свойство служит признаком существования особого класса веществ.

Первое десятилетие XX века: обобщив результаты сотен работ, проделанных десятками ученых разных стран, Кенигсбергер вводит понятие «класса полупроводников» и указывает его основные отличительные признаки.

20—30-е годы XX века: трудами А. Вильсона, В. Шоттки, Н. Мотта, А. Ф. Иоффе, Б. И. Давыдова, Я. И. Френкеля и многих других закладываются основы физики полупроводников.

1947 г.: Бардин, Браттейн и Шокли создают первый полупроводниковый триод — транзистор. Это открытие, отвечавшее самым насущным, самым неотложным практическим запросам, сосредоточило на исследовании полупроводниковых материалов и приборов усилия многих тысяч ученых, инженеров и технологов, объединенных в большие коллективы, использующих для исследований и разработок самую современную аппаратуру и самые совершенные методы исследования. С этого времени прогресс в изучении полупроводников в значительной мере определяет успехи важнейших отраслей промышленности, связи, военного дела, бытовой техники, исследований космоса...

Per aspera ad astra!

СЕКРЕТ ЗОЛОТОЙ СЕРЕДИНЫ

Середина есть точка, ближайшая к мудрости; не дойти до нее — то же самое, что ее перейти.

Конфуций

Познакомившись с самыми основными свойствами полупроводников и с историей их открытия и изучения, мы теперь понимаем, что в понятие «полупроводники» вкладывается смысл гораздо более глубокий, чем просто понятие о веществах, которые проводят электрический ток хуже, чем металлы, но лучше, чем диэлектрики.

В этой части книги, углубляя знакомство с главными героями — электроном и дыркой, мы продолжим изучение физических свойств полупроводников. Узнаем, как возникает «умеренная» — промежуточная между металлами и диэлектриками — электропроводность полупроводников; почему именно полупроводники обладают удивительной способностью менять свои свойства под действием света, тепла и ничтожных добавок других веществ.

Глава 3

ПРИНЦЫ НА ГОРОШИНЕ

Тут все поняли, что это настоящая принцесса, ведь она лежала на двенадцати тюфяках и двенадцати перинках, а все-таки чувствовала горошину. Столь чувствительной могла быть только настоящая принцесса.

Ханс Кристиан Андерсен

Знакомясь во Введении с электрическими свойствами полупроводников, мы убедились в том, что при температуре абсолютного нуля они представляют собой идеальные диэлектрики. Свободных носи-

телей тока — электронов и дырок в полупроводниковых кристаллах нет и их электропроводность равна нулю.

Что произойдет, если кристалл нагреть?

Если не задумываться о количественных соотношениях, а размышлять только о качественной картине, то можно, казалось бы, рассуждать следующим образом. Нам известно, что атомы полупроводника связаны со своими соседями электронными связями. Чтобы разорвать эти связи, нужна энергия $\mathcal{E}_g$. Когда мы увеличим температуру кристалла до такой величины, что средняя энергия теплового движения атомов станет близкой к величине $\mathcal{E}_g$, можно ожидать, что хаотическое тепловое движение частиц начнет разрушать электронные связи. В кристалле появятся свободные носители — электроны и дырки — и он приобретет способность проводить ток.

Это очень правдоподобное и убедительное рассуждение. Проверим его теперь «с цифрами в руках». Для упоминавшихся уже раньше полупроводников — антимонида индия (InSb) и кремния (Si) величины $\mathcal{E}_g$ равны соответственно 0,17 и 1,1 эВ. Для германия (Ge) $\mathcal{E}_g = 0,72$ эВ, для фосфида галлия (GaP) — 2,3 эВ. Средняя энергия теплового движения частиц равняется, как известно, $\frac{3}{2}kT$, где k — постоянная Больцмана, T — абсолютная температура. Величина $k = 1,38 \cdot 10^{-23}$ Дж/К или $8,6 \times 10^{-5}$ эВ/К. Итак, если наше рассуждение справедливо, полупроводники начнут приобретать заметную проводимость при температуре T_0 , которая определится из приближенного равенства $\frac{3}{2}kT_0 \approx \mathcal{E}_g$. Таким образом, для InSb величина $T_0 \approx 2000$ К, для Ge — 8400 К, для Si — 12700 К и для GaP — 27000 К.

Задолго до того, как такая температура будет достигнута, каждое из этих веществ обратится в пар! М-да, уж не вкралась ли какая-нибудь ошибка в наши убедительные рассуждения?

Что произойдет, если блюдце с водой поставить на шкаф?

Проверим наши рассуждения еще раз на каком-нибудь более привычном объекте. Нальем, например, в блюдце немножко воды и поставим его

на шкаф, так, чтобы если вода исчезнет, мы не могли ни в чем заподозрить кошку. Займемся на несколько часов другими делами. Затем освидетельствуем блюдо — вода испарилась. Как люди, имеющие некоторый житейский опыт, мы полагаем, что удивляться тут нечему. Как ученые, мы должны испытать глубокое изумление.

В самом деле. При комнатной температуре ($T = 300\text{ K}$) энергия теплового движения молекул воды $\frac{3}{2}kT \approx 0,04\text{ эВ}$. Удельная теплота парообразования воды, как нетрудно убедиться, заглянув в учебник физики или в справочник, составляет при комнатной температуре $2,4 \cdot 10^6\text{ Дж/кг}$. Зная, сколько в 1 кг воды содержится молекул и сколькоим электронвольтам равняется один джоуль, легко подсчитать*), что на одну молекулу воды энергия парообразования составляет $\mathcal{E} \approx 0,46\text{ эВ}$. Таким образом, чтобы оторвать от поверхности воды в блюде одну молекулу, необходима энергия $\mathcal{E}$, более чем в 10 раз превышающая среднюю тепловую энергию. И тем не менее, молекулы воды покидают блюдо!

Дело в том, что, наряду с молекулами, энергия теплового движения которых порядка kT (их подавляющее большинство), в воде в каждый данный момент времени имеются и молекулы с энергией гораздо меньшей, чем средняя величина, и очень «энергичные» молекулы, энергия которых гораздо больше. Ведь каждая молекула приобретает энергию за счет случайных столкновений с другими молекулами или со стенками сосуда. Молекула в секунду испытывает множество случайных столкновений. Немудрено, что иногда, хотя и очень редко, она приобретает энергию, гораздо большую, чем средняя. Если эта энергия превысит 0,46 эВ и если в этот момент молекула подойдет под подходящим углом к поверхности воды, то она покинет блюдо. Если блюдо не накрыто — молекула улетает безвозвратно. Вода испаряется.

Если блюдо накрыть, между поверхностью воды и крышкой возникнет определенная концентрация улетевших из блюда молекул воды — водяной пар. При плотной крышке молекулам некуда деваться, и, проблуждав некоторое время между крышкой и поверхностью, они вновь вернуться в блюдо. Сколько

*) Проделайте этот расчет.

молекул каждую секунду уйдет с поверхности воды — столько же возвратится обратно. Установится динамическое равновесие.

Вот что произойдет, если кристалл нагреть!

Теперь ясно, в чем состояла ошибка в наших рассуждениях: мы полагали, что средняя тепловая энергия порядка kT присуща всем без исключения частицам. В действительности, в каждый данный момент найдутся частицы с энергией гораздо большей, чем средняя. В том числе и такие, чья энергия больше, чем величина $\mathcal{E}_g$. Часть электронных связей между атомами окажется разорванной: в полупроводнике возникнут электроны и дырки. Уйти из кристалла электроны и дырки не могут. Поэтому, проблуждав некоторое время, электрон и дырка встретятся и рекомбинируют (исчезнут). Установится динамическое равновесие: сколько электронов и дырок будет каждую секунду создаваться в объеме полупроводника, столько же их будет гибнуть в процессе рекомбинации. В полупроводнике при данной температуре установится определенная равновесная концентрация электронов и дырок.

В кристаллах, свойства которых мы сейчас изучаем, электроны и дырки всегда рождаются и гибнут парами, поэтому концентрация электронов n_i в точности равна концентрации дырок p_i . Определим, как зависит равновесная концентрация свободных носителей $n_i = p_i$ от температуры.

Для этого следует определить, сколько свободных носителей создается в единице объема полупроводника в единицу времени и сколько их гибнет за счет рекомбинации. В условиях равновесия эти две величины должны быть равны. Приравнивая их друг к другу, мы и определим равновесную концентрацию носителей.

Тепловая генерация электронов и дырок

Мы уже выяснили, в чем состоит главная особенность процесса создания, *генерации*, электронно-дырочных пар: для их образования за счет хаотического теплового движения нужна энергия $\mathcal{E}_g$,

обычно значительно превышающая среднюю энергию kT . Ясно, что чем ниже температура T и чем больше требуемая энергия $\mathcal{E}_g$, тем менее вероятно такое событие.

Сходные ситуации возникают при исследовании самых разнообразных физических процессов. Пусть для того, чтобы произошло какое-нибудь событие, например, генерация электронно-дырочной пары, испарение молекулы воды или ионизация молекулы газа за счет тепла, необходима относительно большая энергия $\mathcal{E} \gg kT$, которая может быть получена только за счет хаотического теплового движения. В одном из основных разделов физики — статистической физике — доказывается, что вероятность такого события всегда пропорциональна $e^{-\mathcal{E}/kT}$. Число $e \approx 2,718 \dots$ называется основанием натуральных логарифмов. Оно играет важную роль во многих разделах физики и математики *).

Зависимость вида $y = e^x$ называется *экспоненциальной* и встречается в физике чрезвычайно часто. Величина x называется *показателем экспоненты*. В нашем случае $x = -\mathcal{E}/kT$. Давайте посмотрим, как меняется величина экспоненты y при изменении ее показателя x .

Пусть энергия $\mathcal{E}$, необходимая для того, чтобы произошло интересующее нас событие, равняется $2 kT$. Тогда $x = -2$, а величина экспоненты в этом случае равна $y = 0,135$. Допустим, что для осуществления другого события необходима в 5 раз большая энергия $\mathcal{E} = 10 kT$. Тогда $y \approx 4,6 \cdot 10^{-5}$. Обратите внимание: величина требуемой энергии возросла всего в 5 раз, а вероятность события, пропорциональная величине экспоненты y , уменьшилась в 3 000 раз.

Пусть необходимая энергия еще в 5 раз больше: $\mathcal{E} = 50 kT$. Тогда $y \approx 2 \cdot 10^{-22}$. Увеличение требуемой энергии еще в 5 раз уменьшило вероятность события в миллиард миллиардов раз!

Когда абсолютная величина показателя экспоненты много больше единицы, даже небольшие изменения показателя сильно влияют на величину экспоненты.

Итак, для образования электронно-дырочной пары необходима энергия $\mathcal{E}_g$. Таким образом, число элек-

*) О числе e можно прочитать в статье: Кузьмин Е., Ширшов А. — Квант, 1979, № 8.

тронно-дырочных пар K_1 , создаваемых каждую секунду в единице объема полупроводника,

$$K_1 = \alpha e^{-\epsilon_g/kT}. \quad (2)$$

В формуле (2) α — коэффициент пропорциональности, который можно рассчитать только с помощью квантовой механики *).

С другой стороны, каждую секунду в той же единице объема исчезнет за счет рекомбинации некоторое число носителей K_2 . От чего зависит число рекомбинирующих носителей?

Рекомбинация электронов и дырок

Для того чтобы рекомбинировать, электрон и дырка должны встретиться. Обсудим, от чего зависит частота таких встреч. Проведем следующий мысленный эксперимент. Будем следить за каким-нибудь атомом кристаллической решетки и отмечать появление в его окрестности, скажем, дырки. Ясно, что чем больше дырок имеется в единице объема полупроводника, тем чаще они будут появляться: вероятность появления дырки пропорциональна концентрации дырок p . По тем же причинам вероятность появления свободного электрона в окрестностях этого же атома пропорциональна концентрации электронов n . Нас интересует вероятность *одновременного* появления в окрестностях атома и электрона, и дырки. Очевидно, что появление электрона никак не зависит от того, имеется ли в окрестности данного атома дырка. И наоборот: вероятность того, что в данный момент в окрестностях атома имеется свободный электрон, никак не влияет на вероятность появления дырки. Появление в окрестностях атома электрона и дырки — *независимые события*.

Вероятность того, что произойдут два независимых события, пропорциональна произведению

*) Когда-то великий немецкий поэт Генрих Гейне писал, что латинские неправильные глаголы отличаются от латинских правильных глаголов тем, что за первые чаще секут. Есть студенты, которые уверены, что в этом же состоит и основная разница между квантовой и классической механикой. В известном смысле это правильно. Однако имеются и другие существенные различия. Некоторые из них мы в дальнейшем обсудим применительно к свойствам полупроводников.

вероятностей каждого из событий. Почему это так, подробно рассказывается в книге А. М. Колмогорова, И. Г. Журбенко и А. В. Прохорова «Введение в теорию вероятностей» (М.: Наука, 1982. Библиотечка «Квант», вып. 23).

Итак, вероятность одновременного появления в окрестностях атома электрона и дырки, т. е. вероятность их встречи и рекомбинации, пропорциональна произведению концентраций электронов и дырок. Таким образом, число носителей, рекомбинирующих каждую секунду в единице объема:

$$K_2 = \beta n_i p_i. \quad (3)$$

Коэффициент пропорциональности β также может быть вычислен только с помощью квантовой механики. Поскольку $n_i = p_i$, то

$$K_2 = \beta n_i^2 = \beta p_i^2. \quad (4)$$

Собственные полупроводники

Приравнивая число ежесекундно появляющихся пар K_1 числу гибнущих K_2 , получим:

$$\alpha e^{-\mathcal{E}_g/kT} = \beta n_i^2 = \beta p_i^2. \quad (5)$$

Отсюда получаем, что

$$n_i = p_i = \sqrt{\alpha/\beta} e^{-\mathcal{E}_g/2kT} = A e^{-\mathcal{E}_g/2kT}. \quad (6)$$

О формуле (6) стоит поговорить подробно.

Прежде всего следует сказать, что полупроводник, в котором концентрация носителей определяется формулой (6), называется *собственным*. Индекс «i» у n_i и p_i от английского слова intrinsic — внутренне присущий, собственный. Вычисленная по формуле (6) *собственная концентрация* электронов и дырок $n_i = p_i$ определяется температурой T и энергией $\mathcal{E}_g$ — важнейшей характеристикой полупроводника.

Что касается неизвестного коэффициента A , то для различных полупроводников при комнатной температуре его величина лежит в пределах от $\sim 10^{17}$ до $\sim 10^{19}$ см⁻³. С увеличением температуры коэффициент A , вообще говоря, несколько растет, но в интересующем нас диапазоне температур мы можем смело считать A величиной постоянной. Температурная за-

висимость собственной концентрации практически полностью определяется вторым сомножителем в формуле (6) — экспонентой. Чтобы лучше понять, как резко зависит от температуры собственная концентрация, проследим, как меняется с температурой величина n_i в кремнии. Для кремния $\mathcal{E}_g = 1,1$ эВ, $A \approx 2 \cdot 10^{19} \text{ см}^{-3}$.

При температуре жидкого азота ($T = 77$ К) величина n_i , подсчитанная по формуле (6), равняется $1,7 \cdot 10^{-17} \text{ см}^{-3}$. (Температуру жидкого азота физики очень любят. При этой температуре тепловое движение атомов заметно стихает и появляется возможность изучения таких тонких деталей взаимодействия атомов, которые при комнатной температуре маскируются энергичными тепловыми колебаниями. Жидкий азот дешев и имеется, практически, в любой лаборатории.) Итак, при $T = 77$ К нужен куб кремния с длиной ребра в 4 км, чтобы в нем нашлась одна-единственная электронно-дырочная пара.

При температуре 300 К величина $n_i = 1,1 \times 10^{10} \text{ см}^{-3}$ (температуру 300 К, или 27°C , любят почти все. При этой температуре приятно купаться и загорать.). Заметим, что увеличение температуры всего в 4 раза привело к изменению собственной концентрации электронов и дырок на 27 порядков!

При $T = 500$ К $n_i = 5,6 \cdot 10^{13} \text{ см}^{-3}$. Много это или мало? В 1 см^3 кремния содержится приблизительно 10^{23} атомов. Таким образом, даже при 500 К только одна десятиллиардная часть связей между атомами окажется разорванной. В то же время стоит вспомнить, что в металлах при любых температурах плотность электронов составляет $\sim 10^{22} \text{ см}^{-3}$.

Так же резко, как от температуры, собственная концентрация n_i зависит и от энергии $\mathcal{E}_g$. Рассмотрим табл. II. В ней приведены значения n_i для различных полупроводников при комнатной температуре (300 К).

Прежде всего зададим простой вопрос: легко ли изучать электрические свойства полупроводников? В 1 см^3 любого кристалла содержится приблизительно $4 \cdot 10^{23}$ атомов. Чистым принято называть вещество, в котором содержится один посторонний атом на 1000 собственных (концентрация примесей $\sim 0,1 \%$). С химической точки зрения изумительно чистым будет вещество, в котором один посторонний атом приходится

Таблица II

Полупроводники	InSb	Ge	Si
$\mathcal{E}_g$, эВ n_i , см ⁻³	0,17 $1,3 \cdot 10^{16}$	0,72 $2,4 \cdot 10^{13}$	1,1 $1,1 \cdot 10^{10}$

Полупроводники	InP	GaAs	GaP	SiC
$\mathcal{E}_g$, эВ n_i , см ⁻³	1,3 $5,7 \cdot 10^7$	1,4 $1,4 \cdot 10^7$	2,3 0,8	3,2 $2 \cdot 10^{-8}$

на 100 000 собственных. При этом в каждом кубическом сантиметре вещества будет содержаться $\sim 10^{18}$ посторонних атомов. А теперь представим себе, что примесь, попавшая в полупроводник, способна сравнительно легко отдавать свободный электрон или образовывать дырку. (Чуть дальше мы убедимся, что очень многие примеси обладают этим свойством.) Тогда с химической точки зрения изумительно чистый германий (концентрация примесей всего 0,001 %) будет при комнатной температуре содержать созданных примесью электронов в 40 000 раз больше, чем собственных. Кремний — в 10^8 раз больше. Фосфид галлия — в 10^{18} раз больше!

А ведь примеси содержатся и в исходных материалах, из которых получают полупроводниковые кристаллы, и в стенках печей и приборов, где происходит синтез и очистка полупроводников, и в окружающем воздухе: примеси, посторонние примеси подстерегают повсюду.

Да, принцесса на горошине — толстокожий бегемот по сравнению с самым обыкновенным кремнием! Ее Высочество была способна почувствовать горошину сквозь двенадцать тюфяков и двенадцать перинок. Кристалл кремния заметно меняет свои свойства, если на сто миллионов атомов кремния приходится один посторонний атом!

Правда, есть выход, с которым мы уже знакомы: нагревать кристалл. Тогда концентрация собственных носителей резко увеличивается. Но, во-первых, для нагревания кристалла имеется естественный пре-

дел: плавление или разложение его на составляющие элементы. Арсенид галлия, например, при попытке нагреть его слишком сильно просто распадется на мышьяк и галлий. Во-вторых, для кристаллов с большим значением $\mathcal{E}_g$ и нагревание — плохой помощник. Используя формулу (6), легко убедиться, что для GaP или SiC нагревание даже до 1 000 К не избавит эти вещества от «болезненной чувствительности» к примесям. И, наконец, наибольшее практическое значение все-таки имеют свойства полупроводников при комнатной температуре.

Теперь мы можем лучше представить себе трудности, с которыми столкнулись исследователи, начавшие изучать полупроводники. Они со всей возможной тщательностью очищали кристаллы и получали чистейшие вещества с концентрацией примесей $\leq 0,1\%$ (такая очистка на заре истории полупроводников была предметом законной гордости экспериментаторов). Они скрупулезно исследовали концентрацию электронов и дырок и ... Пока речь шла о высоких температурах, данные, полученные в различных лабораториях, еще кое-как согласовывались друг с другом. Но как только начиналось обсуждение результатов, полученных при сравнительно низких температурах, разноречивость в результатах оказывалась такой, как будто бы речь шла о совершенно различных веществах. (Вспомните историю исследования Si_2O , рассказанную в первой части книги.) Теперь мы понимаем в чем дело: в каждой лаборатории изучали один и тот же полупроводник, но со своим набором случайных примесей. И только бесконечному терпению и высокому экспериментальному искусству физиков, занимавшихся исследованием полупроводников, мы обязаны тем, что в хаосе результатов, обусловленных разными примесями и разными их комбинациями, удалось «нащупать» и правильно определить собственные свойства полупроводников.

Важный вывод, к которому можно прийти, изучая табл. II, состоит еще и в том, что при комнатной температуре собственные полупроводники с энергией связи $\mathcal{E}_g \geq 2$ эВ представляют собой превосходные изоляторы. В самом деле, концентрация собственных электронов в фосфиде галлия в 10^{23} раз меньше, чем в металлах. В карбиде кремния — в

10^{30} раз меньше! Даже собственный арсенид галлия, концентрация собственных электронов в котором в 10^{16} раз меньше чем в металлах, является очень хорошим изолятором. Возможность использования таких веществ для изготовления полупроводниковых приборов и устройств целиком зависит от того, какие примеси и в каких количествах удастся ввести в эти вещества. Поэтому примеси в полупроводниках далеко не всегда — мешающая «грязь». Подавляющее большинство практически используемых полупроводниковых материалов содержит тщательно контролируемое количество примесей, обеспечивающих необходимую величину проводимости.

Вот, в частности, почему нельзя разделить четко полупроводники и диэлектрики по величине $\mathcal{E}_g$. Алмаз, например, имеет величину $\mathcal{E}_g \approx 5,6$ эВ. Используя формулу (6), можно подсчитать, что шансы найти в алмазе один-единственный свободный электрон при комнатной температуре имеются только, если искать его в кристалле алмаза величиной с земной шар. Тем не менее, в научной литературе, посвященной полупроводникам, можно встретить выражение «полупроводниковый алмаз». Появилось оно около 30 лет назад, когда в копиях Южной Африки обнаружили голубые алмазы: каким-то образом в процессе естественного роста этих драгоценных камней в них попала примесь бора и сообщила им заметную *примесную проводимость*.

Примесные полупроводники

Донорная примесь. Начнем с самой простой ситуации. Пусть в кристалл кремния каким-то образом попал атом мышьяка (As) и занял место в одном из узлов кристаллической решетки, заместив в нем законного хозяина — атом Si. (Чтобы не рисовать в очередной раз кристаллическую решетку кремния, воспользуемся рис. 6 и будем для определенности считать, что атом As находится в узле 13.) Атом Si имеет четыре валентных электрона; атом As — пять. Четыре валентных электрона мышьяка будут участвовать в связях с соседними атомами кремния (атомами 8, 12, 14 и 18), а пятый? ... Пятый будет удерживаться атомом As, но гораздо слабее, чем остальные четыре, прочно связанные в элек-

тронных орбитах, определяемых всей структурой кристалла кремния. Энергия $\Delta\mathcal{E}$, которая нужна, чтобы разорвать связь этого пятого электрона с атомом мышьяка и превратить его в свободный электрон, гораздо меньше, чем энергия $\mathcal{E}_g$, которую необходимо затратить, чтобы разорвать связь между атомами кремния и образовать электронно-дырочную пару.

Примесь, атомы которой легко отдают электроны, называют *донорной*. Латинский глагол *donare*, от которого произошло слово «донор», означает дарить, жертвовать.

Пусть в каждый кубический сантиметр кристалла введено N_d атомов донора (например, мышьяка). Рассмотрим сначала самую простую ситуацию — температура кристалла $T = 0$ К. Ясно, что в этом случае кристалл опять-таки представляет собой идеальный диэлектрик: как ни мала энергия $\Delta\mathcal{E}$, требуемая, чтобы оторвать пятый электрон от атома мышьяка, при абсолютном нуле температуры ей взяться неоткуда.

Если температура кристалла отлична от нуля, то равновесная концентрация примесных электронов n_d определится из выражения, аналогичного формуле (6):

$$n_d = N_d e^{-\Delta\mathcal{E}/2kT}. \quad (7)$$

Вместо большой величины $\mathcal{E}_g$ в показателе экспоненты стоит теперь величина гораздо меньшая: $\Delta\mathcal{E}$. А мы уже убедились раньше, как сильно меняется величина экспоненты с изменением показателя.

Энергия ионизации $\Delta\mathcal{E}$ (иногда ее называют энергией активации примеси) для мышьяка, введенного в кремний, равняется всего 0,05 эВ, что в 20 раз меньше, чем энергия $\mathcal{E}_g$, требуемая для создания электронно-дырочной пары в кремнии. В соответствии с формулой (7) (и со здравым смыслом) это означает, что ионизация атомов мышьяка, т. е. отрыв лишнего пятого электрона, начнется при гораздо более низких температурах, чем генерация электронно-дырочных пар.

Полезно сравнить концентрацию собственных (n_i) и примесных (n_d) электронов при различных температурах. Для этого, прежде всего, необходимо знать концентрацию примеси N_d в формуле (7). Эта величина может меняться в очень широких пределах.

Предельно низкие значения N_d определяются совершенством технологии: тем, насколько хорошо мы умеем очищать полупроводник от примесей. Для карбида кремния, например, предельно низкие значения концентрации посторонних примесей составляют 10^{16} — 10^{17} см $^{-3}$. Для арсенида галлия — 10^{13} — 10^{14} см $^{-3}$. Кремний, из которого изготавливается сегодня подавляющая часть полупроводниковых приборов, технологи умеют очищать очень тщательно. Уровень «неконтролируемых» примесей в кремнии составляет 10^{10} — 10^{11} см $^{-3}$.

Обсудим теперь, чем определяется максимальное значение N_d в формуле (7). Если атомов примеси ввести в кристалл очень много, они окажутся близко друг от друга и начнут друг на друга влиять. Влияние это приводит к тому, что энергия ионизации $\Delta\mathcal{E}$ начинает зависеть от концентрации примесей N_d . С ростом N_d величина $\Delta\mathcal{E}$ уменьшается и при достаточно больших значениях N_d (10^{18} — 10^{19} см $^{-3}$) обращается в нуль. А это значит, что свободные электроны будут существовать в полупроводнике даже при температуре $T = 0$ К! Введение очень большого количества примесей сообщает полупроводнику свойства металла *). При этом возникает много важных и интересных эффектов, но в этой книге мы их рассматривать не будем.

Итак, величина N_d в формуле (7) может лежать в пределах от $\sim 10^{11}$ до $\sim 10^{19}$ см $^{-3}$. Выберем $N_d = 10^{15}$ см $^{-3}$; это значение N_d не только лежит в середине всего диапазона возможных значений, но и типично для многих практически важных случаев. Теперь мы можем сравнить значения n_i и n_d при разных температурах.

При 77 К расчет по формуле (7) дает для кремния значение $n_d = 2,3 \cdot 10^{13}$ см $^{-3}$ — на 30 порядков больше, чем величина n_i !

Полупроводник, в который введена донорная примесь, называют *электронным или полупроводником n-типа*. Термин «электронный» полупроводник вполне понятен. Что может быть естественней, чем на-

*) О том, почему и как происходит переход полупроводника в металл при увеличении концентрации примесей, рассказывается в книге: Эфрос А. Л. Физика и геометрия беспорядка. — М.: Наука, 1982. Библиотечка «Квант», вып. 19.

зывать так полупроводник, в котором почти все свободные носители тока — электроны. Что касается термина «полупроводник n -типа», то латинская буква « n » — от слова *negativus*, отрицательный, опять-таки напоминает, что такой полупроводник содержит много отрицательно заряженных частиц, в данном случае электронов.

При 300 К величина n_d , рассчитанная по формуле (7), равняется $\sim 4 \cdot 10^{14}$ см $^{-3}$, т. е. в 40 000 раз больше, чем указанное в табл. II значение n_i . В действительности же разница будет еще большей.

Примесное истощение. Дело в том, что формула (7) является не точной, а приближенной. Она справедлива только тогда, когда рассчитанное по ней значение n_d много меньше, чем концентрация введенных примесей N_d , т. е. когда $kT \ll \Delta\mathcal{E}$. Когда температура T так высока, что величина $kT \sim \Delta\mathcal{E}$, все примесные атомы оказываются ионизованными и концентрация примесных электронов просто равняется N_d . Это явление носит название примесного истощения — термин довольно естественный. При достаточно высокой температуре все атомы примеси отдали «лишние» электроны и источник примесных электронов истощился. Дальнейший нагрев кристалла не приводит к росту концентрации примесных электронов.

Снова собственная проводимость. Пусть $T = 1000$ К. Концентрация примесных электронов по-прежнему равняется 10^{15} см $^{-3}$. А концентрация собственных? Расчет по формуле (6) дает значение $n_i \approx 3,3 \cdot 10^{16}$ см $^{-3}$, что в 30 раз больше, чем концентрация примесных электронов. При повышенных температурах собственная концентрация электронов значительно превысила концентрацию примесных. Проводимость кристалла снова определяется не концентрацией введенной примеси, а собственными свойствами полупроводника.

Зависимость концентрации электронов от температуры. Взгляните на рис. 17. Кривая, которая на нем показана, суммирует многое из того, что мы узнали о свойствах собственных и примесных полупроводников. Эта кривая представляет собой типичную зависимость равновесной концентрации свободных электронов в полупроводнике от температуры. Правда, для графического изображения этой зависимости

выбраны несколько необычные координаты: по оси ординат откладывается не концентрация n , а значение ее натурального логарифма, $\ln n$. По оси абсцисс — не температура, а величина, обратная температуре, $1/T$. Читать такой график непривычно, в особенности потому, что продвижению по оси абсцисс

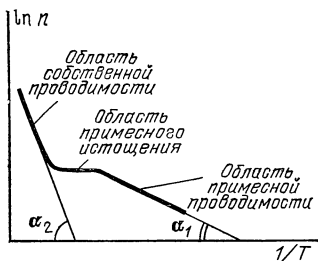


Рис. 17. Зависимость логарифма концентрации электронов от обратной температуры. $\operatorname{tg} \alpha_1 = \Delta \mathcal{E}/2k$, $\operatorname{tg} \alpha_2 = \mathcal{E}_g/2k$.

вправо соответствует не увеличение, а уменьшение температуры. Почему для изображения этой зависимости выбраны именно такие координаты, мы обсудим чуть позже. А пока просто прочитаем график качественно. Начнем с области низких температур, т. е. с правой части графика, где значение $1/T$ велико, и будем двигаться по оси абсцисс справа налево.

При низких температурах, как мы видели, концентрация электронов в полупроводнике определяется наличием примесей. С ростом температуры примесная концентрация растет и на этом участке кривой зависимость $n(T)$ определяется формулой (7). При некоторой температуре концентрация электронов перестает зависеть от температуры. Это — область примесного истощения. Все атомы примеси уже ионизованы, а собственная концентрация все еще гораздо меньше, чем примесная. Наконец, в области еще более высоких температур начинается очень резкий рост концентрации с дальнейшим повышением температуры. Это — область собственной проводимости, где зависимость $n(T)$ определяется формулой (6).

А теперь несколько слов о том, почему для зависимости, показывающей, как связана концентрация электронов в полупроводнике с температурой, выбраны именно такие координаты. Прологарифмируем формулу (6):

$$\ln n_i = \ln A - \mathcal{E}_g/2kT \quad (8)$$

и обозначим $x = 1/T$, $y = \ln n_i$, $a = \ln A$, $b = \mathcal{E}_g/2k$. Тогда формула (8) приобретет знакомый вид: $y =$

$= a - bx$ — уравнение прямой линии. Итак, если по оси ординат откладывать $\ln n$, а по оси абсцисс — $1/T$, то до тех пор, пока справедливо уравнение (6), зависимость $\ln n$ как функции $1/T$ должна представлять собой прямую линию. И что особенно важно, тангенс угла наклона этой прямой $b = \operatorname{tg} \alpha_2 = = \mathcal{E}_g/2k$ просто пропорционален важнейшему параметру полупроводника $\mathcal{E}_g$. В ранних работах по полупроводникам такой способ определения $\mathcal{E}_g$ использовался очень часто. Его и сейчас еще иногда используют при изучении свойств новых полупроводников.

Если такую же операцию логарифмирования проводить с уравнением (7), то легко убедиться, что тангенс угла наклона прямой пропорционален $\Delta\mathcal{E}$. Таким образом, построив зависимость $\ln n$ как функции $1/T$ в области примесной проводимости, мы легко можем определить величину энергии ионизации примеси $\Delta\mathcal{E}$.

Дырки в электронном полупроводнике. Наш рассказ о полупроводниках n -типа будет не полон, если мы оставим без внимания вопрос, что произойдет с дырками в полупроводнике, в который введена донорная примесь.

Прежде всего нужно отметить, что когда атом донора ионизуется и отдает свободный электрон, никакой дырки не образуется. Атом мышьяка в решетке кремния, например, отдав пятый валентный электрон, остается связанным с соседними атомами кремния четырьмя полноценными связями. Никакой свободной связи, на которую могли бы перемещаться электроны с соседних орбит, нет. Дырка не образуется.

На первый взгляд может показаться, что дырок в электронном полупроводнике будет тогда столько же, сколько и в совершенно чистом, собственном. Однако этот первый взгляд неверен. Действительно, при данной температуре T каждую секунду в электронном полупроводнике рвется электронных связей между атомами и образуется электронно-дырочных пар ровно столько же, сколько и в собственном. Но гибнут дырки гораздо чаще. Ведь в электронном полупроводнике свободных электронов гораздо больше, чем в собственном. Значит, и встреча дырки со свободным электроном, встреча, которая заканчивается

рекомбинацией (и исчезновением дырки), будет происходить гораздо чаще. В результате дырок в полупроводнике n -типа оказывается меньше, чем в собственном полупроводнике. Существует очень простая формула, связывающая равновесную концентрацию дырок p и электронов n в полупроводнике:

$$pn = n_i^2(T), \quad (9)$$

где n_i — собственная концентрация носителей при температуре T , определяемая выражением (6). Из формулы (9) следует, что чем больше электронов, тем меньше дырок (и наоборот).

Наше знакомство с примесными полупроводниками далеко не закончено. Мы познакомились с электронным полупроводником, в который введена донорная примесь. Изучим теперь другую ситуацию,

а именно: в полупроводник введена примесь, которая не отдает (как донорная), а легко отбирает электроны у соседних атомов.

Акцепторная примесь*). Не будем на этот раз лениться, и снова нарисуем кристаллическую решетку кремния, в которой некоторые из узлов заняты посторонними примесными атомами. Но теперь примесью является не мышьяк, а бор (рис. 18). Бор (В) трех-

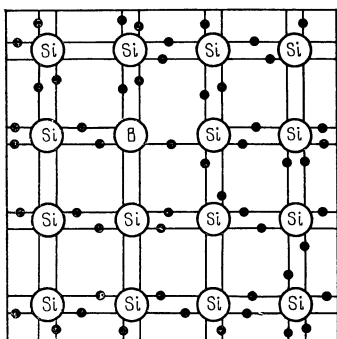


Рис. 18. Атом акцептора (бор) в решетке кремния.

валентен — на внешней электронной оболочке у него находятся три электрона. Чтобы образовать полноценную связь с соседними атомами кремния, одного электрона ему не хватает.

Сравните «рисунок» электронных связей вокруг атома бора на рис. 18 с конфигурацией электронных связей вокруг атома 19 на рис. 6. Не правда ли, очень схожие ситуации? И атому кремния 19 на рис. 6, и атому бора на рис. 18 не хватает одного электрона. Но имеется и глубокое различие. Ведь

*) Латинский глагол *accipere* означает принимать.

все атомы кремния совершенно одинаковы и пустая связь — дырка, которая в данный момент принадлежит атому 19, в следующий момент может оказаться у атома 14, затем 9 и т. д. Для того чтобы дырка перемещалась по кристаллу, никакой энергии не требуется. Атом бора — чужак в решетке кремния. Для того чтобы электрон от соседнего атома кремния перешел к бору, необходимо затратить энергию $\Delta\mathcal{E}$. Эта энергия — энергия активации, невелика (для атома бора в кремнии она составляет всего 0,045 эВ), но все-таки не равна нулю. Пока этот, пусть небольшой, энергетический барьер не перейден, дырки в кристалле нет.

Предположим, однако, что тепловые колебания решетки или квант света сообщили необходимую энергию, и электрон с соседнего атома кремния перешел к бору. Вот теперь ситуация абсолютно ничем не отличается от той, которая показана на рис. 6. Пустая связь образовалась у атома кремния и возникла дырка — свободный носитель положительного заряда, способный переносить электрический ток. Акцепторная примесь создает в кристалле полупроводника свободные носители тока — дырки.

Знакомая песня с одним новым словом. Теперь мы могли бы слово в слово переписать несколько предыдущих страниц, посвященных описанию электронного полупроводника, заменяя слово «электрон» словом «дырка».

Пусть в каждый кубический сантиметр кристалла введено N_a атомов бора. Рассмотрим сначала самую простую ситуацию — температура кристалла $T = 0$ К. Ясно, что в этом случае кристалл представляет собой идеальный диэлектрик: как ни мала энергия активации $\Delta\mathcal{E}$, требуемая, чтобы перевести электрон от атома кремния к атому бора и образовать дырку, при нулевой температуре ей взяться неоткуда.

Если температура кристалла отлична от нуля, то равновесная концентрация примесных дырок p_a определится из выражения, аналогичного формуле (7):

$$p_a = N_a e^{-\Delta\mathcal{E}/2kT} \quad (10)$$

и т. д.

Не будем заводить шутку слишком далеко. Но ситуации с донорной и акцепторной примесью действительно очень похожи. Типичный диапазон кон-

центратий примесных атомов для акцепторной примеси такой же, как и для донорной. Так же, как и в случае донорной примеси, величина $\Delta\mathcal{E}$ для дырки уменьшается с ростом N_a и при достаточно больших значениях N_a обращается в нуль (рис. 19). Поскольку величина $\Delta\mathcal{E}$ гораздо меньше, чем энергия образования электронно-дырочной пары $\mathcal{E}_g$ (кстати, обратите внимание, что энергии активации донорной примеси — мышьяка и акцепторной — бора очень близки), то при относительно низких температурах

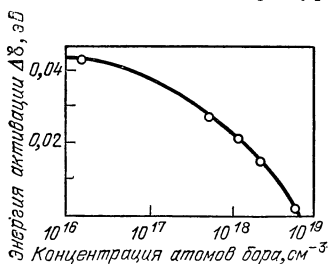


Рис. 19. Зависимость энергии активации дырки $\Delta\mathcal{E}$ от концентрации атомов бора в решетке кремния.

Сплошная кривая — теоретический расчет. Точки — экспериментальные результаты. При концентрации бора $N_a \approx 8 \cdot 10^{18} \text{ см}^{-3}$ энергия активации обращается в нуль — кремний превращается в полуметалл.

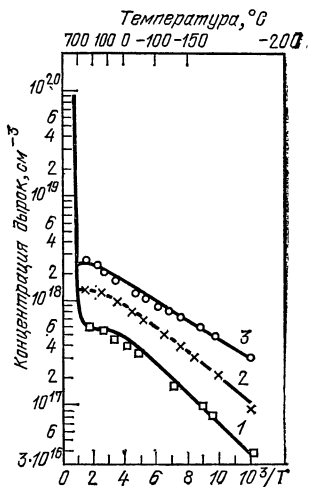


Рис. 20. Зависимость концентрации дырок от температуры для кремния, легированного бором.

Приведены данные для трех различных образцов с различным содержанием бора. Образец 1 — концентрация атомов бора $N_a = 6 \cdot 10^{17} \text{ см}^{-3}$; 2 — $N_a = 1,3 \cdot 10^{18} \text{ см}^{-3}$; 3 — $N_a = 2,2 \cdot 10^{18} \text{ см}^{-3}$. Результаты получены в той же работе Бардина и Пирсона, из которой взята зависимость, показанная на рис. 19.

экспонента в формуле (10) будет гораздо больше, чем экспонента в формуле (6). Это значит, что концентрация примесных дырок p_a будет значительно больше, чем концентрация собственных *). При более высоких температурах, когда $kT \sim \Delta\mathcal{E}$, наступает явление примесного истощения. Наконец при

*) Полупроводник, в который введена акцепторная примесь, называют дырочным или полупроводником *p*-типа. Латинская буква «*p*» — от слова *positivus* (положительный).

еще более высоких температурах собственная концентрация, определяемая формулой (6), превысит примесную.

На рис. 20 показаны экспериментальные зависимости, полученные американскими физиками Бардином и Пирсоном и наглядно иллюстрирующие все изученные нами особенности поведения «мелкой» примеси в полупроводнике. Мелкой принято называть примесь, энергия активации которой $\Delta\mathcal{E}$ значительно меньше, чем энергия образования электронно-дырочной пары $\mathcal{E}_g$.

Рассмотрим рис. 20 внимательно, он очень информативен. Прежде всего обратите внимание на обозначения по осям координат. По оси абсцисс отложена не величина $1/T$, а в тысячу раз большая — $10^3/T$. Это сделано для того, чтобы цифры на оси абсцисс соответствовали небольшим целым числам и график удобно было читать. По оси ординат отложены значения, пропорциональные логарифму концентрации дырок, а цифры возле делений прямо указывают величину концентрации. И это сделано для удобства чтения.

Из рис. 20 видно прежде всего, что в области примесной проводимости зависимости $\ln p$ как функции $1/T$ идут тем более полого, чем больше примеси содержит полупроводник. Этот факт отражает уменьшение $\Delta\mathcal{E}$ при больших концентрациях примеси N_a , который уже иллюстрировался (см. рис. 19). Очень хорошо видна на рис. 20 область примесного истощения, когда концентрация дырок перестает зависеть от температуры и просто равняется концентрации введенной примеси N_a . При высоких температурах ($\geq 700^\circ\text{C}$) наступает область собственной проводимости. Заметим, что чем выше уровень легирования полупроводника примесью (чем больше величина N_a), тем при более высоких температурах начинает проявляться собственная проводимость.

Появление дырки в примесном полупроводнике не сопровождается появлением свободного электрона — это хорошо видно из рис. 18. Наоборот, в дырочном полупроводнике (или полупроводнике p -типа) увеличение концентрации дырок сопровождается уменьшением концентрации электронов. Количественное соотношение между концентрациями электронов и дырок

(9) одинаково справедливо и для n - и для p -типа полупроводников.

Мы подробно изучили поведение мелких примесей в полупроводнике. Внимание, которое мы им уделили, вполне оправдано. При создании почти любого полупроводникового прибора в тщательно очищенный полупроводниковый материал вводится определенное, точно рассчитанное количество мелкой примеси. И именно концентрация примесных электронов (или дырок) определяет, как правило, параметры приборов.

Ах, как кружится голова ...

Примеры, на которых мы изучили поведение мелких примесей — мышьяк и бор в решетке кремния, демонстрируют типичные черты поведения мелкой примеси в любом полупроводнике.

В элементарных полупроводниках, принадлежащих к четвертой группе элементов таблицы Менделеева, германии (Ge) и кремнии (Si), элементы пятой группы — фосфор, мышьяк, сурьма — служат мелкой донорной примесью с энергией ионизации $\sim 0,05$ эВ. Элементы третьей группы — бор, алюминий, галлий, индий — мелкой акцепторной примесью.

В полупроводниковом соединении арсениде галлия (GaAs) мелкими донорами являются сера, селен и теллур — элементы шестой группы таблицы Менделеева. Мелкими акцепторами являются цинк, кадмий, магний и бериллий — элементы второй группы. (Посмотрите в таблицу Менделеева и постарайтесь понять, почему.) Элементы четвертой группы — Ge и Si — являются в GaAs *амфотерными примесями*. Греческое слово *amphoteros* означает обоюдный, двусторонний, двойственный. В зависимости от того, на место какого атома попадают в арсениде галлия атомы Ge или Si, они могут быть и мелкими донорами, и мелкими акцепторами. Если атом Si (или Ge) заместит в кристаллической решетке GaAs атом галлия, он будет играть роль мелкого донора. В случае, если он попадет на место атома мышьяка, — роль мелкого акцептора. (Постарайтесь понять сами, почему это так.)

В карбиде кремния (SiC) относительно мелкой донорной примесью служит азот, мелкими акцепторами — бор и алюминий.

Для каждого полупроводника, для каждого полупроводникового соединения имеется свой набор элементов, являющихся в нем донорами или акцепторами. Часто характер примеси (донорная или акцепторная) и энергия ее активации зависит от того, какое место в кристаллической решетке займет элемент. Кроме того, примесь почти никогда не попадает в полупроводник одна, в чистом виде. Почти всегда в материале содержатся одновременно несколько видов примесей. Многие примеси обладают способностью менять свои свойства при взаимодействии друг с другом и с несовершенствами кристаллической решетки, всегда присутствующими в реальном полупроводнике. А некоторые несовершенства кристаллической решетки, в свою очередь, могут играть роль доноров или акцепторов.

И все это происходит в материалах, свойства которых радикально меняются от добавления примеси в количестве один атом примеси на сто миллионов атомов полупроводника! Не стоит ли еще раз снять шляпу перед технологами, которые научились с таким совершенством очищать полупроводники, и исследователями, разработавшими методы, с помощью которых изучается поведение примесей в полупроводниках?

Остановимся еще на двух важных и интересных вопросах, связанных с поведением примесей в полупроводнике.

Компенсация

Первый вопрос: что произойдет, если ввести в полупроводник одновременно в строго одинаковых количествах мелкую донорную и акцепторную примеси? Например, в кремний ввести совершенно одинаковое количество мышьяка и бора.

Естественный ответ: в полупроводнике возникнет большая концентрация свободных носителей обоих знаков — электронов и дырок.

Этот ответ совершенно неправилен. Правильный ответ состоит в том, что в полупроводнике не будет

ни примесных электронов, ни примесных дырок — он будет вести себя как собственный полупроводник.

Действительно, атомы донорной примеси легко отдадут лишние электроны, которые станут свободными. Атомы акцепторной примеси захватят электроны у своих ближайших соседей по решетке и образуют соответствующее количество дырок. Примесные электроны и дырки начнут хаотически блуждать по кристаллу и рано или поздно встретятся и рекомбинируют. В соответствии с условием задачи, которую мы рассматриваем, количество электронов, которые легко могут стать свободными, в точности равно количеству дырок. (Поскольку концентрация мелких доноров равна концентрации акцепторов.) После того как примесные электроны и дырки встретятся, рекомбинируют и исчезнут, в кристалле не останется больше атомов, которые могут легко образовать электрон или дырку. В самом деле, у каждого атома мышьяка осталось теперь только по четыре валентных электрона. И эти оставшиеся электроны атом мышьяка удерживает почти так же прочно, как и атом кремния свои четыре валентных электрона. То же можно сказать и об атоме бора. Присоединив к себе четвертый электрон, он образовал с соседними атомами кремния связи почти такие же прочные, как и связи между атомами кремния. В результате, чтобы образовать свободный электрон или дырку, нужно теперь затратить большую энергию — почти равную энергии $\mathcal{E}_g$.

Явление, которое мы рассмотрели, носит название *компенсации*: введение акцепторной примеси компенсирует действие донорной (и наоборот).

Если концентрации донорной N_d и акцепторной N_a примесей не равны, то произойдет *частичная компенсация*. Например, если донорной примеси больше, то полупроводник будет вести себя как полупроводник *n*-типа, концентрация доноров в котором равняется $N_d - N_a$.

Перейдем ко второму вопросу. Вопрос этот так важен, что мы посвятим ему отдельную главу.

КОМПЕНСАТОРЫ, СТУПЕНЬКИ, УБИЙЦЫ

Хотите — буду от мяса бешеный,
 — и, как небо, меняя тона —
 Хотите —
 буду безукоризненно нежный,
 не мужчина, а — облако в штанах.
В. Маяковский. «Облако в штанах»

До сих пор, рассматривая поведение примеси в кристаллической решетке полупроводника, мы все время говорили о мелкой примеси. Нет, однако, никаких оснований ожидать, что любая примесь, введенная в полупроводник, будет вести себя как мелкая. То есть, что энергия активации $\Delta\mathcal{E}$ любого примесного донорного атома или любого акцепторного во всех случаях будет много меньше, чем энергия образования собственной электронно-дырочной пары $\mathcal{E}_g$.

Лучше всего изучено к настоящему времени поведение примесей в германии и кремнии. Список элементов, поведение которых в качестве примеси в кремнии достаточно подробно исследовано, насчитывает несколько десятков наименований. Он включает уже знакомые нам фосфор, мышьяк, бор, алюминий, но кроме них — серебро, медь, кадмий, кобальт, золото, железо, кислород, ртуть, платину, молибден, никель, палладий, серу, селен, вольфрам, цинк и многие другие элементы. Каждый из этих элементов характеризуется в качестве примеси в кремнии своей энергией активации $\Delta\mathcal{E}$, и величины этих энергий лежат в пределах от $\Delta\mathcal{E} \sim 0,05$ эВ (мелкие доноры и акцепторы) до $\Delta\mathcal{E} \sim 0,8 - 0,9$ эВ. Последние значения, характерные для самых «глубоких» примесей, уже вполне сравнимы с энергией $\mathcal{E}_g$ в кремнии (1,1 эВ).

Казалось бы, чем «глубже» примесь (т. е. чем больше $\Delta\mathcal{E}$), тем меньшую роль она должна играть. Действительно, из формул (7) и (10) ясно, что с увеличением $\Delta\mathcal{E}$ очень быстро, экспоненциально падает число свободных электронов и дырок, которые будут возникать в кристалле при тех же концентрациях легирующей примеси N_d или N_a . Формулы (7) и (10) вполне справедливы для любых величин $\Delta\mathcal{E}$. Однако вывод о том, что на глубокие примеси можно не обращать внимания, поспешен и неверен.

В драме с участием главных героев — электронов и дырок, глубокие примеси играют три очень важные роли — «компенсаторов», «ступенек» и «убийц».

Прорецензируем эти роли одну за другой.

Начнем с самой простой и отчасти нам уже знакомой роли компенсаторов.

Компенсаторы

Представим себе, что мы хотим получить полупроводник, например GaAs, с очень высоким удельным сопротивлением. Такая задача не так уж редко возникает перед специалистами, изготавливающими полупроводниковые приборы. Пленки или пластины с высоким удельным сопротивлением часто используются в качестве основы, «подложки», на которой затем наращиваются очень тонкие слои этого же полупроводника с различными легирующими примесями.

Самое высокое удельное сопротивление имеет собственный полупроводник: в нем меньше всего свободных носителей заряда. Однако того, что мы уже знаем о примесях в полупроводниках, вполне достаточно, чтобы оценить трудности получения собственных полупроводников. Ничтожная концентрация мелкой примеси и... концентрация электронов или дырок возрастает в сотни миллиардов раз. Во всех случаях попытки совершенно избавиться от мелких примесей сопряжены с длительными, дорогими и сложными операциями по очистке материалов. Но ощутимые результаты они приносят в настоящее время только для германия и кремния. Для GaAs, например, самая совершенная современная технология обеспечивает уровень концентрации мелкой примеси $N \sim 10^{13} \text{ см}^{-3}$. Эта величина в десять миллиардов раз меньше, чем концентрация атомов галлия и мышьяка, однако она все еще в миллион раз больше, чем концентрация собственных электронов и дырок в этом материале при комнатной температуре (см. табл. II). Как же быть?

А очень просто. Не будем особенно тщательно очищать полупроводник. Пусть в нем содержится 10^{15} или даже 10^{16} см^{-3} мелких доноров. А кроме того, мы дополнительно введем в полупроводник примесь и даже в еще большем количестве. Допустим,

концентрация этой дополнительной примеси составит 10^{16} — 10^{17} см⁻³, но только примесь эта будет глубоким акцептором.

Возникнет эффект, с которым мы уже знакомы,— компенсация. Только не мелкими, а глубокими примесями. Глубокие акцепторы захватят все электроны, созданные мелкими донорами, и теперь освободить электроны будет не так-то просто. Энергия активации этих электронов с глубоких примесей велика.

Можно, конечно, пытаться компенсировать мелкую донорную примесь мелкой акцепторной. Но тут требуется величайшая тщательность. Нужно, чтобы величины N_d и N_a равнялись друг другу с большой точностью. Добиться этого очень трудно.

То ли дело глубокая примесь! Даже если ввести ее в концентрации в несколько раз большей, чем концентрация мелкой примеси, которую нужно компенсировать, ничего страшного не произойдет. Все носители, созданные мелкой примесью, будут с гарантией захвачены. А лишние акцепторы, скажем? Ведь они способны захватывать электроны у атомов полупроводника и образовывать дырки. Способны, но... на то они и глубокие,— энергия активации дырки $\Delta\mathcal{E}$ у таких акцепторов велика и, соответственно, концентрация дырок окажется очень небольшой.

В случае GaAs именно так и поступают. Берут не особенно тщательно очищенный GaAs и вводят в него в значительной концентрации хром или кислород. Если до введения глубокой примеси GaAs был полупроводником *p*-типа, то вводят кислород; в GaAs он является глубоким донором. Если исходный GaAs был полупроводником *n*-типа, вводят хром. В GaAs хром является глубоким акцептором. В результате удастся получить монокристаллы GaAs с очень высоким удельным сопротивлением — до 10^7 Ом·см. Подложки из такого полужолирующего арсенида галлия широко используются при изготовлении полупроводниковых приборов на основе GaAs.

Итак, в первом из рассмотренных случаев глубокие примеси играют роль положительного героя. Но ... «наши недостатки являются продолжением наших достоинств». Эта мудрая пословица применима не только к людям. Ее полезно вспомнить и при

обсуждении свойств глубоких примесей. Представим себе другой, тоже очень распространенный случай. Мы хотим получить не полуизолирующую подложку, а материал, скажем, *n*-типа. Мы вводим и вводим в кристалл мелкую донорную примесь, а полупроводник остается высокоомным. Легко догадаться, в чем здесь дело: в полупроводнике присутствует глубокий акцептор. И пока мы его не «перекомпенсируем», т. е. не введем мелкой примеси больше, чем имеется глубоких акцепторов, повышение уровня легирования ни к чему не приведет.

Между тем, полупроводник совсем не безразличен к общему количеству введенной в него примеси. Если примесей ввести слишком много, возникают различные структурные несовершенства, падает скорость движения электронов и дырок в электрическом поле, возникают другие нежелательные явления. Поэтому если мы хотим получить чистые кристаллы с заданной концентрацией электронов (или дырок), мы в первую очередь должны научиться, когда это необходимо, избавляться от глубоких примесей. Технологи прекрасно справляются с этой задачей в германии и кремнии — самых «старых» полупроводниковых материалах. В последние годы большие успехи достигнуты в технологии арсенида галлия. А вот в карбиде кремния, например, минимальный уровень содержания глубоких примесей, которого удастся достигнуть, $\sim 10^{16} \text{ см}^{-3}$.

Ступеньки

Мы знаем, что для того чтобы в собственном полупроводнике возникли электрон и дырка, необходимо затратить энергию $\mathcal{E}_g$. Изобразим процесс возникновения электронно-дырочной пары схематически (рис. 21, *a*). Будем считать, что пока электрон находится на орбите, связывающей между собой атомы кремния, он имеет энергию $\mathcal{E}_1$. Чтобы сделать электрон свободным и образовать электрон и дырку, необходимо преодолеть энергетический барьер высотой $\mathcal{E}_g$. Если энергию $\mathcal{E}_g$ нужно получить от тепловых колебаний решетки, а величина $\mathcal{E}_g$ много больше, чем средняя тепловая энергия kT , то такой процесс очень маловероятен. Как мы знаем, он пропорционален $e^{-\mathcal{E}_g/kT}$ (формула (2)). Возникает

вопрос, нельзя ли энергию, необходимую для того, чтобы освободить электрон, затратить не сразу, а в два этапа, как это показано на рис. 21, б. Сначала сообщить энергию $\mathcal{E}_g/2$, а затем — вторую порцию, тоже, естественно, равную $\mathcal{E}_g/2$. Легко понять, что если бы этого как-нибудь удалось достичь, процесс образования электронов и дырок шел бы гораздо интенсивнее.

Действительно, пусть речь идет, например, об арсениде галлия при комнатной температуре. Величина $\mathcal{E}_g$ равняется 1,4 эВ, $kT = 0,026$ эВ. Вероятность образования электронно-дырочной пары пропорциональна $e^{-\mathcal{E}_g/kT} \approx 5 \times 10^{-25}$. Если снизить величину $\mathcal{E}_g$ в 2 раза, значение экспоненты составит $e^{-\mathcal{E}_g/2kT} \approx 7 \cdot 10^{-13}$. Уменьшение необходимой порции энергии в 2 раза привело к увеличению вероятности процесса в триллион раз!

Ситуация очень схожа с изображенной на рис. 22. Если требуется преодолеть барьер высотой в 2 м без

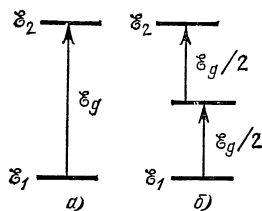


Рис. 21. Схема возникновения электронно-дырочной пары.

а — энергетический барьер преодолевается сразу; б — в два этапа.

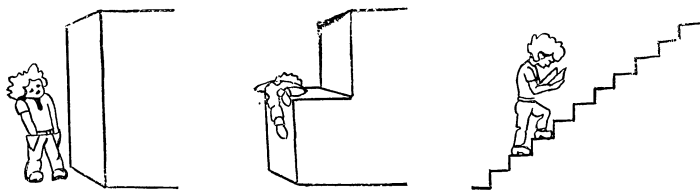


Рис 22. Высота одна и та же, но подняться по лестнице гораздо легче.

каких-либо промежуточных ступенек, то на такой подвиг способны лишь несколько человек из нескольких миллиардов жителей нашей планеты. Тот же барьер, разбитый на две метровых ступеньки, способен преодолеть, наверное, каждый десятый. Разбитый на восемь ступенек по 25 см, такой барьер представляет собой удобную лестницу.

А теперь убедимся, что глубокие уровни, энергия активации которых составляет $\mathcal{E}_g/2$, и являются

теми «ступеньками», теми «уровнями», которые показаны на рис. 21, б.

Пусть в полупроводник введены акцепторные примесные центры с энергией активации $\Delta\mathcal{E} \approx \mathcal{E}_g/2$. С вероятностью, которую мы только что рассчитали, акцептор отберет электрон у соседнего атома полупроводника. Образуется дырка, которая начнет хаотическое путешествие по кристаллу. Раньше, рассматривая поведение мелкого акцептора, об электро-не, захваченном акцепторным атомом, мы забывали. И правильно делали. Маленькая энергия активации $\Delta\mathcal{E}$ соответствует очень маленькой ступеньке вверх по энергетической лестнице (рис. 23, а). Для того чтобы освободить электрон с мелкого акцептора и сделать его свободным, нужна по-прежнему очень большая энергия, практически равная $\mathcal{E}_g$. Другое дело теперь, когда электрон захвачен глубоким акцептором. Тепловые колебания решетки с доста-

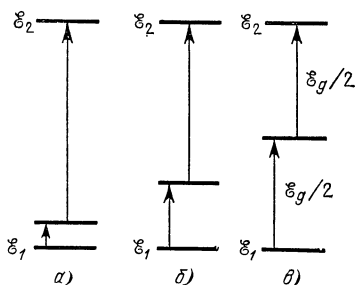


Рис. 23. а) Мелкий акцепторный уровень. Электрон легко захватывается на него, но освободить электрон с такого уровня и сделать свободным почти так же трудно, как образовать электронно-дырочную пару в собственном полупроводнике; б) более глубокий акцепторный уровень; в) уровень, энергия активации которого $\Delta\mathcal{E} = \mathcal{E}_g/2$ обеспечивает наибольшую скорость генерации электронно-дырочных пар.

точно большой вероятностью могут оторвать электрон от акцептора и сделать его свободным. В результате такого двухступенчатого процесса в кристалле образовалась электронно-дырочная пара. А акцептор снова приобрел возможность отобрать электрон у соседнего атома полупроводника и образовать следующую дырку и т. д. Скорость генерации электронно-дырочных пар резко возрастает.

Заметим, что наиболее эффективен такой двухступенчатый процесс именно для примеси с энергией активации $\Delta\mathcal{E} \approx \mathcal{E}_g/2$. Действительно, если энергия активации акцептора меньше, чем $\mathcal{E}_g/2$, то электрон будет сравнительно легко на него захватываться, но вот освободить его и сделать свободным — будет трудно (рис. 23, б). Полная аналогия с лестницей: если

нужно забраться на высоту 2 м и единственная ступенька имеет высоту 30 см, то на эту-то ступеньку взобраться легко, а вот оставшиеся 1 м 70 см преодолеть почти так же трудно, как и исходную высоту. Самый оптимальный вариант — две одинаковые ступеньки (рис. 23, в).

Такой же удобной «ступенькой-уровнем» для образования электронно-дырочных пар служат и глубокие доноры, энергия активации которых близка к величине $\mathcal{E}_g/2$. Проанализируйте сами подробнее процессы образования электронно-дырочных пар с участием глубоких доноров.

Итак, *введение глубоких примесей резко увеличивает скорость генерации электронно-дырочных пар*. Но позвольте! Как же быть с тем, что мы уже знаем о свойствах примесей? Как быть с формулой (9), из которой следует, что произведение концентраций электронов и дырок для каждого полупроводника зависит только от температуры, а вовсе не от того, какие в него введены примеси? Как быть с формулами (7) и (10), из которых следует, что чем больше энергия активации $\Delta\mathcal{E}$, тем меньшую концентрацию электронов (или дырок) будет такая примесь создавать при тех же значениях N_a и T ?

Оснований для беспокойства, тем не менее, нет никаких. Формулы (7), (9) и (10) справедливы для любых уровней — и глубоких, и мелких. Ведь эти формулы описывают значения концентрации электронов и дырок, которые устанавливаются в результате двух конкурирующих процессов — генерации и рекомбинации.

Введение в полупроводник *глубоких центров* резко увеличивает скорость генерации электронов и дырок. Но ... так же *резко увеличивается и скорость рекомбинации через эти примесные центры*. В результате формулы (7), (9) и (10) по-прежнему правильно описывают равновесные значения электронов и дырок.

В связи с таким утверждением естественно возникают два вопроса. Первый — почему глубокие центры увеличивают скорость рекомбинации электронов и дырок? Второй — если введение глубоких примесей увеличивает и скорость генерации, и скорость рекомбинации электронно-дырочных пар, так что равновесная концентрация электронов и дырок все равно определяется только энергией активации $\Delta\mathcal{E}$, то какое

значение имеет то обстоятельство, что именно примеси с энергией $\Delta\mathcal{E} \approx \mathcal{E}_g/2$ являются наиболее эффективными «ступеньками»? Как в этом убедиться? Существуют ли «пьесы», в которых глубокие примеси-ступеньки выступают в главной роли?

Ответ на первый вопрос будет получен немножко дальше. Там, где мы будем рецензировать выступление глубоких центров в роли «убийц».

Что касается ответа на второй вопрос, то следует сказать, что роль глубоких примесей-ступенек является одной из самых важных во многих полупроводниковых приборах и устройствах. В частности, во всех полупроводниковых диодах и транзисторах.

Однако для того чтобы центры-ступеньки смогли проявить себя, необходимы определенные условия.

Равновесные и неравновесные условия. До сих пор мы все время обсуждали самую простую *равновесную* ситуацию, когда число создаваемых тепловым движением носителей в точности равнялось числу носителей, гибнущих за то же время вследствие рекомбинации. В таких условиях проявить себя примеси-ступеньки не в состоянии. Во сколько раз увеличивается, благодаря их присутствию, скорость генерации носителей, во столько же раз растет и скорость рекомбинации. Равновесная концентрация носителей определяется формулами (7), (9) и (10).

Чтобы сделать явной деятельность центров-ступенек, нужно нарушить равновесие между процессами генерации и рекомбинации, создать резко *неравновесную* ситуацию. Удобнее всего это осуществить, выбрав условия, когда процесс генерации будет идти так же, как и раньше, в условиях равновесия, а процесс рекомбинации будет полностью устранен. Как этого добиться?

Оказывается, не так уж сложно. Нужно создать в какой-либо области полупроводника сильное электрическое поле. Тогда свободные носители, электроны и дырки, будут очень быстро выноситься из этой области электрическим полем и их концентрация в этой области будет очень маленькой. Раз так, скорость рекомбинации, пропорциональная произведению концентраций электронов и дырок (см. формулу (3)), резко уменьшится — рекомбинировать будет некому. А процесс генерации? На него электрическое поле влияет слабо, Электрическое поле, способное

практически полностью очистить область полупроводника от свободных носителей, на скорость генерации может почти не повлиять *).

Электрическое поле, выносящее свободные носители, обуславливает появление тока в цепи. И ток этот тем больше, чем больше свободных носителей генерируется в полупроводнике. Величина тока в цепи прямо пропорциональна скорости генерации электронно-дырочных пар.

Теперь условия для успешного выступления примесей в роли ступенек созданы.

Пусть сначала участок сильного поля создан в полупроводнике, в котором нет глубоких центров. Измерим протекающий в таких условиях ток. Затем введем в полупроводник глубокие центры и снова измерим величину тока. Прodelывая такие эксперименты, можно убедиться, что наличие даже небольшой ($10^{11} - 10^{12} \text{ см}^{-3}$) концентрации примесей с энергией активации $\Delta\mathcal{E} \approx \mathcal{E}_g/2$ увеличивает ток в тысячи и десятки тысяч раз. С другой стороны, примеси, чья энергия активации заметно отличается от величины $\mathcal{E}_g/2$, влияют на ток гораздо слабее. Мелкие доноры и акцепторы практически не влияют на скорость генерации электронно-дырочных пар.

Итак, в резко неравновесной ситуации, когда процесс рекомбинации электронов и дырок устранен, глубокие центры получают возможность блеснуть в роли «энергетических ступенек», помогающих рождению электронно-дырочных пар.

«Убийцы»

Рассмотрим теперь другую, также неравновесную ситуацию, когда процесс рекомбинации преобладает над процессом генерации. Нам предстоит с удивлением убедиться, что в таких условиях примеси-ступеньки, помогавшие созданию свободных носителей, превратятся в примеси-«убийцы», уничтожающие электроны и дырки. Вообще-то говоря, мы уже несколько подготовлены к такой метаморфозе.

*) Вспомните, какие сильные поля удерживают валентные электроны на их орбитах (см. Введение). Поля, удерживающие носители на глубоких центрах, также очень велики. Это ясно из того, что нужна большая энергия $\Delta\mathcal{E}$, чтобы освободить носитель, захваченный глубоким центром

Обсуждая, каким образом формулы (7), (9) и (10) могут остаться справедливыми в условиях, когда глубокие центры резко увеличивают скорость генерации, мы отмечали, что для этого имеется только одна возможность. Те примеси, которые резко увеличивают скорость генерации электронов и дырок, должны также резко увеличивать и скорость рекомбинации.

Как сделать это свойство глубоких примесей явным? Или, на языке детективных романов, как изобличить убийцу?

Избыточные носители. Фототок. Один из возможных экспериментов-разоблачений показан на рис. 24. Образец, изготовленный из исследуемого полупроводникового материала, подключается последовательно с амперметром к источнику постоянного напряжения (рис. 24, а). В цепи течет ток, определяемый сопротивлением образца. Осветим образец светом, энергия квантов которого $h\nu$ больше, чем энергия создания электронно-дырочной пары в материале $\mathcal{E}_g$. В образце возникнут избыточные носители заряда, электроны и дырки (рис. 24, б). Сопротивление образца уменьшится и амперметр зарегистрирует увеличение тока. Избыточный ток, возникший благодаря дополнительным носителям, созданным светом, *фототок*, будет протекать до тех пор, пока свет будет падать на поверхность образца.

Выключим свет. Ясно, что через некоторое время концентрация электронов и дырок достигнет *равновесного* (темнового) значения. Избыточные электроны и дырки рекомбинируют, «вымрут». Почему?

Понятно, почему, — равновесие между генерацией и рекомбинацией после выключения света будет нарушено. Процессы генерации будут создавать в единице объема в единицу времени определенное число носителей. В равновесии это число в точности равно числу носителей, гибнущих за счет рекомбинации. Но сразу после выключения света в образце существуют *избыточные носители*. Электронов и дырок больше, чем в равновесном состоянии. А ведь скорость рекомбинации пропорциональна числу электронов и дырок в единице объема (формула (3)). Поэтому электроны и дырки будут гибнуть быстрее, чем создаваться: концентрация носителей будет уменьшаться, пока не достигнет равновесного (темнового)

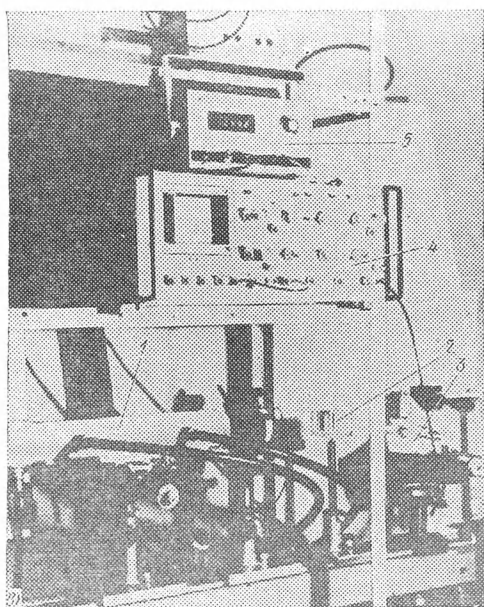
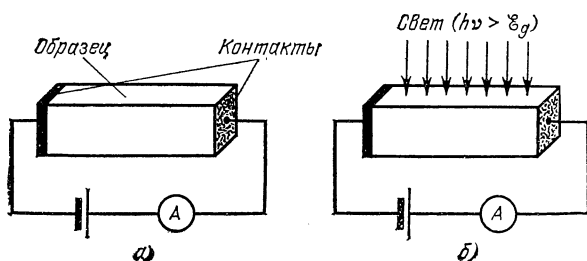


Рис. 24. Измерение времени жизни носителей заряда в полупроводнике.

а), б) Принципиальная схема опыта; в) фотография экспериментальной установки для измерения времени жизни носителей в кремнии (Физико-технический институт им. А. Ф. Иоффе АН СССР): 1 — импульсный твердотельный лазер; 2 — набор фильтров, регулирующих интенсивность излучения; 3 — держатель с образцом; 4 — скоростной осциллограф, регистрирующий спад фототока; 5 — цифровой вольтметр.

значения, при котором скорости процессов генерации и рекомбинации сравниваются.

За уменьшением концентрации неравновесных носителей можно следить, наблюдая спад фототока в цепи. Время возвращения тока к темновому значению определяется *временем жизни* избыточных носителей заряда. Чем скорее ток возвращается к темновому значению, чем меньше время жизни, тем быстрее, следовательно, идет рекомбинация. Если введение глубоких примесей уменьшает время жизни избыточных, неравновесных носителей, значит, эти примеси ускоряют процесс рекомбинации.

Измеряя скорость уменьшения фототока в чистом полупроводнике и в полупроводнике, содержащем примеси, можно изобличить «убийц».

Измерение времени жизни неравновесных носителей. Перед тем как обсудить «следственный эксперимент по изобличению убийц» подробнее, выясним несколько важных технических деталей.

Характерные времена жизни избыточных носителей τ_p (для дырок) и τ_n (для электронов) в различных полупроводниках меняются в очень широких пределах. В зависимости от типа полупроводника, концентрации и природы глубоких примесей, равновесной концентрации носителей, температуры и ряда других факторов, значения τ_n , τ_p лежат в пределах от $\sim 10^{-10}$ до 10^{-2} с. Ясно, что никакой стрелочный прибор, никакой амперметр не успеет зарегистрировать спад фототока. Для регистрации таких быстропротекающих процессов используются электронные осциллографы. Ясно также, что выключать свет следует очень резко. Если в процессе выключения интенсивность света будет спадать за времена, сравнимые с τ_p и τ_n , или медленнее, мы не сможем узнать, наблюдаем ли мы «вымирание» избыточных носителей, или просто прослеживаем, как меняется фототок по мере ослабления света.

Если значения τ_n , τ_p больше, чем микросекунда ($1 \text{ мкс} = 10^{-6} \text{ с}$), то для регистрации таких процессов можно использовать широко распространенные и доступные осциллографы. А для резкого выключения света часто используют быстро вращающийся диск с узкой прорезью или отверстием на периферии диска.

При τ_n, τ_p порядка наносекунды ($1 \text{ нс} = 10^{-9} \text{ с}$) для регистрации спада фототока используются современные скоростные осциллографы. Для формирования импульсов света используют специальные электрооптические модуляторы. Иногда короткие импульсы света получают, используя лазеры.

При $\tau_n, \tau_p \sim 10^{-10} - 10^{-11} \text{ с}$ измерение времени жизни по спаду фототока находится на грани современных технических возможностей. Только в нескольких лабораториях мира имеются соответствующие экспериментальные установки.

Теперь, представляя, какая сложная техника (рис. 24, в) может в действительности скрываться за амперметром, изображенным на рис. 24, а, вернемся к физике исследуемого процесса.

Изготовим сначала образец из возможно более чистого полупроводникового материала и измерим в нем время жизни носителей τ . Если для исследования мы выбрали Si или Ge, то измеренные значения τ могут быть очень велики: $\geq 2 \cdot 10^{-3} \text{ с}$. Затем введем в образец небольшое, строго контролируемое количество примеси и вновь измерим время жизни. Продолжая такие опыты с различными примесями, можно изучить, как влияют на скорость рекомбинации те или иные примесные центры.

Введение мелких примесных центров будет увеличивать равновесную концентрацию носителей в образце, но практически никак не повлияет на их время жизни.

А вот введение глубоких примесей, энергия активации которых $\Delta \mathcal{E}$ близка к $\mathcal{E}_g/2$, немедленно изобличит в них «убийц». Эти примеси, даже в ничтожной концентрации, могут уменьшить время жизни в сотни и тысячи раз. Так, введение золота или платины в кремний в концентрации $\sim 10^{13} \text{ см}^{-3}$ ускоряет гибель неравновесных носителей в тысячу раз. Время жизни τ уменьшается с $\sim 10^{-3}$ до $\sim 10^{-6} \text{ с}$.

Наши «принцы на горошине» верны себе и в очередной раз обнаруживают чувствительность, оставляющую далеко позади способность андерсеновской принцессы: скорость рекомбинации тысячекратно возросла при введении $0,00\ 000\ 001 \%$ примеси!

Итак, убийцы изобличены. Более того, убийцами оказались именно те, на кого с самого начала пало подозрение. И все-таки хороший следователь не соч-

тет дело законченным. Необходимо ответить на вопрос «почему». Почему именно примеси с большой энергией активации служат *активными центрами рекомбинации*?

Герои гибнут в пучине Мальстрема. Представим себе электрон и дырку, блуждающие по кристаллу.

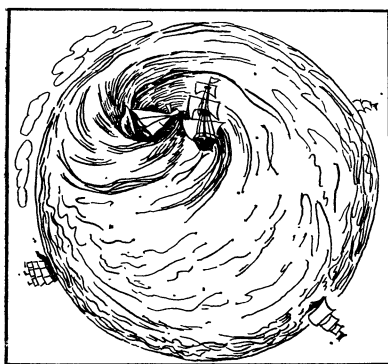
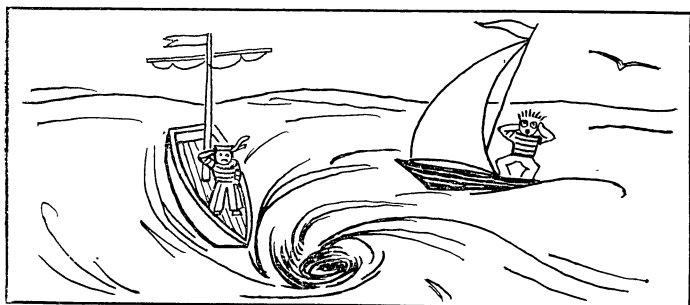


Рис. 25. Вероятность столкновения кораблей в водовороте гораздо больше, чем в открытом море.

Чтобы встретиться, рекомбинировать и исчезнуть, им необходимо оказаться совсем рядом, в окрестности одного и того же атома кристаллической решетки.

Мы можем уподобить процесс рекомбинации электронов и дырок крушению двух кораблей при столкновении их в открытом море. Такой процесс возможен, но, к счастью, маловероятен. Время жиз

ни кораблей (электронов и дырок) велико. Скорость рекомбинации мала.

Теперь представим себе, что в море возник чудовищный водоворот, который втягивает в воронку все корабли, оказавшиеся в его пределах. (Тем, чья фантазия отказывает представить себе такое зрелище, следует перечитать рассказ Эдгара По «Низвержение в Мальстрем».) Ясно, что в таких условиях вероятность встречи кораблей значительно повышается (рис. 25). Глубокие уровни и служат такими «водоворотами».

Пусть в кристалле имеется глубокий примесный центр с большой энергией активации $\Delta\mathcal{E}$. Рассмотрим для определенности донорный центр, который отдал свой лишний электрон и теперь положительно заряжен. Такой положительно заряженный центр служит источником сильного притяжения для свободного электрона, появившегося поблизости. Когда в окрестности ионизованного донора появится электрон, он будет «втянут в водоворот» — захвачен на примесный центр. Если бы это был центр «мелкий» ($\Delta\mathcal{E} \leq kT$), тепловое движение почти немедленно выбросило бы захваченный электрон. А вот глубокий центр, *центр рекомбинации*, будет удерживать захваченный электрон до тех пор, пока в его окрестности не появится дырка. Как только это произошло — электрон и дырка рекомбинируют. Убийца сделал свое дело, истребив электронно-дырочную пару и ... готов снова приняться за свое.

Вот теперь следствие окончено*) и дело можно передавать в суд. Вернее — на суд физиков и технологов. Если необходимо, чтобы электроны и дырки в приборе «жили» долго, — суд приговорит полупроводник к возможно более тщательной очистке от примесей. Иногда, напротив, электроны и дырки должны гибнуть по возможности быстрее. Такое требование часто возникает при создании быстродействующих полупроводниковых приборов. Тогда в материал специально вводят примеси, создающие эффективные центры рекомбинации.

*) Обдумайте самостоятельно роль глубокого акцептора в качестве центра рекомбинации.

СКВОЗЬ ИГОЛЬНОЕ УШКО

Все начали бегать, когда кому захотелось, и бежали, кто куда хотел, и останавливались, когда кто пожелает.

Л. Кэрролл. «Алиса в стране чудес»

В двух предшествующих главах мы узнали, как рождаются и гибнут свободные носители заряда: электроны и дырки. Познакомимся теперь с тем, как проходит жизнь наших героев.

Атомы любого вещества находятся в непрерывном тепловом движении. Очень частые, происходящие сотни миллиардов и триллионы раз в секунду, столкновения свободных носителей с атомами кристаллической решетки заставляют электроны и дырки проводить жизнь в непрерывном хаотическом *тепловом движении*. Если к кристаллу приложено электрическое поле, носители заряда, не прекращая хаотического теплового движения, начинают направленно перемещаться под действием этого поля. Отрицательно заряженные электроны дрейфуют к положительному электроду; положительно заряженные дырки — к отрицательному. *Направленное движение свободных носителей под действием электрического поля* представляет собой электрический ток. Электрический ток, как мы узнаем, может возникнуть, даже если к кристаллу и не приложено внешнее электрическое поле. Когда в какой-либо области кристалла свободных носителей (например, электронов) больше, чем в соседних областях, носители под действием хаотического теплового движения начинают направленно перемещаться из области с более высокой в область с более низкой концентрацией. Такое явление называется *диффузией*, а электрический ток, возникший под влиянием диффузии — *диффузионным током*.

В этой главе мы рассмотрим основные типы движения свободных носителей: тепловое движение, движение в электрическом поле и диффузию.

Тепловое движение

Энергия теплового движения частицы равняется, как известно, $\frac{3}{2} kT$. Приравнявая эту энергию кинетической энергии частицы $m\bar{v}_T^2/2$, полу-

чим, что средняя скорость хаотического движения равняется:

$$v_T = (3kT/m)^{1/2}. \quad (11)$$

Применяя эту формулу для электрона, масса которого $m = 9,1 \cdot 10^{-31}$ кг, получим при $T = 300$ К для тепловой скорости электрона значение $\sim 10^5$ м/с. Сто километров в секунду — на порядок больше второй космической скорости, приобретя которую, тело покидает Солнечную систему! В «микроскопических» масштабах, однако, такая скорость ничего особенного из себя не представляет. В ускорителях электронов сообщают скорость, в 3000 раз большую (практически, равную скорости света). А скорости $v_T \sim 10^5$ м/с соответствует кинетическая энергия электрона всего в несколько сотых долей электронвольта, недостаточная даже для того, чтобы электрон мог покинуть кристалл полупроводника и вылететь наружу. (Для такого «подвига» электрон, как правило, должен обладать энергией в несколько десятых долей электронвольта или даже в несколько электронвольт.)

Тем не менее взирать на значение $v_T \sim 10^5$ м/с свысока никаких оснований нет. Чуть дальше мы убедимся, что, как правило, нужны очень сильные электрические поля, чтобы скорость направленного вдоль поля (а не хаотического) движения электронов или дырок могла достичь значений $\sim 10^5$ м/с.

Движение в электрическом поле

Представим себе, что к полупроводнику приложено электрическое поле E . В электрическом поле E на электрон (или дырку) действует сила $F = eE$. Под действием этой силы носитель движется вдоль поля с ускорением $a = eE/m$. Если носитель движется без столкновений, то через время t его скорость в направлении поля $v = at = eEt/m$.

Однако в кристалле свободные электроны и дырки испытывают столкновения очень часто.

Непосредственно после столкновения носитель может двигаться с равной вероятностью в любом направлении. Это означает, что скорость направленного движения после столкновения равняется нулю (рис. 26).

Пусть среднее время между двумя последовательными столкновениями равняется τ_0^*). (В действительности время между столкновениями может быть как меньше, так и больше τ_0 (рис. 26).) Расчет, проделанный с учетом того, что время между двумя

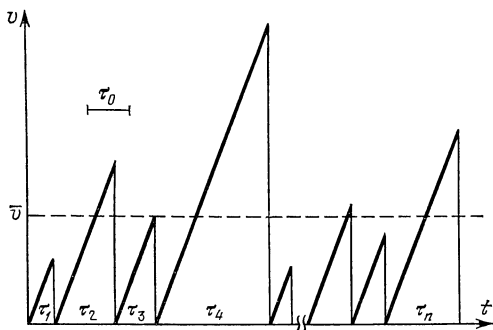


Рис. 26. Направленное движение свободных носителей в электрическом поле с учетом столкновений. Среднее время между столкновениями τ_0 ; средняя скорость направленного движения $\bar{v}$.

последовательными столкновениями — величина случайная, показывает, что *средняя дрейфовая скорость* носителей вдоль электрического поля $\bar{v}$ равняется произведению ускорения a на среднее время между двумя столкновениями τ_0^{**}):

$$\bar{v} = e\tau_0 E/m. \quad (12)$$

Средняя дрейфовая скорость электронов $\bar{v}$ пропорциональна напряженности электрического поля E . Коэффициент пропорциональности

$$\mu = e\tau_0/m \quad (13)$$

носит название *подвижности*.

Подвижность. Из формулы (12) видно, что единицы, в которых измеряется подвижность $[\mu] = \text{м}^2/(\text{В} \cdot \text{с})$.

Величина подвижности свободных носителей, электронов и дырок, является очень важной характери-

*) Мы обозначили среднее время между столкновениями τ_0 , чтобы не путать это обозначение с введенным раньше временем жизни носителей τ .

**) Подробнее о выводе формулы (12) можно прочитать в «Фейнмановских лекциях по физике». — М.: Мир, 1965, вып. 4, с. 82—90.

стикой полупроводника. Значения подвижности во многом определяют пригодность материала для изготовления полупроводниковых приборов. Кроме того, величина подвижности, как мы сейчас убедимся, позволяет судить о степени чистоты полупроводникового кристалла, о совершенстве его структуры и о характере взаимодействия свободных носителей с кристаллической решеткой.

Проанализируем формулу (13), определяющую подвижность. Прежде всего зададим вопрос, что такое τ_0 . Время движения носителя между какими двумя столкновениями определяет этот параметр? Естественно предположить, что τ_0 — это время движения носителя между двумя соседними атомами решетки. Тогда мы очень легко можем вычислить значение τ_0 , а следовательно, и величину подвижности μ .

Вспомним, что двигаясь под действием электрического поля, электрон (или дырка) отнюдь не прекращает хаотического, теплового движения. Скорость теплового движения носителей v_T , как мы видели, по порядку величины равняется 10^5 м/с. Скорость направленного движения под действием электрического поля, как правило, гораздо меньше. Поэтому время, за которое носитель пролетает между двумя соседними атомами, определяется скоростью v_T .

Расстояние между соседними атомами решетки a_0 приблизительно одинаково во всех твердых телах и равняется, как мы знаем, $\sim 5 \cdot 10^{-10}$ м. Двигаясь с тепловой скоростью $v_T \sim 10^5$ м/с, электрон в среднем преодолевает расстояние между атомами за время $\tau_0 \sim 5 \cdot 10^{-15}$ с. Такой величине τ_0 отвечает, в соответствии с формулой (13), значение подвижности $\mu \sim 10^{-3} \text{ м}^2/(\text{В} \cdot \text{с})$.

Поскольку величины a_0 и v_T приблизительно одинаковы для всех кристаллов, подвижность μ , в соответствии с нашим предположением о характере столкновений, также должна быть приблизительно одинакова для всех твердых тел.

Наша гипотеза дает возможность предсказать и температурную зависимость подвижности. Действительно, расстояние между атомами решетки a_0 меняется с температурой слабо. Скорость теплового движения v_T пропорциональна, как следует из формулы (11), корню квадратному из абсолютной температуры T :

$v_T \sim T^{1/2}$. Среднее время между столкновениями $\tau_0 = a_0/v_T \sim T^{-1/2}$. Поэтому можно ожидать, что подвижность, пропорциональная в соответствии с формулой (13), величине τ_0 , будет уменьшаться с ростом температуры: $\mu \sim T^{-1/2}$.

А теперь сравним наши теоретические предсказания с экспериментальными результатами, приведенными в табл. III.

Т а б л и ц а III

Полупроводники		InSb	Ge	Si	InP	GaAs	GaP	SiC
Подвижность электронов, м ² /(В · с)	300 К 77 К	8 80	0,39 4	0,13 2,5	0,55 5	1 30	0,05 0,2	0,02—0,1 0,05—0,3
Подвижность дырок, м ² /(В · с)	300 К 77 К	0,07 0,8	0,19 4	0,05 0,45	0,07 0,55	0,04 0,35	0,01 0,02	2 · 10 ⁻⁴ — —4 · 10 ⁻³ ~10 ⁻³

М-да, сравнение обескураживающее. Значения подвижности у различных полупроводников отличаются в тысячи и десятки тысяч раз. Величины подвижности у некоторых полупроводников превосходят предсказанное нами значение μ на 4—5 порядков. Температурная зависимость подвижности... На первый взгляд кажется, что температурную зависимость, хотя бы качественно, мы предсказали правильно. При увеличении температуры от 77 до 300 К подвижность и электронов, и дырок уменьшается; у некоторых полупроводников отношение подвижностей при 77 и при 300 К действительно близко к 2. (Температура возросла в ~ 4 раза — подвижность уменьшилась вдвое.) Увы, однако и здесь наша совесть нечиста. Подвижность электронов в Si при 77 К превосходит ее значение при комнатной температуре в 20 раз, в GaAs — в 30, в InP и Ge — в 10 раз. Кроме того, если посмотреть, как ведет себя подвижность при еще более низких температурах, то последние иллюзии рухнут. На рис. 27 показаны температурные зависимости подвижности, измеренные в GaAs в диапазоне температур от 4 до 300 К. Видно, что рост подвижности с понижением температуры от комнатной до ~ 80 К сменяется за-

тем падением подвижности с дальнейшим уменьшением температуры.

Одним словом, мы оказались плохими теоретиками.

Что ж, попробуем подойти к делу с другой стороны. Исходя из экспериментально наблюдаемых значений подвижности, рассчитаем, какое расстояние должен пройти электрон, прежде чем он испытает столкновение. Может быть тогда мы лучше поймем, с чем именно он сталкивается?

Возьмем в качестве примера GaAs при 77 К. Величине $\mu = 30 \text{ м}^2/(\text{В} \cdot \text{с})$ соответствует, исходя из формулы (13), значение $\tau_0 \sim 2 \cdot 10^{-10} \text{ с}$. В соответствии с формулой (11), при $T = 77 \text{ К}$ $v_T \approx 5 \cdot 10^4 \text{ м/с}$. Таким образом, между двумя столкновениями электрон проходит путь $l \sim v_T \tau_0 \sim 10^{-5} \text{ м}$. При расстоянии между атомами порядка $5 \cdot 10^{-10} \text{ м}$ это соответствует 20 000 атомным расстояниям!

Поистине, это кажется невероятным. Прежде чем испытать столкновение (с чем?), электрон минует, не замечая, 20 000 атомов. Древняя мудрость гласит, что легче верблюду пройти сквозь игольное ушко, чем богатому попасть в царствие небесное. Но тут поневоле кажется, что все-таки и богачу легче попасть в рай, чем электрону пролететь мимо десятков тысяч атомов и не столкнуться ни с одним из них.

И тем не менее, именно так и обстоит дело. Мы можем доказать это прямым экспериментом. Величина $\mu \sim 30 \text{ м}^2/(\text{В} \cdot \text{с})$ характерна при 77 К для очень чистого GaAs. Из рис. 27 видно, что наличие примеси, концентрация которой составляет $\sim 10^{14} \text{ см}^{-3}$ (один примесный атом на миллиард атомов решетки!), снижает величину подвижности

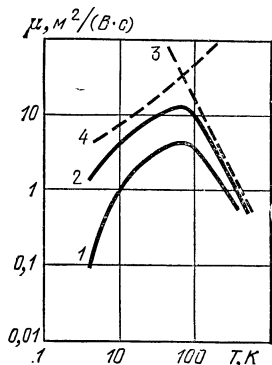


Рис. 27. Зависимость подвижности от температуры в арсениде галлия, содержащем различные концентрации примесных атомов.

1 — концентрация мелких доноров $N_d = 2,5 \cdot 10^{14} \text{ см}^{-3}$; 2 — $N_d = 6,5 \cdot 10^{13} \text{ см}^{-3}$. Пунктирными линиями показаны теоретические зависимости при рассеянии только на колебаниях решетки (кривая 3) или только на примесных центрах (кривая 4).

приблизительно в 3 раза. Введение 10^{17} атомов примеси в 1 см^3 GaAs уменьшает μ до $\sim 1\text{ м}^2/(\text{В}\cdot\text{с})$. Совершенно ясно, что в таких условиях подвижность определяется рассеянием на примесях. Электрон чувствует только атомы примеси и не замечает атомов решетки полупроводника, хотя их в миллион раз больше.

Как можно понять эти чудеса? Никак, если рассматривать электрон и дырку как частицы,— шарики, которые должны пробиться через частокол из десятков тысяч атомов и не задеть ни один из них.

Разгадка всех этих и многих других чудес, связанных с подвижностью, содержится в названии замечательной статьи Д. А. Франк-Каменецкого «Электрическое сопротивление — квантовое явление»*), которую мы очень рекомендуем прочитать.

Именно квантовой природой микрочастиц, в том числе электронов и дырок, объясняются те кажущиеся противоречия, с которыми мы сталкиваемся, анализируя понятие подвижности.

В соответствии с основными представлениями квантовой механики электрон (и дырка), как и каждая частица, обладают не только корпускулярными, но и волновыми свойствами. Применительно к свету такая двойственность, или, как пышно выражаются физики, *корпускулярно-волновой дуализм*, уже не кажется удивительной. К ней привыкли. Анализируя явления интерференции и дифракции, мы подчеркиваем волновую природу света. Обсуждая явление фотоэффекта, представляем световой поток в виде потока частиц — фотонов. Представление электрона в виде волны менее привычно. Однако то обстоятельство, что электрон обладает свойствами не только частицы, но и волны, доказано не менее убедительно и наглядно, чем для света. Еще на заре развития квантовой механики замечательные опыты Дэвиссона и Джермера наглядно продемонстрировали явление дифракции электронов.

То, что кажется немислимым для электрона — частицы, представляется вполне естественным, если не забывать о волновой природе электрона.

Любая волна (звуковая, световая, радиоволна) может распространяться, не отражаясь, в среде, со-

*) См.: Квант, 1970, № 9.

держашей какие-либо рассеивающие центры. При этом, конечно, необходимо выполнение ряда условий, и главнейшее из них состоит в том, что эти центры должны быть расположены в виде идеальной периодической решетки. Малейшее отклонение от идеальной периодичности: изменение расстояния между центрами, замена одного центра другим, хотя бы немного отличающимся по своим свойствам, даже отсутствие центра там, где он должен был бы находиться в решетке, вызывает рассеяние волны.

Точно так же и электрон, двигаясь в идеальном кристалле, не содержащем посторонних атомов, вакансий (пустых мест в решетке), искажений, при температуре абсолютного нуля не испытывал бы рассеивающих столкновений и двигался свободно.

Механизмы рассеяния. Если кристалл содержит примеси, хотя бы и в очень незначительной концентрации, электронная волна будет рассеиваться на них и среднее время между двумя актами рассеяния, между двумя столкновениями, τ_0 , будет определяться концентрацией примеси (так называемое *примесное рассеяние*). Так же будет обстоять дело, если в кристалле содержатся какие-либо структурные несовершенства, пустые места в узлах кристаллической решетки и т. д. Если температура кристалла отлична от нуля, атомы, составляющие решетку кристалла, хаотически колеблются и в любой момент времени некоторые атомы будут находиться ближе друг к другу, чем это было в идеальной решетке при $T = 0$ К, а некоторые — дальше друг от друга. Идеальная периодичность нарушится. Возникнет рассеяние на колеблющихся атомах решетки (так называемое *рассеяние на колебаниях решетки*). Это рассеяние будет, естественно, тем сильнее, чем выше температура, т. е. чем больше будет амплитуда колебаний.

Ход температурной зависимости подвижности $\mu(T)$, показанный на рис. 27, типичен для многих полупроводников. При высоких температурах величина подвижности практически не зависит от концентрации примесей. Она определяется рассеянием на колебаниях кристаллической решетки. Если бы в кристалле отсутствовали примеси и несовершенства, с понижением температуры подвижность носителей увеличивалась бы безгранично. Но кристалл,

совершенно свободный от примесей и дефектов, получить невозможно. Из рис. 27 ясно видно, что чем меньше примеси содержит кристалл, тем большие значения подвижности могут быть получены. Однако с дальнейшим понижением температуры в конце концов начинает преобладать рассеяние на примесях и подвижность уменьшается.

Кстати, почему подвижность падает с ростом температуры при рассеянии на колебаниях решетки, мы понимаем. Из рис. 27 видно, что при рассеянии на примесях подвижность, напротив, растет с увеличением температуры. Чем это объясняется?

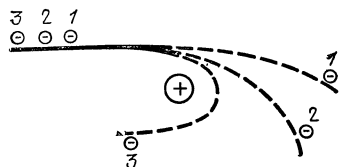


Рис. 28. Рассеяние носителя на заряженном примесном центре. Электроны (1, 2 и 3) подлетают к примесному центру по одной и той же траектории, но с разными скоростями. Электрон 1 имеет самую большую скорость; притяжение примесного атома почти не искажает его первоначальной траектории. Самую маленькую скорость имеет электрон 3; притяжение заряженной примеси заставляет его повернуть обратно.

На рис. 28 схематически показано, как происходит рассеяние электрона на заряженном атоме примеси. Чем больше скорость электрона, пролетающего мимо атома примеси, тем меньше искажается его траектория под действием электрического поля ато-

ма. Средняя скорость теплового движения электронов растет с увеличением температуры (см. формулу (11)). Поэтому рассеяние электронов на примесях с ростом температуры уменьшается, и, как следствие, возрастает подвижность.

На рис. 28 показан один частный случай примесного рассеяния. Было бы неплохо самостоятельно разобрать три других случая: рассеяние электрона на отрицательно заряженном примесном центре и рассеяние положительной дырки на положительно и отрицательно заряженных ионах примеси. Кроме того, полезно проанализировать аналогию между примесным рассеянием свободных носителей в полупроводниках и опытом Резерфорда по рассеянию α -частиц атомами тяжелых элементов. Эти опыты описаны в школьном учебнике физики (10 класс).

А теперь снова обратимся к формуле (13). В правую часть ее, определяющую значение подвижности, входят всего три величины. Однако мы уже убеди-

лись на примере одной из них — τ_0 , как обманчива простота этой формулы.

Эффективная масса. В начале этого раздела, когда мы наивно полагали, что в промежутке между двумя столкновениями электрон летит в пустом пространстве между атомами, мы имели все основания считать величину m массой свободного электрона. Теперь мы, однако, уже знаем, что дело обстоит совсем не так просто. В промежутке между двумя актами рассеяния электрон и дырка могут проходить расстояния, измеряемые тысячами и десятками тысяч постоянных решетки. Это означает, что во время свободного движения электрона и дырки на них, помимо внешнего электрического поля E , воздействуют сильнейшие поля, создаваемые ионами кристаллической решетки и валентными электронами атомов, между которыми пролетают свободные носители (вспомните Введение!). Невозможно ожидать, чтобы это обстоятельство не отразилось на движении электронов и дырок. Правомерно даже поставить вопрос: имеют ли вообще смысл те рассуждения, которые привели к выводу формул (12) и (13) и, следовательно, можно ли этими формулами пользоваться?

Ответ на этот вопрос может дать только квантовая механика, учитывающая одновременно и волновые и корпускулярные свойства электронов и дырок. Ответ этот таков: благодаря тому, что электрические поля атомов кристаллической решетки расположены в пространстве строго периодически (и только поэтому), формула (12) может быть использована для описания движения электронов и дырок в кристалле полупроводника. Однако под массой m , стоящей в знаменателе выражений (12) и (13), следует понимать не массу свободного электрона в вакууме m_0 , а величину, разную для различных полупроводников, и называемую «эффективной массой» электрона или дырки в кристалле. Для обозначения эффективной массы используют обычно символ m^* . Замена массы свободного электрона m_0 на эффективную массу m^* как раз и отражает влияние строго периодического поля кристаллической решетки на движение электрона.

Поскольку параметры атомов и структура решетки различны в различных полупроводниках, следует ожидать, что эффективные массы электронов

и дырок для разных полупроводников будут отличаться друг от друга. Именно так дело и обстоит.

Значения эффективных масс электронов m_e^* и дырок m_h^* для некоторых полупроводников приведены в табл. IV *). Обратите внимание, что числа, приведенные в таблице, представляют собой отношение эффективной массы электронов (дырок) к массе свободного электрона в вакууме $m_0 = 9,1 \cdot 10^{-31}$ кг.

Из табл. IV видно, что значения m_e^* и m_h^* действительно могут отличаться очень существенно от массы свободного электрона m_0 . В антимониде индия (InSb), например, эффективная масса электрона почти в 80 раз меньше, чем масса свободного электрона! Немудрено, что для электронов в InSb характерны такие высокие значения подвижности (см. табл. III).

Т а б л и ц а IV

Полупроводники	InSb	Ge	Si	InP	GaAs	GaP	SiC
m_e^*/m_0	0,013	0,12	0,26	0,07	0,07	0,35	0,6
m_h^*/m_0	0,18	0,28	0,49	0,2	0,45	0,86	1,2

Поскольку эффективные массы электронов и дырок очень заметно отличаются друг от друга в различных полупроводниках, ясно, что скорости теплового хаотического движения свободных носителей, определяемые формулой (11), также различны в различных материалах при одной и той же температуре T . При $T = 300$ К величина средней тепловой скорости для электронов составляет в InSb $\sim 10^6$ м/с, в GaAs $\sim 4,5 \cdot 10^5$ м/с, а в SiC $\sim 1,5 \cdot 10^5$ м/с. При той же температуре тепловые скорости дырок в этих же материалах составят, соответственно, $2,5 \cdot 10^5$, $1,7 \cdot 10^5$ и 10^5 м/с.

Горячие электроны. Теперь, когда мы уточнили значения скорости неупорядоченного, хаотического

*) Для величин, описывающих параметры дырки, индекс «h» используется очень часто. Он происходит от английского слова hole — дырка.

движения носителей в различных материалах и обстоятельно познакомились с понятием подвижности, рассчитаем, какое электрическое поле необходимо приложить к полупроводнику, чтобы скорость направленного вдоль поля движения носителей сравнялась с тепловой скоростью v_T . Соответствующее значение поля E_0 , равное, очевидно, $E_0 = v_T/\mu$, составит при комнатной температуре $\sim 10^5$ В/м для n -InSb, $\sim 5 \cdot 10^5$ В/м — для n -GaAs и $\sim 10^6$ В/м — для n -SiC. Для полупроводников p -типа (дырочных) соответствующие значения E_0 еще выше, поскольку подвижность дырок, как правило, меньше, чем подвижность электронов (см. табл. III).

Для того чтобы лучше представить себе, что такое напряженность электрического поля $\sim 5 \cdot 10^5$ В/м, заметим, что поле такой величины возникнет в образце длиной в 1 см, если приложить к нему напряжение 5 000 В.

Когда приложенное к полупроводнику поле так велико, что скорость направленного движения носителей по порядку величины приближается к скорости теплового движения, электроны и дырки приобретают необычные свойства. Носители, разогнанные полем до больших скоростей и энергий, называют *горячими* носителями.

О том, что такое горячие электроны, как могут существовать горячие электроны в холодном кристалле, при каких условиях электроны становятся горячими, рассказано в статье В. А. Фабриканта «Закон Джоуля — Ленца» (Квант, 1972, № 10)*). Горячие носители взаимодействуют с решеткой иначе, чем носители в слабом электрическом поле. У горячих электронов и время рассеяния τ_0 , и эффективная масса m^* , а, следовательно, и подвижность зависят от напряженности электрического поля E . При этом возникает множество интересных и важных эффектов. О некоторых из них мы расскажем в третьей части книги.

*) Все оценки в этой статье проделаны, однако, для металлов, где недостижимы сильные электрические поля. Повторите проделанные в статье В. А. Фабриканта расчеты для полей $\sim 10^6$ — 10^7 В/м, которые могут быть получены в полупроводниках.

Проводимость

Продолжим обсуждение электрических свойств полупроводников.

Закон Ома в дифференциальной форме. Установим несколько простых соотношений, вытекающих из закона Ома. Пусть к образцу длиной L , имеющему поперечное сечение S и изготовленному из материала с удельным сопротивлением ρ , приложено напряжение V . Через образец течет ток

$$I = \frac{V}{R} = \frac{VS}{\rho L}. \quad (14)$$

Учтем, что плотность тока $j = I/S$, поле в образце $E = V/L$, и введем величину проводимости $\sigma = 1/\rho$ ($[\sigma] = \text{Ом}^{-1} \cdot \text{м}^{-1}$). Тогда формулу (14) можно переписать в виде:

$$j = \sigma E. \quad (15)$$

Формулу (15) иногда также называют законом Ома, поскольку она, в сущности, полностью эквивалентна обычной формуле закона Ома: $I = V/R$. Однако в отличие от обычной формулировки закона Ома, формула (15) устанавливает связь не между величинами, характеризующими весь образец в целом, — сопротивлением R , током I и напряжением V , а между величинами, определенными в каждой точке образца — плотностью тока j , проводимостью σ и напряженностью поля E . Поэтому формулу (15) называют еще и законом Ома в дифференциальной форме.

Сравним теперь формулу (15) с хорошо известным (Физика, 9 класс) выражением для плотности тока j в виде:

$$j = en_0\vec{v} = en_0\mu_n E. \quad (16)$$

В формуле (16) n_0 — концентрация свободных электронов в полупроводнике n -типа. Для дырочного полупроводника вместо n_0 войдет концентрация дырок p_0 , а вместо подвижности электронов μ_n — дырочная подвижность μ_p .

Из формул (15) и (16) следуют выражения для проводимости полупроводников n - и p -типа:

$$\text{а) } \sigma = en_0\mu_n, \quad \text{б) } \sigma = ep_0\mu_p. \quad (17)$$

Вычислять значения концентрации электронов и дырок мы умеем для самых разных случаев; и для собственного полупроводника *) (формула (6)), и для примесных (формулы (7) и (10)), и для случая примесного истощения. С поведением подвижности мы также хорошо знакомы. Таким образом, мы можем теперь анализировать электропроводность полупроводников в самых разных, даже довольно сложных случаях.

Такие простые зависимости. Вернемся к самому началу книги и рассмотрим рис. 1. Убедимся, что приобретенные нами знания дают возможность исчерпывающе объяснить показанную на рисунке зависимость.

Одной десятиллионной доле процента примеси соответствует концентрация примесных атомов фосфора в кремнии $N_d \approx 10^{14} \text{ см}^{-3} = 10^{20} \text{ м}^{-3}$. Фосфор, как мы знаем, является в кремнии мелким донором с энергией ионизации $\Delta\mathcal{E} \approx 0,05 \text{ эВ}$. Рассматривается область достаточно высоких температур, близких к комнатной, для которой энергия теплового движения $\frac{3}{2}kT \approx 0,04 \text{ эВ}$ близка к энергии активации примеси $\Delta\mathcal{E}$. Для такой ситуации характерно, как мы знаем, явление примесного истощения. Все атомы фосфора ионизованы, и кремний содержит в каждом кубическом сантиметре $\sim 10^{14}$ примесных электронов. Расчет по формуле (6) ясно показывает, что вкладом собственных электронов и дырок в проводимость при данной температуре можно пренебречь.

Подвижность электронов в кремнии μ_n при комнатной температуре мы знаем. Она равняется (см. табл. III) $\sim 0,13 \text{ м}^2/(\text{В} \cdot \text{с})$. Следует ожидать, что проводимость кремния при комнатной температуре составит $\sigma = en_0\mu_n = eN_d\mu_n \approx 2 \text{ Ом}^{-1} \cdot \text{м}^{-1}$. Как видно из рис. 1, именно такое значение проводимости и наблюдается в действительности при 25°C . Совершенно ясно также, почему проводимость уменьшается с ростом температуры: в рассматриваемом диапазоне температур подвижность определяется рассеянием на колебаниях решетки (см. рис. 27) и умень-

*) В случае, когда в проводимость вносят заметный вклад и электроны, и дырки, $\sigma = e(n_0\mu_n + p_0\mu_p)$. Так бывает или в собственных полупроводниках, или если полупроводник освещен светом большой интенсивности.

шается с ростом температуры, а концентрация электронов n_0 от температуры не зависит.

Рассмотрим более сложный случай. На рис. 29 показаны зависимости проводимостей трех образцов арсенида галлия, измеренные в широком диапазоне температур. Можем ли мы объяснить ход этих зависимостей? Безусловно.

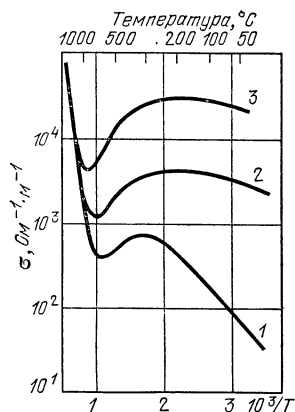


Рис. 29. Температурная зависимость проводимости арсенида галлия для трех образцов с различной концентрацией донорной примеси N_d . Концентрация N_d возрастает с увеличением номера образца ($N_{d3} > N_{d2} > N_{d1}$).

Обсудим сначала ход кривой 1. Начнем с правой части рисунка (области относительно низких температур). Видно, что в области температур от 30 до 200°C проводимость экспоненциально возрастает с ростом температуры. (Аналогично увеличивалась концентрация дырок с возрастанием температуры на рис. 20.) Экспоненциальный характер зависимости $\sigma(T)$ возникает из-за экспоненциального возрастания концентрации электронов с ростом температуры, описываемого формулой (7)*).

На первый взгляд кажется, что ситуация не совсем ясна. Ведь проводимость зависит не только от концентрации носителей, но и от их подвижности. А подвижность тоже изменяется с изменением температуры. Дело, однако, в том, что подвижность зависит от температуры сравнительно слабо. Слабая зависимость подвижности от температуры почти не сказывается на ходе проводимости, когда концентрация очень резко (экспоненциально) изменяется с температурой.

Но вот с возрастанием температуры наступает примесное истощение и концентрация от температуры зависеть перестает (ср. с рис. 20). Теперь ход зависимости $\sigma(T)$ целиком определяется зависи-

*) Определите из рисунка величину энергии ионизации донора $\Delta\mathcal{E}$ для данного случая.

мостью подвижности от температуры $\mu(T)$. Примесь, определяющая концентрацию электронов в GaAs в рассматриваемом случае, оказывается достаточно глубокой. Поэтому явление примесного истощения проявилось при довольно высокой температуре, где подвижность определяется рассеянием носителей на тепловых колебаниях решетки и уменьшается с ростом температуры (ср. с рис. 27).

При еще более высокой температуре снова начинается резкое, экспоненциальное возрастание проводимости с температурой. Это — область собственной проводимости, в которой концентрация электронов и дырок определяется формулой (6).

Образец, для которого ход зависимости $\sigma(T)$ определяется кривой 2, содержит большую концентрацию примесей N_d . Так как энергия ионизации примеси $\Delta\mathcal{E}$ уменьшается с ростом N_d (ср. с рис. 19), то в области примесной проводимости зависимость $\sigma(T)$ более слабая, чем для кривой 1. Кроме того, с уменьшением $\Delta\mathcal{E}$ уменьшается температура, при которой наступает примесное истощение. Видно, что падение подвижности с температурой проявляется на зависимости $\sigma(T)$ для кривой 2 при более низкой температуре, чем для кривой 1. Наконец, в области собственной проводимости все тонкие различия в концентрации примеси и ходе подвижности сменяются стремительным возрастанием с температурой собственной концентрации электронно-дырочных пар.

Измерение проводимости. Знание величины проводимости материала, с которым исследователь собирается работать, так же необходимо для специалиста в области полупроводников, как для инженера-строителя знание предела упругости и прочности материала, из которого должно быть возведено здание. И определить экспериментально проводимость полупроводникового образца, в принципе, совсем несложно. Для этого достаточно измерить сопротивление R образца известной длины L и поперечного сечения S . Затем проводимость σ найдется из очевидного выражения:

$$\sigma = 1/\rho = L/RS. \quad (18)$$

Однако если попробовать проделать такой опыт, то для большинства полупроводниковых материалов

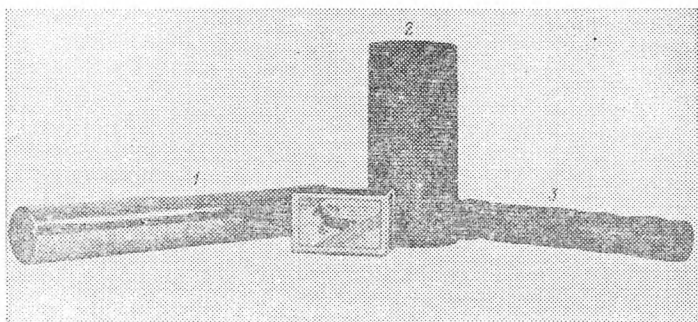


Рис. 30. Такие хорошо изученные и широко применяемые полупроводники, как Ge (1), Si (2), GaAs (3) современная технология позволяет получать в виде монокристаллов объемом в десятки и сотни кубических сантиметров. Из каждого такого монокристалла можно, в принципе изготовить миллиарды высокочастотных диодов и транзисторов, миллионы обычных транзисторов, работающих в бытовых электронных приборах, десятки тысяч мощных диодов и тиристоров, работающих в выпрямительных устройствах электропоездов и мощных линий электропередач.

(Ge, Si, GaAs, GaP и очень многих других) результат окажется неожиданным.

Представьте себе, что в наших руках — массивный монокристалл кремния или арсенида галлия

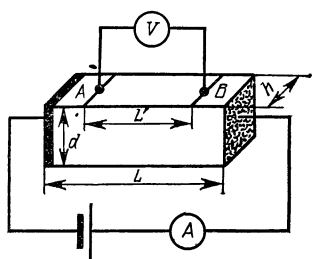


Рис. 31. Измерение проводимости полупроводникового образца.

(рис. 30). Его проводимость известна — кто-то измерил ее и сообщил нам. Вырежем из этого монокристалла образец правильной формы, удобнее всего в форме прямоугольного параллелепипеда (рис. 31)*. Естественно ожидать, что сопротивление такого образца будет равняться $R = L/\sigma h a$. Однако если прикоснуться к боковым граням образца из-

мерительными концами тестера (прибора для измере-

*) Дело это, кстати, не простое. Полупроводники довольно трудно поддаются механической обработке и нам придется воспользоваться абразивным инструментом, режущая кромка которого содержит сверхтвердые материалы корунд, эльбор или алмаз,

ния сопротивлений), то прибор покажет значение в миллионы раз больше, чем ожидаемое. Дело в том, что между металлическими концами тестера и полупроводником возникает узкий приповерхностный слой с очень большим сопротивлением, и прибор будет измерять сопротивление этого слоя, а не объемное сопротивление полупроводника.

Чтобы обойти эту трудность, боковые грани образца необходимо снабдить так называемыми «омическими контактами», т. е. контактами, в которых слой с высоким сопротивлением отсутствует. Как правило, такие контакты получают, вплавляя в боковые грани определенные металлы (индий, олово) или сплавы (германий — золото, германий — серебро). Изготовить хороший омический контакт нелегко. Приходится подбирать материал контакта, температуру и режим сплавления, атмосферу, в которой следует производить сплавление (часто это бывает атмосфера водорода или инертного газа), определенным образом подготавливать поверхность боковых граней. Почти всегда технологам удается преодолеть все трудности, и в конце концов тестер, подключенный к боковым граням образца, показывает величину сопротивления, равную расчетному значению.

Но как быть, если величина σ нам неизвестна и ее то и надлежит определить? Измеряя сопротивление какого-либо образца прямо на боковых гранях, никогда не знаешь, какая доля измеряемого сопротивления приходится на объем образца, а какая — на долю контактных сопротивлений.

Выход показан на рис. 31. Падение напряжения следует измерять не в тех точках, в которых оно приложено к образцу! Ток I , текущий через образец, измеряется амперметром. Этот ток создает падение напряжения V между точками A и B , в которое сопротивление контактов, очевидно, не входит. Падение напряжения измеряется вольтметром, внутреннее сопротивление которого гораздо большее, чем сопротивление между точками A и B . Значение σ рассчитывается затем по формуле $\sigma = IL'/Vhd$.

Измерение проводимости — одна из самых простых экспериментальных задач. И мы так подробно разбирали этот пример только для того, чтобы показать,

как много неожиданностей может встретиться при постановке простейшего, в принципе, эксперимента и о скольких вещах нужно знать и помнить, чтобы получить правильный результат даже в очень простом случае.

Знать проводимость полупроводника физику-экспериментатору необходимо. Но недостаточно. В самом деле, одно и то же значение проводимости может быть обусловлено большой подвижностью и малой концентрацией носителей в полупроводнике и наоборот — большой концентрацией и малой подвижностью (см. формулу (17)). Между тем, в разных эффектах эти два параметра проявляют себя совершенно по-разному. Например, способность полупроводникового диода блокировать приложенное напряжение или чувствительность к свету (фоточувствительность) фотосопротивления определяются, в основном, концентрацией носителей, а быстроедействие многих типов приборов — величиной подвижности. Важно поэтому уметь измерять значения концентрации и подвижности по отдельности.

Способов измерения концентрации n_0 (или p_0) и подвижности существует великое множество. Однако самым распространенным является способ, основанный на эффекте Холла. Этот эффект, открытый американским физиком Холлом в 1879 г., вот уже больше ста лет служит таким же необходимым (и почти таким же надежным) инструментом при исследовании свойств полупроводников, каким мастеров каменщика служит вот уже несколько тысяч лет строителям. Мы подробно познакомимся с этим эффектом в третьей части книги.

Диффузия

Перейдем теперь к изучению третьего типа движения носителей в полупроводнике — диффузии.

Понятие «диффузия» — от латинского *diffusio* (разлитие) — применимо к очень многим явлениям в газах, жидкостях и твердых телах.

Запах одеколона, пролитого в комнате, в конце концов распространится по всей квартире. Даже если отключить батареи, закрыть форточки и законопатить щели, так, чтобы воздух был совершенно неподвижен,

молекулы одеколона попадут в другие комнаты благодаря процессу диффузии.

На поверхность полупроводниковой пластины, тщательно очищенной от примесей, наносится слой вещества, содержащего примесь в большом количестве. Через некоторое время эту примесь можно обнаружить в объеме полупроводниковой пластины, далеко от поверхности полупроводника. Примесь попадет в объем благодаря процессу диффузии *). Время, в течение которого процесс диффузии заметно проявит себя, зависит от характера примеси, от того, в каком полупроводнике осуществляется диффузия, и от температуры. Скорость диффузии примесей в полупроводнике очень сильно растет с увеличением температуры. При комнатной температуре нужны десятилетия, чтобы примесь проникла в объем полупроводника, а при высокой температуре процесс диффузии может занять несколько часов или даже несколько минут.

Дадим воде в банке хорошенько отстояться и осторожно положим на поверхность воды кристалл марганцевокислого калия, именуемого в просторечии марганцовкой. Кристалл тяжелее воды. Он упадет на дно, оставив в толще воды резкий, хорошо различимый розовый след. Оставим банку в покое на некоторое время, а затем посмотрим, что произошло. Узкий розовый след расплзся в толстую трубочку. На дне — розовое облачко вокруг упавшего кристалла. Через несколько дней, даже если очень тщательно защитить банку от толчков и сотрясений, перемешивающих слои жидкости, вся вода в банке приобретет однородную розовую окраску. Молекулы марганцовки окрасили воду, распространившись, благодаря диффузии, по всему объему банки.

Что общего между тремя описанными явлениями? Именно то, что составляет существо процесса диффузии, — спонтанное, самопроизвольное, происходящее без каких-либо внешних воздействий перемещение вещества из места, где его много, туда, где его мало.

Процесс диффузии есть прямое следствие хаотического теплового движения атомов (или молекул). Или, если выразаться еще пышнее, — прямое следствие

*) О диффузии см. книгу: *Бокштейн Б. С.* Атомы блуждают по кристаллу. — М.: Наука, 1983. Библиотечка «Квант», вып. 28.

атомно-молекулярного строения вещества. Понять, почему совершенно неупорядоченное хаотическое движение частиц приводит к направленному перемещению вещества из области высокой концентрации в область концентрации низкой, очень важно. И очень просто.

Рассмотрим пример с пролитым одеколоном. Мысленно выделим вокруг лужицы пролитого одеколона полусферу, через которую молекулы проходят совершенно беспрепятственно. Будем подсчитывать молекулы одеколона, прошедшие через поверхность полусферы изнутри наружу и наоборот. Молекулы одеколона под ударами молекул воздуха беспорядочно мечутся во все стороны, хаотически двигаясь во всех направлениях. Те молекулы одеколона, которые находятся поблизости от мысленно выделенной поверхности, получив толчок в нужном направлении, могут пересечь поверхность. Если бы концентрация молекул одеколона была одинакова по обе стороны мысленной перегородки, то поток молекул, направленных изнутри наружу, был бы в точности равен потоку молекул снаружи внутрь. Действительно, при равной концентрации внутри и снаружи эти состояния (внутри и снаружи) абсолютно ничем не отличаются. Сколько молекул одеколона под влиянием случайных блужданий выйдет изнутри за перегородку, столько же войдет снаружи в выделенный объем.

В разбираемом примере дело обстоит не так. Концентрация молекул одеколона над пролитой лужицей велика и спадает по мере удаления от нее. Снаружи от выделенной поверхности концентрация одеколона ниже, чем внутри. Поэтому больше молекул одеколона пересечет поверхность изнутри, чем снаружи. Под влиянием совершенно хаотических блужданий возникает направленный изнутри наружу поток молекул одеколона. В этом и состоит процесс диффузии, всегда приводящий к выравниванию концентраций.

Из рассмотренной картины ясно, что поток, вызванный диффузией, будет тем больше, чем резче меняется концентрация вблизи от мысленно выделенной поверхности. Именно от того, насколько резко спадает концентрация в рассматриваемой области, и зависит, насколько поток молекул одеколона изнутри наружу превысит встречный поток — снаружи внутрь. Скорость изменения концентрации n с коор-

динатой x , естественно, характеризуется производной от концентрации по координате dn/dx . (Для величины dn/dx существует специальное название — градиент концентрации.) Итак, следует ожидать, что диффузионный поток носителей J_D будет пропорционален градиенту концентрации:

$$J_D = -D \, dn/dx. \quad (19)$$

Почему перед правой частью уравнения (19) появился знак минус, помогает понять рис. 32. Поток носителей всегда направлен навстречу градиенту концентрации и стремится, выравнивая концентрацию, уменьшить градиент.

Коэффициент пропорциональности D носит название *коэффициента диффузии*.

Коэффициент диффузии. Установим, прежде всего, размерность коэффициента диффузии. Поток носителей J_D численно равен числу носителей, пересекающих в единицу времени единичную площадь. Таким образом, размерность величины $[J_D] = \text{м}^{-2} \cdot \text{с}^{-1}$. Размерность концентрации $[n] = \text{м}^{-3}$, а градиента концентрации, очевидно, $[dn/dx] = \text{м}^{-4}$. Отсюда ясно, что коэффициент диффузии имеет размерность $[D] = \text{м}^2/\text{с}$.

От чего зависит коэффициент диффузии? Какими свойствами молекул он определяется?

Мы могли бы вывести точное выражение для коэффициента диффузии. Идея такого вывода очень напоминает способ, которым в курсе физики (9 класс) устанавливается связь между давлением газа p и скоростью, концентрацией и массой молекул газа (основное уравнение молекулярно-кинетической теории). Центральное место в выводе этого уравнения занимает подсчет числа молекул, ударяющихся в единицу времени о заданную площадку (поршень). В задаче о диффузии ситуация очень похожа: необходимо найти число молекул, которые пересекают в единицу време-

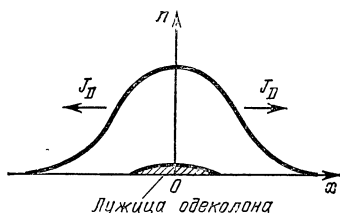


Рис. 32. Распределение молекул одеколона над пролитой лужицей.

В области $x > 0$ градиент dn/dx отрицателен (n уменьшается с ростом x). Диффузионный поток носителей J_D положителен: направлен вдоль положительного направления оси x . В области $x < 0$ градиент положителен; диффузионный поток — отрицателен.

ни выделенную мысленно площадку. Существенная разница состоит, однако, в том, что в отличие от задачи с поршнем, молекулы в задаче о диффузии будут пересекать площадку и слева направо, и справа налево. Результирующий поток J_D будет равен разности этих двух потоков.

Попробуйте вывести сами выражение для коэффициента диффузии, перечитав внимательно гл. 14 учебника физики за 9 класс *).

Здесь мы определим коэффициент диффузии D , воспользовавшись так называемыми соображениями размерности. Определение какой-либо величины из размерных соображений при первом знакомстве производит впечатление ловкого фокуса. В действительности же это очень изящный и широко используемый в современной физике метод. О нем рассказывается в статье Ю. Брука и А. Стаченко «Метод размерностей помогает решать задачи» (Квант, 1981, № 6).

Итак, вернемся вновь к вопросу, который мы поставили в начале раздела. От чего зависит коэффициент диффузии?

Естественно ожидать, что коэффициент диффузии будет зависеть от длины свободного пробега молекул l .

Действительно, чем больше l , тем реже сбивают молекулу с «истинного пути» удары других молекул. В рассмотренном нами примере с пролитым одеколоном ясно, что если бы удары молекул воздуха не сообщали движению молекул одеколона характер случайных, хаотических блужданий, то, стартовав с поверхности пролитой лужицы со скоростью v_T , молекула одеколона достигла бы границы комнаты размером L через время $t \sim L/v_T$. В действительности же она окажется там через гораздо более длительный промежуток времени. Чем чаще молекулу одеколона будут сбивать с пути удары молекул воздуха, т. е. чем меньше длина свободного пробега l , тем медленнее будет диффундировать молекула одеколона. Тем меньше будет коэффициент диффузии D .

*) Если правильный ответ не получится, не огорчайтесь, задача не так уж проста. Подробный вывод выражения для коэффициента диффузии содержится во всех курсах физики для высших учебных заведений (см., например: *Савельев И. В.* Курс общей физики. — М.: Наука, 1981, т. I, с. 408).

Естественно также ожидать, что коэффициент диффузии будет увеличиваться с ростом тепловой скорости молекул v_T .

От чего еще может зависеть коэффициент диффузии? От времени между столкновениями τ_0 ! Правильно, но сама величина τ_0 выражается через l и v_T : $\tau_0 = l/v_T$.

От чего еще может зависеть величина D ? Да пожалуй, больше ни от чего. Может быть, от массы молекул? Вряд ли. Нас ведь интересует не энергия, переносимая молекулами, и не импульс, а просто число молекул.

Итак, предположим, что мы, хорошо представляя себе физическую картину диффузии, правильно определили переменные, от которых зависит коэффициент диффузии D . Тогда осталась главная в методе размерностей, но в нашем случае очень простая задача: так скомбинировать переменные (в нашем случае l и v_T), чтобы комбинация имела размерность определяемой величины. Когда имеешь две величины: одну с размерностью метр на секунду (v_T), а другую с размерностью метр (l), то нетрудно сообразить, что нужно сделать, чтобы получить из них величину D с размерностью метр в квадрате на секунду. Перемножить:

$$D \sim lv_T. \quad (20)$$

Подробный расчет, который, как мы надеемся, вы попытаетесь проделать сами, дает возможность установить численный коэффициент:

$$D = \frac{1}{3}lv_T. \quad (21)$$

Диффузионный ток. Оставим на время молекулы одеколona и обратимся к нашим старым знакомым — свободным носителям заряда в полупроводниках.

Представим себе, что в какой-то части полупроводника (для определенности — электронного) концентрация носителей больше, чем в соседних областях. Такая ситуация может возникнуть, например, если эта часть полупроводника нагрета или освещена. Благодаря диффузии, как мы знаем, возникнет поток электронов из области с более высокой в область с более низкой концентрацией. Но направленный поток заряженных носителей, по определению, пред-

ставляет собою электрический ток! Плотность такого *диффузионного тока* мы можем легко определить, зная распределение носителей

$$j_D = eJ_D = -eD \, dn/dx. \quad (22)$$

Плотность диффузионного тока равняется плотности потока носителей, определяемой уравнением (19), умноженной на заряд электрона (или дырки).

При вычислении диффузионного тока следует помнить о знаке заряда носителей, поведение которых изучается. Если их заряд отрицателен (электроны), то хотя поток электронов будет направлен навстречу градиенту носителей, диффузионный ток электронов будет направлен в противоположную сторону, в сторону возрастания концентрации $j_{eD} = |e|D_e dn/dx$. В этой формуле D_e — коэффициент диффузии электронов.

Для дырок (положительный заряд) диффузионный ток направлен в ту же сторону, что и поток дырок, — в сторону уменьшения концентрации: $j_{pD} = -|e|D_p dp/dx$, D_p — коэффициент диффузии дырок.

Диффузионный ток — это самый настоящий ток, ничем не уступающий хорошо нам знакомому току, возникающему под действием электрического поля E и определяемому (для электронов) формулой (16). При прохождении диффузионного тока выделяется джоулево тепло. Так же как и ток, возникающий под действием электрического поля, диффузионный ток вызывает отклонение магнитной стрелки.

Диффузионный ток играет очень важную роль при работе многих полупроводниковых приборов, в частности, диодов и транзисторов.

Чтобы уметь вычислять диффузионный ток, необходимо знать коэффициенты диффузии электронов и дырок, D_e и D_p . Не следует ли нам привести таблицу значений D_e и D_p для различных полупроводников при разных температурах, аналогично табл. III для значений подвижности?

Соотношение Эйнштейна. Нет, не следует! Между величиной коэффициента диффузии любых частиц и их подвижностью существует простое соотношение.

Перепишем выражение (21) в виде

$$D = v_T^2 \tau_0 / \beta \quad (23)$$

и подставим в него значение квадрата тепловой скорости из формулы (11):

$$D = v_T^2 \tau_0 / 3 = kT \tau_0 / m. \quad (24)$$

Сравнив теперь выражение (24) с формулой (13) для подвижности частиц, получим

$$D = kT \mu / e. \quad (25)$$

Зная величину подвижности при данной температуре, из формулы (25) легко найти значение коэффициента диффузии.

Связь между коэффициентом диффузии и подвижностью, устанавливаемая соотношением (25), справедлива не только для электронов и дырок, но и для любых частиц, заряженных и незаряженных *), движущихся в поле силы тяжести или в электрическом поле. Эта связь является наглядным отражением того обстоятельства, что и направленному движению частиц под действием силы, и процессу диффузии мешает один и тот же процесс: столкновения частиц, происходящие через средний промежуток времени τ_0 при средней тепловой скорости частиц v_T .

Связь между коэффициентом диффузии и подвижностью установлена Альбертом Эйнштейном и соотношение (25) называется соотношением Эйнштейна.

Скорость диффузии. Обсудим теперь важный вопрос о скорости диффузии. Или, возвращаясь к примеру с одеколоном, определим время t , через которое запах одеколона, пролитого в углу комнаты, можно будет почувствовать на расстоянии L от этого угла.

Решим поставленную задачу двумя способами. Методом размерностей, который позволит нам буквально в течение минуты получить ответ и удивиться его неожиданности. И приемом, который называется «путь пьяницы» и позволяет подробно проследить за процессом диффузии.

Итак, какое время t потребуется частицам с коэффициентом диффузии D , чтобы продиффундировать на заданное расстояние L ?

*) Для частиц, не обладающих электрическим зарядом, величину подвижности μ определяют как отношение средней скорости частицы $\bar{v}$ к силе F , под действием которой устанавливается эта скорость: $\mu = \bar{v}/F$. В этом случае связь между коэффициентом диффузии и подвижностью имеет вид: $D = kT\mu$.

Поскольку ответ может, очевидно, определяться только величинами L , измеряемой в метрах, и D — в метрах в квадрате на секунду, то после некоторого размышления можно убедиться, что единственный способ получить величину, имеющую размерность времени, состоит в том, чтобы разделить квадрат длины L на коэффициент диффузии D . Таким образом, на поставленный вопрос метод размерностей дает следующий ответ: $t \sim L^2/D$ или

$$L \sim \sqrt{Dt}. \quad (26)$$

Точный вывод, с которым можно ознакомиться в уже упоминавшемся ранее вып. 4 «Фейнмановских лекций по физике», позволяет определить численный коэффициент в формуле (26). Оказывается, что значение коэффициента зависит от того, происходит ли диффузия частиц в плоскости или в трехмерном пространстве. Однако во всех случаях этот коэффициент порядка единицы.

Соотношение (26) устанавливает необычную зависимость между временем диффузии и пройденным за

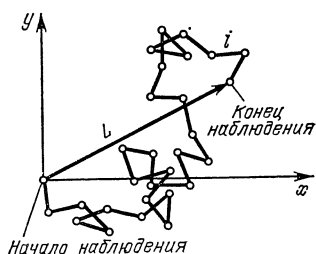


Рис. 33. В результате N случайно направленных шагов длиной l каждый, объект наблюдения удаляется от начальной точки на расстояние $L \approx l\sqrt{N}$ по прямой.

это время путем. Путь, пройденный молекулами в процессе диффузии, пропорционален не времени t , а корню квадратному из величины t . Что это значит?

Задача о случайных блужданиях. Представим себе, что мы в состоянии следить за движением отдельной молекулы одеколона. Под ударами молекул воздуха она хаотически мечется; каждый следующий удар бросает страдальцу в совершенно произвольном направлении (рис. 33). В промежутке между столкновениями молекула продвигается на расстояние, равное средней длине свободного пробега l . На каком расстоянии L от исходной точки наблюдения окажется молекула, испытав N столкновений?

Сформулированная задача о случайных блужданиях известна в математической статистике под назва-

нием «путь пьяницы». И для пьяницы, и для молекулы (по разным причинам) направление каждого следующего шага совершенно произвольно и никак не зависит от направления предыдущего шага. Поэтому интересующую нас задачу мы можем сформулировать и так: на какое расстояние L по прямой уйдет пьяный человек, сделавший N шагов длиной l .

Проведем на рис. 33 две взаимно перпендикулярных оси x и y и будем характеризовать каждый шаг $1, 2, 3, \dots, i, \dots, N$ его проекциями на эти оси $\Delta x_1, \Delta x_2, \Delta x_3, \dots, \Delta x_i, \dots, \Delta x_N$; $\Delta y_1, \Delta y_2, \Delta y_3, \dots, \Delta y_i, \dots, \Delta y_N$. Величины Δx_i и Δy_i , очевидно, с равной вероятностью могут быть как положительными, так и отрицательными. Их значения лежат в пределах от $-l$ до $+l$. Кроме того, ясно, что $\Delta x_i^2 + \Delta y_i^2 = l^2$.

Подсчитаем, чему будет равна величина L^2 после N случайных шагов:

$$L_N^2 = (\Delta x_1 + \Delta x_2 + \Delta x_3 + \dots + \Delta x_i + \dots + \Delta x_N)^2 + (\Delta y_1 + \Delta y_2 + \Delta y_3 + \dots + \Delta y_i + \dots + \Delta y_N)^2. \quad (27)$$

Если выполнить возведение в квадрат, то величина L_N^2 будет представлять собой сумму квадратов всех слагаемых $\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta y_1^2 + \Delta y_2^2 \dots$ плюс сумму попарных произведений вида $2\Delta x_1\Delta x_2 + 2\Delta x_1\Delta x_3 + \dots + 2\Delta y_1\Delta y_2 + \Delta y_1\Delta y_3 + \dots$.

Поскольку при случайных блужданиях величины Δx_i и Δy_i с равной вероятностью могут быть как положительными, так и отрицательными, при суммировании по большому числу шагов значение суммы попарных произведений равняется нулю. А каждая пара суммы квадратов вида $\Delta x_i^2 + \Delta y_i^2$ просто равняется l^2 . Поэтому после большого числа N шагов

$$L_N^2 = Nl^2 \quad \text{или} \quad L_N = l\sqrt{N}. \quad (28)$$

Итак, совершив N шагов длиной l каждый, пьяница (или диффундирующая молекула) удалится от точки, с которой началось путешествие, на расстояние L , определяемое формулой (28).

Но... число «шагов» молекулы, атома, электрона или пьяницы пропорционально времени t и обратно пропорционально среднему времени между столкновениями τ_0 : $N \sim t/\tau_0$. Длина свободного пробега

частицы l равняется, очевидно, $l = v_T \tau_0$. Таким образом $L \sim v_T \tau_0 \sqrt{t/\tau_0} \sim \sqrt{v_T^2 \tau_0 t}$. Или, учитывая выражение (23) для коэффициента диффузии, $L \sim \sqrt{D t}$.

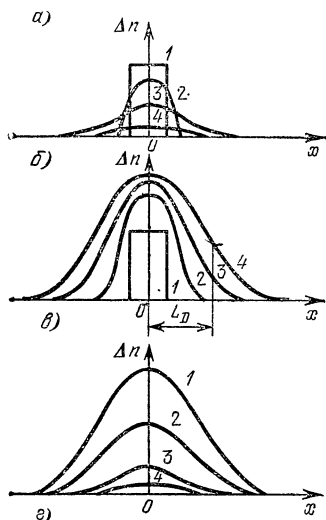
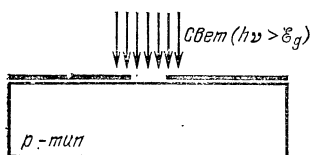


Рис. 34. Носители, возникающие в освещенной области полупроводника, распространяются в область тени благодаря диффузии.

а) Образец p -типа закрыт металлической пластиной с прорезанной в ней щелью; б) распределение носителей при короткой вспышке света в моменты времени $t_1 - t_4$ ($t_1 < t_2 < t_3 < t_4$); в) распределение носителей в моменты времени $t_1 - t_4$ при включении света ($t_1 < t_2 < t_3 < t_4$). Свет остается включенным очень долго; г) распределение носителей при выключении света ($t_1 - t_4$) ($t_1 < t_2 < t_3 < t_4$).

Итак, путь L , пройденный частицами в процессе диффузии за время t , определяется коэффициентом диффузии и пропорционален не времени, а корню квадратному из времени t .

Теперь, вооруженные общими представлениями о диффузии, перейдем к рассмотрению одного из самых важных в физике полупроводников диффузионных процессов — диффузии неосновных носителей.

Диффузия неосновных носителей. Неосновными носителями называются дырки в полупроводнике n -типа и электроны в полупроводнике p -типа. Такой термин представляется тем более естественным, что электроны в электронном, а дырки в дырочном полупроводнике называются *основными носителями*.

Обсудим опыт, схема которого показана на рис. 34.

Полупроводниковый образец (для определенности, p -типа) освещается светом с энергией кванта $h\nu$, большей, чем энергия образования электронно-ды-

рочной пары $\mathcal{E}_g$. Поверхность образца закрыта металлической пластиной с прорезанной в ней щелью. Металл света не пропускает, так что электронно-

дырочные пары возникают под действием света только непосредственно под щелью (рис. 34, а). Выберем источник света относительно небольшой интенсивности так, чтобы дополнительная концентрация основных носителей (дырок), возникших под действием света Δp , была бы гораздо меньше, чем исходная равновесная концентрация дырок p_0 . Неравновесная концентрация неосновных носителей (электронов) Δn , возникших благодаря действию света, конечно, равняется неравновесной концентрации дырок: $\Delta n = \Delta p$. Однако равновесная концентрация электронов n_p в полупроводнике p -типа гораздо меньше, чем концентрация дырок (см. формулу (9)). Поэтому изменению концентрации основных носителей на доли процента ($\Delta p \ll p_0$) может соответствовать изменение концентрации неосновных носителей на много порядков ($\Delta n = \Delta p \gg n_p$).

Включим свет и тут же выключим его (рис. 34, б). Допустим, что вспышка света была такой короткой, что возникшие под действием света электроны не успели ни расплыться под влиянием диффузии, ни рекомбинировать. В этом случае распределение электронов сразу же после окончания импульса света будет соответствовать кривой 1. Что произойдет с неравновесными носителями (электронами) в дальнейшем?

Во-первых, как мы знаем, неравновесные носители будут рекомбинировать с характерным временем жизни τ_n . Во-вторых, — перемещаться под действием диффузии из области высокой концентрации в область низкой. Кривые 2—4 (рис. 34, б) показывают, что происходит с электронным пакетом благодаря совместному действию процессов рекомбинации и диффузии.

Процесс очень напоминает картину, которую можно наблюдать в безветренный летний день, когда промчавшийся по проселочной дороге грузовик оставляет за собой шлейф пыли. Столб пыли расплывается к обочинам и постепенно оседает на землю.

Электроны, возникшие под действием света, рекомбинируют за время порядка времени жизни τ_n . В соответствии с выражением (26) это означает, что, «умирая», они успевают продвинуться из освещенной области, где они возникли, в область тени на

расстояние порядка *диффузионной длины* L_D :

$$L_D = \sqrt{D\tau_n}. \quad (29)$$

Прделаем следующий опыт. Включим свет и оставим его включенным неопределенно долго (рис. 34, *в*). Сразу же после включения света неравновесные электроны существуют только в освещенной области (кривая 1). Затем электронный пакет начинает, благодаря диффузии, расплываться, однако процесс генерации носителей светом преобладает над процессом рекомбинации. Концентрация в освещенной области продолжает возрастать (кривые 2 и 3). Кривая 4 соответствует стационарному состоянию. Распределение электронов, соответствующее кривой 4, будет сохраняться неизменным, пока свет останется включенным. Число носителей, создаваемых светом, в точности равно числу носителей, погибающих за счет рекомбинации. В каждой точке число носителей, доставляемых диффузией из соседней области более высокой концентрации, равняется сумме числа носителей, уходящих из этой точки в соседнюю область более низкой концентрации, и числа рекомбинирующих носителей.

Созданные в освещенной области неравновесные носители проникают в область тени на расстояние порядка нескольких диффузионных длин L_D . Картину, подобную показанной на рис. 34, *в*, можно иногда наблюдать на деревенской околице, когда сорванец, забравшийся в пыль, самозабвенно ворошит ее палкой. В воздух поднимается мутное облако. Оно растет, расплывается и, если паренек достаточно терпелив, а день — безветренный, достигает стационарного состояния. Сколько пыли парнишка взмывает в воздух, столько оседает ее за то же время. Лучше не приближаться к сорванцу на расстояние нескольких «диффузионных длин» — последствия для вашего костюма могут оказаться печальными.

Рис. 34, *г* показывает, что произойдет, если выключить свет. Под действием рекомбинации концентрация избыточных носителей постепенно уменьшается.

Итак, во второй части книги мы установили причины, по которым «принцы на горошине» — полупроводники — обладают способностью заметно изменять

проводимость при изменении температуры. Удивительная чувствительность полупроводников к температуре связана с тем, что не только подвижность, как в металлах, но и концентрация свободных носителей — электронов и дырок в полупроводниках резко зависит от температуры. По той же причине проводимость полупроводников может меняться при освещении: свет (с определенной длиной волны) создает в полупроводнике дополнительные электроны и дырки.

В следующей, третьей, части книги мы рассмотрим, как используются замечательные свойства полупроводников.

ПРОФЕССИИ ПРИНЦЕВ

Кто честным кормится трудом —
Таких зову я знатью.

Р. Бернс

У полупроводников — тысячи профессий. Даже простое перечисление их потребовало бы книги значительно более толстой, чем та, которая сейчас лежит перед вами.

Важные и интересные профессии полупроводников связаны с диодами и транзисторами. Именно эти приборы и интегральные схемы на их основе определяют современный уровень электроники, микроэлектроники, вычислительной техники. Для понимания принципов работы диодов и транзисторов, помимо тех свойств полупроводников, с которыми мы познакомились в предыдущих главах, требуется также обстоятельное знакомство со сложными и интересными явлениями, возникающими на границе между различными полупроводниками ($p-n$ -переход, гетеропереход) или между полупроводником и металлом (барьер Шоттки).

Но это — тема для другой книжки.

Полупроводниковые приборы, о которых мы расскажем в этой части, по конструкции предельно просты. Они представляют собой всего-навсего небольшие кристаллики полупроводника с контактами. Однако, благодаря замечательным физическим свойствам полупроводников, даже эти простейшие приборы способны решать множество трудных, важных и интересных задач в самых разных областях науки и техники.

ТЕРМОРЕЗИСТОРЫ

Требуются очень глубокие знания, чтобы заметить простейшие, но подлинные отношения вещей между собой.

Г. К. Лихтенберг

Терморезистор, термосопротивление, термистор — слова-синонимы. Резистор (resistor) по-английски означает сопротивление *), но в последние годы это слово так прочно вошло в русский язык, что практически не требует перевода. А термистор — просто широко распространенное сокращение от слова «терморезистор».

Все три слова означают одно и то же: кусочек полупроводникового материала с омическими контактами — пожалуй, самый простой из всех возможных полупроводниковых приборов.

А между тем...

Термисторы используются для измерения и регулирования температуры в диапазоне от ~ 1 К до температуры расплавленной стали (~ 1800 К). Для температурной стабилизации различных элементов электротехнической и радиоэлектронной аппаратуры. Для пожарной сигнализации. Для измерения низких давлений в вакуумной аппаратуре, измерения мощности различных сверхвысокочастотных устройств, измерения скорости движения жидкостей и газов. Для контроля тепловых режимов машин и механизмов.

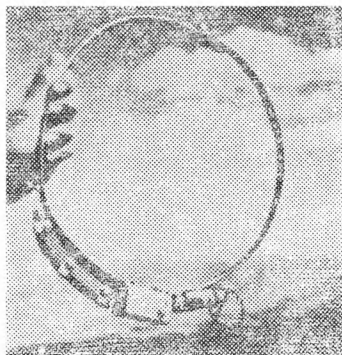


Рис. 35. Телеметрический ошейник для медведя гризли в Йеллоустонском Национальном Парке (США). На ошейнике смонтирован полупроводниковый передатчик, позволяющий непрерывно следить за местоположением медведя и, отчасти, за его самочувствием: температура тела животного измеряется с помощью миниатюрного терморезистора.

*) От латинского *resistere* — сопротивляться.

Для контроля температуры тяжелобольных в реанимационных палатах и для дистанционного наблюдения за состоянием здоровья редких и ценных животных (рис. 35). Для изучения спектра Солнца и звезд. Термосопротивления используются также в качестве бесконтактных переменных сопротивлений, реле времени, автоматических потенциометров, генераторов, усилителей и модуляторов низкой частоты, стабилизаторов и предохранителей.

Не правда ли, неплохой послушной список для такого скромного прибора? И, кстати сказать, славная родословная. Первые термисторы были изготовлены в 1890 г. великим немецким физиком Нернстом.

Как же терморезистор ухитряется выполнять все эти разнообразные функции? Начнем с самого простого — измерения и регулирования температуры.

Просто термосопротивление

Сопротивление полупроводника очень сильно зависит от температуры. По изменению сопротивления термистора можно судить об изменении температуры. Очень просто.

В принципе, действительно, несложно. Разберемся, однако, внимательнее в вопросах, с которыми приходится сталкиваться тем, кто конструирует и использует термисторы.

Прежде всего, какой полупроводниковый материал выбрать для изготовления термосопротивления?

Низкие температуры. Допустим, речь идет об измерении низких температур порядка нескольких кельвинов. Мы помним, что при температуре абсолютного нуля полупроводники превращаются в изоляторы. Если температура отлична от нуля, но очень мала, удельное сопротивление полупроводника может быть недопустимо велико. Воспользовавшись формулой (6) и данными, приведенными в табл. II, легко установить, что даже в InSb — полупроводнике с очень небольшой энергией образования электронно-дырочной пары $\mathcal{E}_g$, концентрация собственных электронов и дырок $n_i = p_i$ при $T \sim 1$ К составит $\leq 10^{-400}$ см⁻³! Конечно, в реальных условиях сопротивление образца, изготовленного из такого материала, будет определяться не собственной проводи-

мостью, а утечкой тока по поверхности, неизбежными случайными примесями и загрязнениями и т. д. Сопротивления разных образцов, даже изготовленных в одинаковых условиях, будут отличаться в сотни раз, сопротивление каждого образца будет нестабильным. Короче говоря, практически такие термосопротивления никуда не будут годны.

Мы знаем о полупроводниках достаточно много, чтобы догадаться, в чем состоит выход. Необходимо легировать полупроводник мелкой примесью. В германии энергия ионизации мелких примесей (доноров и акцепторов) $\Delta\mathcal{E}$ составляет всего около 0,01 эВ. Но для температур порядка 1 К это не такая уж малая величина. Значение экспоненты $e^{-\Delta\mathcal{E}/2kT}$ в формулах (7) и (10) составляет при $\Delta\mathcal{E} = 10^{-2}$ эВ и $T = 1$ К величину $\approx 5,6 \cdot 10^{-26}$. Так что при концентрации, например, доноров $N_d \sim 10^{15} - 10^{16}$ см⁻³ примесная концентрация n_d составит $\sim 10^{-10}$ см⁻³. И такая величина n_d соответствует недопустимо большому удельному сопротивлению.

Что еще можно придумать? Мы готовы к ответу на этот вопрос. Взглянем еще раз на рис. 19. Энергия активации примеси убывает с ростом концентрации примесных центров! Прекрасно. Подбирая нужную концентрацию мелких доноров N_d , можно уменьшить энергию активации $\Delta\mathcal{E}$ до необходимой величины и получить желаемую концентрацию n_d , а следовательно, и удельное сопротивление полупроводника ρ .

Итак, мы нашли необходимое решение. Исчерпана ли этим наша задача? Отнюдь нет. Мы поняли, как создать термосопротивление, чувствительное в области низких температур. Не следует, однако, забывать о явлении примесного истощения. Если ввести в полупроводник только одну примесь с очень малой энергией ионизации, то созданное на основе такого материала термосопротивление будет чувствительным только в очень узком температурном интервале. В области чуть более высоких температур наступит явление примесного истощения, и (см. рис. 17) концентрация свободных носителей перестанет зависеть от температуры. Между тем, как правило, желательно иметь возможность, не меняя датчика (термосопротивления), вести измерения в достаточно широком диапазоне температур.

Значит? Значит, необходимо ввести в полупроводник не одну, а много примесей с различными энергиями активации $\Delta\mathcal{E}$. И при этом так выбрать величины $\Delta\mathcal{E}$, чтобы при температуре, соответствующей наступлению примесного истощения для одного уровня, для другого начиналась бы эффективная тепловая ионизация. Помня, как чувствительны полупроводники к различного рода примесям, мы вполне можем оценить практические трудности *точно дозированного, воспроизводимого многокомпонентного легирования полупроводников.*

Тем не менее современная технология успешно справляется с этой труднейшей задачей. На рис. 36, а

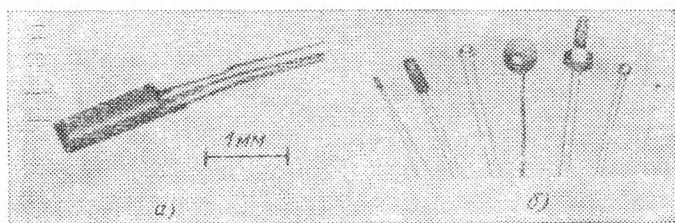


Рис. 36. Терморезисторы.

а) Термистор на основе многокомпонентно легированного германия. Диапазон рабочих температур 1,3 — 273 К. Воспроизводимость при многократных измерениях температуры не хуже 0,0005 К. Термочувствительный элемент (пленка германия толщиной ~ 200 мкм и диаметром $\sim 0,6$ мм) имеет массу всего 0,0001 г. Миниатюрность прибора позволяет измерять температуру в небольших объемах и обеспечивает низкую тепловую инерционность прибора; б) различные конструктивные оформления кремниевых термосопротивлений.

показан миниатюрный германиевый термистор, позволяющий проводить измерения в области температур от 1,3 до 273 К *).

Форма, размеры, конструктивное оформление корпуса термистора могут быть самыми разнообразными и зависят от конкретной задачи, для решения которой предназначен прибор. На рис. 36, б показаны термисторы на основе кремния. Кремниевые термо-

*) Если внимательно приглядеться к рис. 36, а, то можно увидеть, что из корпуса прибора выходят не два, а четыре вывода. Для повышения точности сопротивления измеряется так, как показано на рис. 31. По двум выводам на терморезистор подается измерительный ток, с двух других — снимается измеряемое напряжение.

сопротивления используются для измерения и контроля температур в диапазоне от 10 до 400 К.

Высокие температуры. Первое требование, которому должен удовлетворять полупроводниковый материал, пригодный для изготовления высокотемпературных термосопротивлений, нам, как физикам, совершенно ясно. Энергия образования электронно-дырочной пары $\mathcal{E}_g$ в таком материале должна быть большой.

Кроме того, когда речь идет о высоких температурах, первостепенное значение приобретают такие качества материала, как химическая и структурная стабильность: не так-то просто найти материал, способный выдержать многократный нагрев, скажем, до 1000 °С и при этом не изменить своих свойств. Очень непросто подобрать и контакты, которые должны выдерживать высокие температуры и иметь приблизительно такой же, как и сам полупроводник, коэффициент температурного линейного расширения (Физика, 9 класс). Если последнее требование не удовлетворяется, после нескольких циклов нагрева и охлаждения контакты могут просто отлететь или вызвать растрескивание полупроводника.

Из полупроводников, уже упоминавшихся в нашей книге, требованиям большой величины $\mathcal{E}_g$, высокой химической и структурной стабильности в наибольшей степени отвечают полупроводниковые алмазы и карбид кремния (SiC). Эти материалы действительно применяются для изготовления высокотемпературных термосопротивлений, предназначенных для работы вплоть до температур 300—400 °С.

Для еще более высоких температур, вплоть до 1200—1500 °С, применяются материалы, которые из-за малой величины проводимости при комнатной температуре относятся скорее к диэлектрикам, чем к полупроводникам: окись алюминия (Al_2O_3), титанаты магния, цинка и кадмия (MgOTiO_2 , ZnOTiO_2 , CdOTiO_2), смеси окислов титана и кобальта.

Не очень высокие и не очень низкие температуры. Ну, хоть тут-то дело должно обстоять, действительно, просто! На роль материалов, пригодных для изготовления термисторов, могут претендовать десятки (если не сотни) различных полупроводников.

И... возникает очень непростая задача — как отобрать из многих претендентов самый достойный:

простой, технологичный и воспроизводимый в изготовлении, надежный и стабильный в эксплуатации, дешевый (очень немаловажное требование для прибора, который выпускается миллионными сериями).

В результате долгой, упорной, хотя и бескровной борьбы трон в королевстве термосопротивлений достался принцам из семейства окислов переходных металлов: марганца (Mn), кобальта (Co), никеля (Ni) и меди (Cu). На основе этих материалов промышленностью изготавливаются десятки типов термосопротивлений. Большинство из них предназначено для работы в области температур от -60 до $+120^\circ\text{C}$, хотя некоторые типы термисторов на основе оксидных полупроводников могут работать и при более низкой (вплоть до $\sim 10\text{ K}$) и при более высокой (до 300°C) температуре. Конструктивное оформление термосопротивлений может быть самым разнообразным. Термисторы изготавливаются в виде стержней, трубок, дисков, шайб, бусинок, пластин и т. д. Значение сопротивлений термисторов различных типов лежит в пределах от нескольких ом до нескольких миллионов ом. (Конечно, по самому назначению терморезистора его сопротивление сильно зависит от температуры. Значение сопротивления указывается обычно для какой-либо температуры вблизи середины рабочего интервала. Чаще всего — для $+20^\circ\text{C}$.)

Температурный коэффициент сопротивления. Из какого бы материала ни был изготовлен термистор, для каких бы применений он ни был предназначен, в каком бы температурном диапазоне ни использовался, — важнейшим параметром термистора является *температурный коэффициент сопротивления*.

Температурный коэффициент сопротивления α_T показывает, как меняется сопротивление термистора с изменением температуры:

$$\alpha_T = \frac{1}{R} \cdot \frac{\Delta R}{\Delta T}. \quad (30)$$

Выражение (30) означает следующее. Пусть при некоторой температуре T сопротивление терморезистора равняется R . Если изменить температуру на величину ΔT , сопротивление изменится. Оно будет равняться теперь $R + \Delta R$. Таким образом, в соответствии с выражением (30), температурный коэф-

эффициент сопротивления представляет собой относительное изменение сопротивления при изменении температуры на один градус.

Сопротивление термистора определяется его геометрическими размерами и удельным сопротивлением материала, из которого изготовлен термистор. Геометрические размеры при изменении температуры меняются мало и изменение сопротивления термистора определяется зависимостью от температуры удельного сопротивления ρ :

$$\alpha_T = \frac{1}{\rho} \cdot \frac{\Delta \rho}{\Delta T}. \quad (31)$$

Для металлов, как мы знаем (Физика, 9 класс), величина α_T очень слабо зависит от температуры, положительна (удельное сопротивление металла растет с ростом температуры) и для чистых металлов составляет приблизительно $0,0035 \text{ K}^{-1}$. При изменении температуры на один градус сопротивление возрастает приблизительно на $0,35 \%$.

Область рабочих температур термистора всегда выбирается так, чтобы при изменении температуры сильно (экспоненциально) изменялась концентрация носителей в полупроводнике, из которого термистор изготовлен. При этом, как мы уже знаем (см. гл. 5), изменение подвижности мало влияет на температурную зависимость проводимости σ . Зависимость σ от температуры будет определяться выражением, аналогичным формулам (6), (7) и (10):

$$\sigma = \sigma_0 e^{-\Delta \mathcal{E}/2kT}. \quad (32)$$

Так как удельное сопротивление $\rho = 1/\sigma$, то

$$\rho = \rho_0 e^{\Delta \mathcal{E}/2kT}. \quad (33)$$

Зная зависимость $\rho(T)$, определяемую выражением (33), легко найти величину α_T , дифференцируя зависимость $\rho(T)$ по температуре:

$$\alpha_T = \frac{1}{\rho} \cdot \frac{d\rho}{dT} = -\frac{\Delta \mathcal{E}}{2k} \cdot \frac{1}{T^2} [\text{K}^{-1}]. \quad (34)$$

Таким образом, как показывает выражение (34), для полупроводников величина α_T сильно зависит от температуры ($\alpha_T \sim 1/T^2$) и отрицательна (удельное сопротивление полупроводника при нагревании уменьшается).

Что касается величины α_T , то ее, как правило, выгодно иметь по возможности большей. Чем больше α_T , тем чувствительней сопротивление термистора к изменению температуры, тем легче зафиксировать, следовательно, небольшие изменения температуры и тем точнее можно измерять и контролировать температуру. На первый взгляд кажется, что, в соответствии с выражением (34), нетрудно получить очень большие значения α_T , в особенности при низких температурах. Стоит только выбрать полупроводник с достаточно большой энергией активации $\Delta\mathcal{E}$, и... Но мы уже достаточно опытные, чтобы не поддаваться обольщению первых впечатлений. Если отношение $\Delta\mathcal{E}/kT$ слишком велико — концентрация носителей, а следовательно, и величина проводимости становятся недопустимо малыми (см. раздел «Низкие температуры» этой главы).

Практически величина α_T для термисторов различных типов лежит в пределах от $-0,02$ до $-0,08 \text{ K}^{-1}$. При изменении температуры на один градус сопротивление термистора изменяется на 2—8 % — на порядок больше, чем в металлах.

Теперь, познакомившись с основными свойствами термисторов, обсудим, как они ухитряются справляться со своими многочисленными обязанностями.

Измерение температуры. Измерение температуры сводится просто к измерению сопротивления термистора. Зная величину сопротивления, по формуле, аналогичной формуле (33),

$$R = R_0 e^{B/T}, \quad (35)$$

можно определить температуру; для практических расчетов, однако, из формулы (35) удобно исключить постоянную R_0 .

Пусть при температуре T_1 сопротивление термистора равняется

$$R_1 = R_0 e^{B/T_1}. \quad (36)$$

При температуре T_2 сопротивление будет равняться

$$R_2 = R_0 e^{B/T_2}. \quad (37)$$

Разделив (37) на (36), получим:

$$R_2 = R_1 e^{B\left(\frac{1}{T_2} - \frac{1}{T_1}\right)} \quad \text{или} \quad T_2 = \frac{B + T_1 \ln R_2/R_1}{B/T_1}. \quad (38)$$

Зная величину сопротивления R_1 при какой-либо температуре T_1 и значение постоянной B , по формуле (38) можно определить температуру при известном сопротивлении R_2 . Значения постоянной B и сопротивления R_1 при некоторой температуре T_1 , соответствующей обычно середине рабочего интервала температур, приведены в паспорте термистора. Величина постоянной B лежит в пределах от ~ 25 К для самых низкотемпературных до $\sim 25\,000$ К для самых высокотемпературных термисторов.

К термосопротивлениям, предназначенным для особенно точных (прецизионных) измерений, и к термисторам, рассчитанным на работу в очень широком интервале температур, в котором закон изменения удельного сопротивления $\rho(T)$ (33) не выполняется, прилагается заранее снятый для каждого прибора отдельно график зависимости сопротивления от температуры. В этом случае, измерив сопротивление термистора, значение температуры просто определяют по графику.

Температурная компенсация. Многие электроизмерительные приборы должны работать в довольно широком интервале температур. Например, с требованием, чтобы вольтметр или амперметр могли нормально работать в диапазоне температур от -30 до $+50^\circ\text{C}$, сохраняя заданную точность, приходится встречаться довольно часто. Между тем, в большинстве вольтметров и амперметров одним из самых основных узлов является легкая проволочная рамка, поворачивающаяся в поле постоянного магнита при протекании по ней измеряемого тока. Рамка изготавливается из меди. Температурный коэффициент сопротивления меди $\alpha_T = +0,0039\text{ K}^{-1}$. Таким образом, при изменении температуры на 80 градусов сопротивление рамки гальванометра изменится на $\sim 30\%$!

Ни о каком сохранении точности в таких условиях говорить, конечно, не приходится.

Понятно, что нужно сделать, чтобы исправить положение. Последовательно с металлической рамкой следует включить термистор, обладающий отрицательным температурным коэффициентом сопротивления.

Тогда при увеличении температуры сопротивление рамки будет расти, сопротивление термистора — падать, и изменение сопротивления термистора

компенсирует температурную зависимость сопротивления меди.

Правда, может возникнуть законный вопрос. Величина α_t у металлов от температуры зависит слабо, а у термисторов — сильно. Кроме того, сами величины α_t для металлов и полупроводников заметно различаются. Как в таких условиях можно обеспечить точную компенсацию?

В практических схемах используются параллельные и последовательные соединения терморезисторов с высокостабильными манганиновыми или константановыми сопротивлениями. Параметры элементов подбираются так, чтобы обеспечить компенсацию в широком диапазоне температур *).

Использование термисторных схем температурной компенсации позволяет в десятки раз снизить чувствительность рабочих параметров приборов к изменению температуры внешней среды.

Термистор — переменное сопротивление. Каждому из нас приходится иметь дело с переменными сопротивлениями: они служат, в частности, для регулировки тембра и громкости в радиоприемниках, проигрывателях, телевизорах, магнитофонах. И каждый владелец такого прибора знает, что после трех-четырех лет эксплуатации вращение оси переменного сопротивления (потенциометра) начинает сопровождаться неприятными потрескиваниями и хрипами. Иногда звук совсем пропадает, а иногда без видимой причины резко меняется громкость. «Ну, все. Контакт стерся — пора менять», — делает заключение домашний или приглашенный эксперт.

В обычном переменном сопротивлении металлический контакт перемещается по токопроводящему слою и постепенно стирает его. Сопротивление выходит из строя. А ведь мы поворачиваем ручку потенциометра всего 5—10 раз в день.

Между тем, существуют схемы (например, схемы автоматического регулирования), в которых необходимо изменять величину сопротивления несколько раз в минуту. Сколько же прослужит в такой схеме обыкновенный потенциометр?

*) О схемах компенсации можно прочитать, например, в книге: Нечаев Г. К., Удалов Н. П. Реле и датчики с полупроводниковыми термосопротивлениями. — М.: Госэнергоиздат, 1961.

Незаменимыми в такого рода схемах оказываются так называемые термисторы с косвенным подогревом. Такой прибор представляет собой термистор, вблизи которого располагается миниатюрная подогревная обмотка. При пропускании по этой обмотке тока она нагревается, нагревает термистор и изменяет его сопротивление. Никаких механических контактов, износа и связанных с ними неприятностей!

Сопротивление обмотки подогревателя составляет обычно несколько десятков ом, ток подогрева — 20—40 мА. Так что мощность, необходимая для управления величиной сопротивления, невелика — порядка десятых или даже сотых долей ватта. Миниатюрные размеры термистора и подогревной обмотки позволяют сконструировать термисторы с косвенным подогревом, обладающие маленькой тепловой инерцией: при изменении тока в обмотке новое значение сопротивления устанавливается уже через 5—20 с.

Важным преимуществом термисторов с косвенным подогревом перед обычными потенциометрами является возможность очень легко регулировать сопротивление дистанционно, на любом расстоянии от управляемого объекта или схемы.

Термисторы с косвенным подогревом серийно выпускаются нашей промышленностью в нескольких модификациях и используются в радиотехнических, электротехнических и телемеханических устройствах.

Очень эффективным оказывается использование термисторов с косвенным подогревом для измерения скорости движения жидкостей или газов. Принцип измерения очень прост. При неизменном токе в подогревной обмотке температура термистора будет тем меньше (а сопротивление, соответственно, тем больше), чем быстрее обтекается термистор потоком воздуха или жидкости, в которую термистор погружен. Кто не верит — пусть лизнет палец и подует на него.

Высокий температурный коэффициент сопротивления позволяет использовать термисторы для измерения даже очень небольших скоростей потоков жидкостей и газов, вплоть до ~ 1 мм/с.

Болометры. Греческое слово *bolē* означает луч. Обычное назначение болометра состоит в измерении энергии теплового излучения, как правило, очень слабого. Источником такого излучения может быть

свет звезд или Солнца, прошедший через спектрометр и разложенный на тысячи линий, энергия в каждой из которых очень мала. Изобрел болометр в 1880 г. американский астрофизик Сэмюэл Ленгли.

По существу, полупроводниковый болометр представляет собой термосопротивление, однако, назначение болометра — изменять свое сопротивление под действием очень слабых тепловых потоков, — предъявляет особые требования к конструкции прибора.

Рабочий элемент болометра представляет собой очень тонкую, от долей микрометра до нескольких микрометров, пленку полупроводникового материала, нанесенную на стеклянную или кварцевую подложку. Для исключения влияния колебаний температуры окружающей среды на показания болометра, как правило, используются два прибора с одинаковыми характеристиками, включенные в специальную (так называемую «мостовую» схему *). В такой схеме изменение сопротивления одного болометра под действием колебаний температуры практически полностью компенсируется точно таким же изменением сопротивления размещенного рядом идентичного болометра — компенсатора.

С помощью современных электронных схем оказывается возможным зарегистрировать изменение температуры болометра на десятиллионные доли градуса. Благодаря этому полупроводниковые болометры позволяют обнаружить излучение, мощность которого составляет десятиллионные доли ватта.

Секреты буквы S

Чтобы измерить величину сопротивления любого резистора, в том числе и терморезистора, необходимо пропустить через него ток. Измеряется ли сопротивление известным нам с 7-го класса методом вольтметра и амперметра, стрелочным тестером или современным цифровым электронным измерителем, — принцип измерения остается одним и тем же. Через элемент пропускается известный ток. Измеряется падение напряжения на элементе при этом токе. Отношение напряжения V к току I определяет сопротивление R .

*) С этой схемой мы подробно ознакомимся в гл. 7.

Если в руки вам когда-нибудь попадет инструкция по пользованию термистором, в особенности, «деликатным», миниатюрным термистором для измерения низких температур или болометром, то в начале этой инструкции, в середине и в конце будет многократно повторено предупреждение: «Не пропускайте через прибор ток, больший указанного в настоящей инструкции!».

Может показаться, что составители инструкции обеспокоены тем, чтобы неумелый экспериментатор не сжег прибор. Ничуть не бывало. Пропускание тока, даже во много раз большего, чем указанное значение, ничем плохим прибору не грозит. Авторы инструкции в данном случае заботятся не о приборе, а о нас. Если наша цель — измерить с помощью термистора температуру окружающей среды, то прежде всего следует позаботиться о том, чтобы термистор не разогревался выше температуры среды проходящим через него током!

Но иногда так тянет нарушить даже самую полезную инструкцию!

Давайте попробуем пропустить через термистор ток, гораздо больший, чем тот, который указан в инструкции. При этом, памятуя завет Козьмы Пруtkова: «Бросая в воду камешки, смотри на круги, ими образуемые; иначе такое бросание будет пустою забавой», — будем одновременно измерять и напряжение на термисторе. Иначе говоря, изучим вольтамперную характеристику термистора при больших токах.

Помимо естественной любознательности, дополнительным оправданием нарушения инструкции нам может послужить следующее обстоятельство. Как мы помним, термосопротивление представляет собой просто кусочек полупроводника с двумя омическими контактами. Таким образом, изучая вольтамперную характеристику термистора, мы в действительности исследуем вид вольтамперной характеристики типичного полупроводника при больших плотностях тока.

S-транная вольтамперная характеристика. Типичная вольтамперная характеристика термистора (напомним, что в то же время — это типичная вольтамперная характеристика полупроводника!) показана на рис. 37. Выглядит она очень необычно, поэтому проследим за ее ходом особенно внимательно.

Начнем с области вблизи начала координат, где токи и напряжения малы. В этой области вольтамперная характеристика представляет собой прямую линию. Или, как иногда говорят, вольтамперная характеристика линейна. Ток пропорционален напряжению, выполняется закон Ома. Точка *A* отмечает границу линейного участка вольтамперной характеристики.

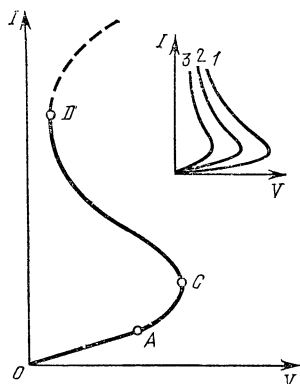


Рис. 37. Типичная вольтамперная характеристика полупроводника. На вставке показаны три вольтамперные характеристики одного и того же образца, снятые при трех различных температурах окружающей среды $T_1 < T_2 < T_3$.

Когда протекающий ток будет больше, вольтамперная характеристика начинает заметно отклоняться от прямой линии. С ростом тока напряжение на образце растет медленнее, чем это предписывается законом Ома. Можно выражаться и иначе: с ростом напряжения на образце ток растет быстрее, чем по линейному закону. Понимаем ли мы, что происходит? Безусловно. При протекании тока образец разогревается джоулевым теплом. Температура образца растет, растет, следовательно, и концентрация электронов в нем.

Больше становится концентрация свободных электронов (или дырок) — падает удельное сопротивление материала и уменьшается сопротивление образца. Значит, при данном напряжении на образце величина тока оказывается больше, чем если бы образец током не разогревался. Дальнейшее увеличение тока влечет за собой еще больший разогрев образца и еще большее увеличение концентрации электронов. Вот почему так резко начинает возрастать ток, когда разогрев образца джоулевым теплом становится заметным (участок *A—C* вольтамперной характеристики).

Влияние разогрева на вольтамперную характеристику образца заслуживает более подробного анализа.

Пусть по образцу течет ток I ; при этом падение напряжения на образце равняется V . В образце выделяется мощность $P = IV$. Если бы эта мощность целиком шла на нагрев образца и не рассеивалась бы в окружающее пространство, то за достаточно большое время образец разогрелся бы до сколь угодно высокой температуры и перестал бы существовать.

К счастью, этого не происходит. Как только образец разогревается проходящим током до температуры, большей, чем температура окружающей среды T_0 , он начинает отдавать тепло. За счет теплопроводности окружающего воздуха, конвекции и излучения. Через некоторое время устанавливается тепловое равновесие.

Каждая точка вольтамперной характеристики соответствует тепловому равновесию между нагревом образца протекающим током и его охлаждением за счет отвода тепла в окружающую среду. Образец при этом нагрет до некоторой температуры $T' > T_0$ и отдает в окружающую среду ровно столько же энергии в единицу времени P , сколько получает.

Даже на участке OA (рис. 37), где вольтамперная характеристика выглядит линейной, точные измерения зафиксировали бы небольшой разогрев образца протекающим током и незначительное отклонение от закона Ома.

Когда ток становится большим чем значение, отвечающее точке C , на первый взгляд, происходит нечто удивительное. *Ток растет, а напряжение на образце падает.* Однако в таком ходе вольтамперной характеристики, в сущности, нет даже ничего особенно странного. (Следствия, к которым приводит такой вид вольтамперной характеристики, действительно выглядят удивительно. Но об этом — позже.) Концентрация свободных носителей в полупроводнике очень резко, экспоненциально, зависит от температуры (см. формулы (6), (7), (10)). Поэтому с увеличением тока концентрация носителей, а следовательно, и проводимость образца могут расти так быстро, что *нужно меньшее чем прежде напряжение, чтобы поддержать тот же или даже больший ток*, если образец уже сильно разогрет джоулевым теплом и концентрация носителей в нем заметно возросла.

Чтобы еще отчетливее понять, что происходит, перепишем уравнение (14) в таком виде:

$$V = IR = I\rho(I) \frac{L}{S} = \frac{I}{\sigma(I)} \frac{L}{S}. \quad (39)$$

Удельное сопротивление $\rho(I)$ и проводимость $\sigma(I)$ в данном случае зависят от протекающего через образец тока I . Из уравнения (39) видно, что если с ростом тока проводимость растет быстрее, чем по линейному закону, то напряжение на образце V будет падать с ростом тока. Именно это и происходит на падающем участке вольтамперной характеристики (участок $C-D$) из-за очень резкой зависимости σ от температуры, характерной для полупроводников.

Подчеркнем, что с ростом тока I на участке $C-D$ мощность, выделяющаяся в образце $P = VI$, растет (несмотря на падение напряжения V). Это ясно из следующего простого рассуждения. Напряжение может падать с ростом тока, только если увеличивается проводимость σ . Величина σ увеличивается только, если с увеличением тока растет температура образца. А для того чтобы росла температура, необходимо, чтобы мощность, выделяющаяся в образце, увеличивалась с ростом тока.

Из рис. 37 видно, что еще большее увеличение тока (участок выше точки D) снова приводит к росту напряжения на образце V . Этот участок возникает на вольтамперной характеристике образца, когда протекающий ток разогреет его до температуры, соответствующей температуре примесного источника.

В этом случае дальнейшее повышение температуры уже не будет сопровождаться ростом проводимости σ и напряжение начнет расти с увеличением тока.

Участок выше точки D показан на вольтамперной характеристике пунктиром. Дело в том, что далеко не во всех ситуациях этот участок оказывается достижимым. Иногда и, увы, не так уже редко образец или контакты к нему, или и то и другое сгорают раньше, чем на вольтамперной характеристике удается наблюдать второй возрастающий участок.

Вольтамперная характеристика, показанная на рис. 37, называется S -образной (за сходство ее формы с латинской буквой S).

S-образная вольтамперная характеристика *многозначна по напряжению*: одному и тому же значению напряжения V на образце могут соответствовать два или даже три значения протекающего через образец тока I . И, соответственно, два или три значения температуры образца: чем больше ток, тем выше и температура образца.

Интересно рассмотреть, что происходит с вольтамперной характеристикой полупроводника, если снимать ее при различных температурах окружающей среды T_0 . Как будет меняться вольтамперная характеристика с изменением T_0 , показывает вставка к рис. 37. Чем больше T_0 , тем меньше начальное (при малых V и I) сопротивление образца, и тем больший ток соответствует тому же напряжению. С повышением T_0 напряжение, которому соответствует начало падающего участка вольтамперной характеристики, уменьшается. Вольтамперная характеристика «прижимается» к оси токов и «спрямляется». При достаточно высокой температуре падающий участок на вольтамперной характеристике может исчезнуть вовсе. И это понятно. Если температура окружающей среды T_0 так велика, что соответствует примесному истощению, то повышение температуры образца за счет джоулева разогрева не будет сопровождаться ростом проводимости. Следовательно, исчезнет и причина, вызывающая появление на вольтамперной характеристике падающего участка.

Волшебная прямая. Прежде чем начать знакомиться с чудесами, на которые способен образец с *S*-образной вольтамперной характеристикой, необходимо научиться определять значения токов и напряжений в цепи, в которую такой образец включен. Или, выражаясь строго научно, необходимо научиться рассчитывать цепи, содержащие элемент с *S*-образной характеристикой.

Для начала попробуем свои силы на более простой задаче. Рассчитаем цепь, показанную на вставке к рис. 38,а. К двум последовательно соединенным сопротивлениям R_1 и R_2 приложено напряжение V_0 . Необходимо определить ток, текущий в цепи, и падения напряжений на сопротивлениях. Задача, конечно, детская, и нам ничего не стоит ее решить, воспользовавшись стандартными формулами (Физика, 9 класс). Мы решим эту задачу, однако, другим,

графическим методом, вся красота и польза которого станет очевидна немного позже.

Графический метод решения иллюстрируется рис. 38, а.

Построим сначала вольтамперную характеристику сопротивления R_2 . Эта характеристика, естественно, будет представлять собой прямую линию, выходящую из начала координат под углом φ_2 (закон

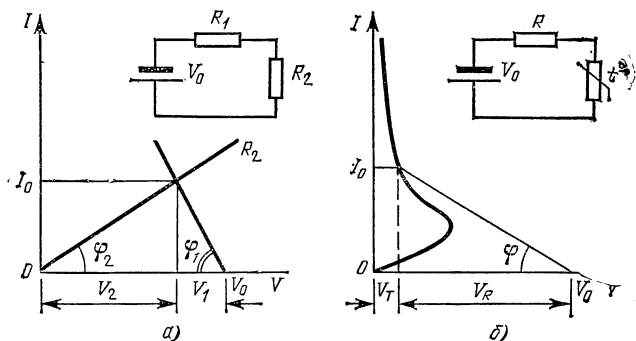


Рис. 38. Графический метод расчета электрической цепи.

а) Для двух последовательных сопротивлений; б) для цепи, содержащей термистор (или любой другой образец со сложной вольтамперной характеристикой).

Ома). При этом $\operatorname{tg} \varphi_2 = 1/R_2$ (или $R_2 = \operatorname{ctg} \varphi_2$). Затем отложим на оси напряжений V величину, равную V_0 , и проведем из этой точки прямую под углом φ_1 к оси напряжений. Угол φ_1 выбран так, что $\operatorname{tg} \varphi_1 = 1/R_1$. Две прямые пересекутся. Утверждается, что ордината точки пересечения I_0 определит ток, текущий в цепи последовательных сопротивлений, а величины напряжений V_2 и V_1 будут равны падениям напряжений на сопротивлениях R_2 и R_1 соответственно.

Легко убедиться, что так оно и есть. Должны выполняться два условия. Во-первых, сумма падений напряжений $V_1 + V_2$ должна равняться V_0 . Что это так, видно непосредственно из рисунка. Во-вторых, падения напряжений на сопротивлениях должны быть пропорциональны величинам сопротивлений. И в этом легко убедиться, так как

$$\begin{aligned} V_1 &= I_0 / \operatorname{tg} \varphi_1 = I_0 R_1; & V_2 &= I_0 / \operatorname{tg} \varphi_2 = I_0 R_2; \\ V_1 / V_2 &= R_1 / R_2. \end{aligned} \quad (40)$$

Решим теперь задачу посложнее. Рассчитаем цепь, показанную на вставке к рис. 38, б: к последовательно соединенным сопротивлению и термистору*) приложено напряжение V_0 . Требуется определить ток в цепи и падение напряжения на каждом из элементов схемы.

Поступим точно так же, как и раньше. Построим вольтамперную характеристику термистора. Отложим на оси напряжений величину V_0 . Проведем прямую линию из точки V_0 под углом φ к оси напряжений. При этом $\operatorname{tg} \varphi = 1/R$. Точка пересечения этой прямой с вольтамперной характеристикой термистора определит искомый ток I_0 , падение напряжения на сопротивлении V_R и напряжение на термисторе V_T^{**}).

Не правда ли, просто? А теперь представьте себе, что этот метод графического решения вам незнаком, и задачу нужно решить аналитически. Поверьте, что чем больше вы будете осознать все те трудности, которые при этом возникнут, тем больше захочется назвать так все просто разрешающую прямую «волшебной». В научной литературе, однако, предпочитают эпитеты менее поэтические. Сопротивление R , включенное последовательно с термистором, обычно называется нагрузочным сопротивлением. Прямую, соответствующую этому сопротивлению, называют нагрузочной прямой. А сам метод графического расчета — *методом нагрузочной прямой*.

А теперь — обратимся к чудесам, которые сулит вольтамперная характеристика термистора и освоенный нами метод нагрузочной прямой.

З-транные прыжки. В предыдущем разделе, обсуждая свойства волшебной прямой, мы, как бы ненамерком, умолчали об одном деликатном обстоятельстве. Мы сделали вид, что нагрузочная прямая всегда пересекается с вольтамперной характеристикой только в одной точке.

Такая ситуация возможна, но совсем не обязательна. Нагрузочная прямая I на рис. 39, действительно,

*) Обратите внимание на то, как принято обозначать термистор. Знак нелинейного сопротивления дополняется значком t° , показывающим, что это сопротивление зависит от температуры и что его температурный коэффициент сопротивления отрицателен.

**) Докажите сами, что предложенное построение приводит к правильному результату.

будет пересекаться с вольтамперной характеристикой только в одной точке при любом приложенном напряжении V_0 . И из рис. 39 сразу видно, почему это так и когда точка пересечения только одна: когда угол φ достаточно мал или, что то же самое, когда нагрузочное сопротивление R велико.

Если нагрузочное сопротивление малó (а угол φ , соответственно, велик), нагрузочная прямая может

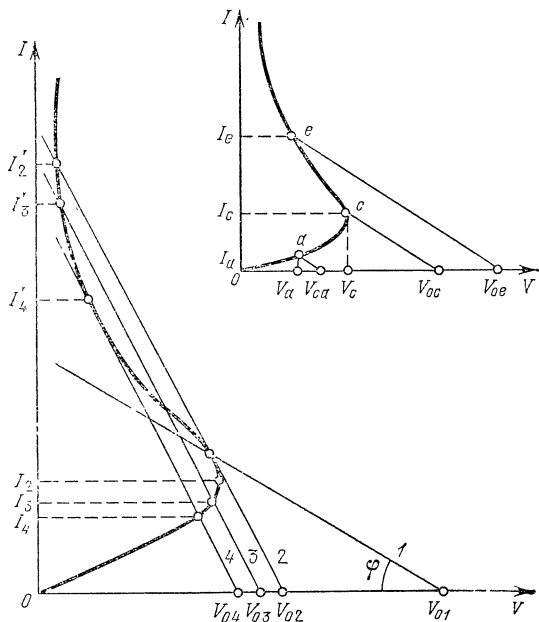


Рис. 39. Терпение. Готовьтесь к прыжкам.

пересекаться с вольтамперной характеристикой в двух (прямые 2 и 4) или даже в трех (прямая 3) точках. Что это значит?

Формальный ответ очень прост. При одном и том же напряжении батареи V_0 и нагрузочном сопротивлении R имеется не одно, а два (или три) возможных решения.

Но нас, как физиков, такой ответ удовлетворить не может. Почему это так? Какое решение реализуется на самом деле? Ведь не может же быть, чтобы через образец одновременно протекали два различных тока! А может быть, при одних условиях реа-

лизуется одно решение, а при других — другое. Тогда, что это за условия?

Обсудим подробно все эти важные и интересные проблемы.

Начнем с самого простого случая, показанного на вставке к рис. 39. Нагрузочное сопротивление R выбрано большим, так что нагрузочная прямая всегда пересекает вольтамперную характеристику термистора только в одной точке. Давайте проследим, что происходит, когда сохраняя R постоянным, мы будем увеличивать напряжение V_0 . Случай a — самое маленькое напряжение. В образце течет ток I_a , падение напряжения на термисторе равно V_a . Каждую секунду в образце выделяется энергия $P = I_a V_a$. Образец нагреет до температуры $T_a > T_0$.

При увеличении напряжения до величины V_{0c} ток возрастает до значения I_c , а напряжение на образце — до величины V_c . В этом состоянии образец нагреет до более высокой температуры: $T_c > T_a > T_0$.

Пусть мы еще больше увеличили напряжение батареи V_0 и попали на падающий участок вольтамперной характеристики. Вид этой характеристики говорит нам совершенно однозначно: на падающем участке тепловое равновесие существует только при вполне определенном условии. Если увеличился ток через образец, напряжение на образце должно уменьшиться! Куда же оно денется? Перераспределится между нагрузочным сопротивлением и образцом. Уменьшится на образце — увеличится на сопротивлении нагрузки. Прекрасно! Но сопротивление нагрузки должно быть способно принять на себя это избыточное напряжение. Попросту говоря, оно должно быть достаточно велико. Если, скажем, сопротивление нагрузки равно нулю (а нам никто не мешает рассмотреть и этот случай), то ясно, что, увеличив ток через образец на падающем участке, мы попросту не найдем, куда девать избыточное напряжение. Оно останется приложенным к образцу и... Условие теплового равновесия окажется нарушенным.

На вставке к рис. 39 нагрузочное сопротивление выбрано достаточно большим. Там все обстоит хорошо.

А вот прямая 2 на рис. 39 соответствует случаю, когда нагрузочное сопротивление недостаточно велико. При напряжении батареи $V_0 = V_{02}$ тепловое

равновесие между образцом и окружающей средой еще возможно. Но стоит еще хотя бы чуть-чуть увеличить напряжение — нагрузочное сопротивление окажется не в состоянии принять на себя то напряжение, на которое должно уменьшиться напряжение термистора, чтобы сохранить тепловое равновесие *). Что же произойдет? К термистору окажется приложенным большее напряжение, чем то, которое соответствует тепловому равновесию. Выделяющаяся в образце мощность $P = IV$ не будет успевать отводиться в окружающее пространство. Образец начнет разогреваться.

Разогрев образца приводит к увеличению в нем концентрации носителей, уменьшению сопротивления, а значит, — к росту тока. С ростом тока тепловыделение, если не уменьшить в достаточной мере падение напряжения на образце, будет еще больше возрастать. А уменьшить напряжение мы не в состоянии — величины нагрузочного сопротивления не хватает. Значит, образец будет разогреваться еще сильнее. Еще больше увеличится концентрация, еще больше возрастет ток, еще больше увеличится тепловыделение... Чем же это кончится?

Нагрузочной прямой 2 соответствует оптимистический исход. При достаточно большом токе через образец I'_2 тепловое равновесие вновь оказывается возможным! Но... при гораздо более высокой температуре образца, гораздо большем токе и значительно меньшем падении напряжения.

Легко увидеть, что чем меньше величина нагрузочного сопротивления, тем большему току I'_2 (и большей температуре образца) будет соответствовать *второе устойчивое состояние*. Если сопротивление малó, вполне может оказаться, что температура, соответствующая току I'_2 , окажется выше, чем температура разрушения образца или контактов. В этом случае, увы, — исход окажется печальным и для образца, и для инженера, проектировавшего схему,

*) Попробуйте сами определить величину нагрузочного сопротивления, при котором для данной вольтамперной характеристики S-типа всегда реализуется только одно решение. Это — трудная задача, требующая внимания, фантазии и аккуратности.

Но допустим, что схему проектировал грамотный инженер, и в точке, соответствующей току I'_2 , с образцом ничего плохого не происходит.

Подумаем, что случится, если, не меняя нагрузочного сопротивления, уменьшить напряжение V_0 до значения V_{03} ? Ничего чрезвычайного. Ток уменьшится до величины I'_3 . Но обратите внимание на одну любопытную деталь. Если бы мы, перед тем как установить напряжение $V_0 = V_{03}$, не превысили бы напряжения V_{02} , то образец, при том же $V_0 = V_{03}$, находился бы в состоянии, характеризуемом не током I'_3 , а током I_3 .

Если бы мы постепенно поднимали напряжение от нуля до значения V_{03} — мы «путешествовали» бы по восходящей ветви вольтамперной характеристики и остановились бы в точке, где $I = I_3$. Но если по пути мы превысили значение $V_0 = V_{02}$ — произошел переброс на падающую ветвь характеристики. И теперь уже напряжению $V_0 = V_{03}$ будет соответствовать состояние образца, характеризующееся током I'_3 .

Что произойдет, если продолжать снижать напряжение?

Как только значение V_0 снизится до величины V_{04} — произойдет прыжок из состояния, характеризующегося током I'_4 , в состояние, где ток равен I_4 . Образец возвратится на восходящую ветвь вольтамперной характеристики. А если сказать проще — остынет! Почему?

Ну, теперь ответ на этот вопрос никаких трудностей для нас не представляет.

Для того чтобы находясь в точке вольтамперной характеристики, где $I = I'_4$, позволить себе уменьшить ток и сохранить при этом тепловое равновесие, необходимо увеличить напряжение на образце. Это добавочное напряжение может взяться только с нагрузочного сопротивления R , если... сопротивление достаточно велико. В противном случае при снижении тока на величину ΔI увеличение напряжения на образце ΔV окажется меньшим, чем нужно. Выделяющаяся в образце мощность $P = IV$ не будет успевать компенсировать уход тепла в окружающее пространство. Образец начнет остывать. Остывание приведет к увеличению сопротивления образца, а значит, к дальнейшему уменьшению тока. Следовательно, выделяющаяся мощность окажется еще меньше. Образец будет

остывать, пока не окажется возможным новое состояние равновесия — в точке, где $I = I_4$.

Итак, при достаточно малом сопротивлении нагрузки, в цепи, содержащей элемент с S-образной вольтамперной характеристикой, могут осуществляться переходы (прыжки) с одной ветви вольтамперной

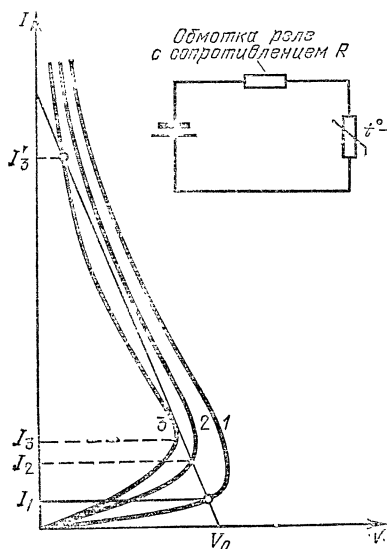


Рис. 40. При комнатной температуре вольтамперная характеристика термистора имеет вид кривой 1. По обмотке реле течет небольшой ток I_1 . Температура повышается. Вольтамперная характеристика начинает изменяться (кривая 2). Температуре, о которой должен сигнализировать датчик, соответствует вольтамперная характеристика, имеющая вид кривой 3.

меняться (см. вставку к рис. 37). Ток в цепи обмотки реле растет. Он становится равным I_2 , затем, при еще более высокой температуре, — I_3 . Однако даже ток I_3 еще не приводит к срабатыванию реле (рис. 40). Но вот температура повысилась еще чуть-чуть. Ток резко возрастает до величины $\approx I'_3$. Реле замыкает контакты — ревет пожарная сирена,

характеристики на другую. Причем (и это очень важно!) небольшие изменения параметров могут привести к очень большим изменениям тока.

Сбор урожая. Дело известное: «Корни учения горьки, плоды — сладки». Нам пора собирать урожай.

Представим себе, что термистор с вольтамперной характеристикой в виде кривой 1 (рис. 40) питается от источника напряжения V_0 . Нагрузочным сопротивлением служит обмотка сигнального реле. Ток I_1 мал и не вызывает срабатывания реле. Пусть эта миниатюрная схема установлена в каком-то помещении. ...Возникает пожар. Температура окружающего воздуха повышается. Вольтамперная характеристика термистора начинает

Точно такая же схема может быть использована для дистанционного контроля тепловых режимов машин и механизмов, для предупреждения перегрева зерна в элеваторах, для срочного вызова врача при опасном повышении температуры больного и во многих других случаях.

Схема, показанная на рис. 40, может быть использована и для автоматического регулирования уровня жидкости в сосуде. Жидкости, как правило, отводят тепло гораздо лучше, чем воздух. Поэтому вольтамперные характеристики термистора, опущенного в жидкость и находящегося в воздухе, будут различны, даже если температура воздуха и жидкости одна и та же. Действительно, при одном и том же токе I сопротивление образца будет тем выше, чем лучше отводится от него тепло. Поэтому в жидкости тому же току I будет отвечать большее падение напряжения на термисторе, чем в воздухе. Вольтамперная характеристика термистора, погруженного в жидкость, сдвинется направо по оси напряжений по отношению к соответствующей характеристике в воздухе.

Будем считать, что кривая I на рис. 40 соответствует вольтамперной характеристике термистора в воде, а кривая 3 — в воздухе. Установим датчик в бачке с водой на желаемом уровне. Пока термистор находится в воде, ему соответствует вольтамперная характеристика I , по обмотке реле течет небольшой ток I_1 , контакты реле разомкнуты и вода в бак не наливается. Стоит воде опуститься ниже заданного уровня, как термистор начнет нагреваться протекающим через него током (тепло в воздухе отводится хуже!), его вольтамперная характеристика начнет сдвигаться влево. Затем — знакомый прыжок вверх, на падающий участок характеристики. Реле замыкает контакты, срабатывает исполнительный механизм и в бак начинает поступать вода.

Как только вода снова поднимется до уровня датчика, термистор, оказавшись в воде, начнет остывать. Его вольтамперная характеристика начнет сдвигаться вправо. В определенный момент произойдет прыжок с падающей на восходящую ветвь вольтамперной характеристики. (Взгляните на рис. 39: прыжку, о котором мы говорим, соответствует переход вдоль нагрузочной характеристики от тока I_4' к току I_4 .)

Контакты реле размыкаются — вода в бак поступать перестает.

Дополнив схему, содержащую термистор, конденсатором, можно сконструировать реле времени, генераторы импульсов, модуляторы и множество других схем. Мы не будем их рассматривать: поняв физические явления, лежащие в основе прыжков, и научившись использовать метод нагрузочной прямой, вы вполне можете разобраться в их работе самостоятельно. Необходимые ссылки на статьи и книги, где описаны такие схемы, можно найти в монографии И. Т. Шефтеля *).

Следует упомянуть еще об одном важном обстоятельстве. *S*-образная вольтамперная характеристика возникает во многих полупроводниковых приборах, широко используемых в современной электронике. Методы анализа таких приборов, усвоенные на примере простейшего полупроводникового устройства — термистора, оказываются полезными и эффективными при изучении работы любого прибора с *S*-образной вольтамперной характеристикой.

Глава 7

ТЕНЗОРЕЗИСТОРЫ

Известна характерная особенность людей: до того, как будет сделано какое-либо замечательное открытие, они удивляются, как это оно может быть сделано, а после того, как оно сделано, удивляются, как это оно не было открыто раньше.

Ян Коменский

Тензорезистор — это полупроводниковый резистор, у которого величина сопротивления зависит от механического усилия, сжимающего или растягивающего резистор. Корень «тензо», входящий в название прибора, происходит от латинского *tendo* — растягиваю.

В отличие от изученного нами в предыдущей главе термистора, имеющего десятки самых разнообразных применений, тензорезисторы применяются почти исключительно для измерения деформаций и

*) Шефтель И. Т. Терморезисторы. — М.: Наука, 1973.

механических напряжений. Как и подобает профессионалу, тензорезистор способен измерять эти величины в очень широком диапазоне значений и в самых разнообразных условиях. С помощью тензорезисторов можно измерять относительное изменение длины, лежащее в пределах от стомиллионной доли процента ($\sim 10^{-10}$) до величины $\sim 5 \cdot 10^{-3}$. (При относительном удлинении $\sim 5 \cdot 10^{-3}$, или 0,5 %, разрушается подавляющее большинство металлов и сплавов.) Тензорезисторы способны измерять деформации, возникающие за десятиллионные доли секунды, — и следить за медленно, в течение часов, суток, месяцев накапливающимися деформациями в стальных и железобетонных конструкциях. С помощью тензорезисторов деформации могут быть измерены на самом лютom морозе, когда-либо зарегистрированном на планете Земля ($\sim -100^\circ\text{C}$), и при температурах от $+300$ до 400°C , нередко возникающих при работе различных машин и механизмов, проведении многих технологических процессов.

Зачем нужно измерять деформации

Деформации (от латинского *deformatio* — изменение формы), с которыми приходится встречаться в физике, технике, биологии, лежат в пределах от смещений, соответствующих размеру атома, до изменений, легко замечаемых невооруженным глазом.

Наш мозг способен фиксировать и ощущать как звук смещение чувствительных элементов уха, соответствующее размеру атома.

С другой стороны, деформация мышц, управляющих нашим дыханием, движением рук и ног, составляет сантиметры и десятки сантиметров.

В технических устройствах деформация в 1 мкм (10^{-6} м) иногда делает невозможной работу прибора. С другой стороны, верхняя часть высотных зданий может колебаться под порывами ветра с амплитудой порядка метра — и ни к каким опасным последствиям это обстоятельство не приводит.

В физике и технике значительно более важную роль, чем абсолютная, играет относительная деформация ϵ , т. е. отношение изменения размера к его первоначальному значению. Если речь идет, например,

об изменении длины, то

$$\varepsilon = (l - l_0)/l_0 = \Delta l/l_0, \quad (41)$$

где l_0 — начальная длина, l — конечная.

Понятно, почему относительная деформация так важна. В соответствии с законом Гука (Физика, 9 класс) относительная деформация пропорциональна механическому напряжению, испытываемому конструкцией, деталью, определенной частью детали. Если относительная деформация ε превысит допустимое (определенное для каждого материала) значение, возможны серьезные неприятности. Хрупкие материалы трескаются и разрушаются, пластичные — начинают течь — необратимо менять форму, и в конце концов также разрушаются.

Создается новая модель самолета. По результатам расчета конструкторов в металле воплощаются корпус, шасси, крылья, турбины, винты, хвостовое оперение — десятки и сотни блоков, из которых состоит современный воздушный лайнер или сверхзвуковой истребитель. Как измерить деформации и напряжения, возникающие в различных элементах при взлете, посадке, маневре, порыве ветра? Какие участки крыла или фюзеляжа самые «напряженные», — где относительные деформации достигают наибольшего, может быть, опасно большого значения? Как следует изменить форму и конструкцию детали, чтобы снять избыточные напряжения и деформации? Запустить самолет в серийное производство, не ответив на эти вопросы, значит взять на себя ответственность за возможную гибель людей.

Турбина современной ГЭС представляет собой исполинский цилиндр массой в сотни тонн, на боковой поверхности которого расположены выступы очень сложной формы — лопатки. Мощная струя воды, бьющая по лопаткам, вращает турбину и тем самым ротор генератора электростанции. Выход из строя одной-единственной лопатки в лучшем случае означает остановку турбины, прекращение подачи энергии потребителям, необходимость извлекать из шахты машинного зала огромный агрегат и ремонтировать его. В худшем случае сломанная лопатка может привести к тяжелой аварии, разрушениям, человеческим жертвам. Какие нагрузки испытывает лопатка? В каких сечениях деформации и напряже-

ния особенно велики? Может быть, где-то нужно увеличить толщину, а где-то изменить форму?

Выдержит ли мост одновременный натиск паводка и ураганных весенних ветров?

Какие деформации возникают в конструкции при обычной, повышенной, экстремальной нагрузке?

Без ответа на эти вопросы невозможно проектирование, создание, испытание современных технических сооружений.

Как можно измерить деформацию

Исторически самый первый способ определения величины деформации состоял в том, что изменения в размерах тела, возникающие под влиянием нагрузки, измерялись с помощью какого-либо, по возможности более точного, механического приспособления. Например, линейки, штангель-циркуль, микрометра. Этот самый простой способ измерения деформаций имеет множество недостатков. Точность его невелика. Его можно использовать только для исследования статических нагрузок на неподвижную деталь. Такой способ абсолютно непригоден для исследования труднодоступных деталей, находящихся внутри механизма. Наконец (и это — очень важный недостаток!) такой метод дает возможность определить только изменение размеров всего тела, т. е. среднюю деформацию. А ведь в деталях сложной формы напряжения и деформации на каком-либо небольшом участке могут иногда в десятки раз превосходить средние значения.

Механические тензометры. Механические тензометры или, как их иногда называют, экстензометры позволяют увеличить точность измерения деформаций и производить измерения не только на всей детали целиком, но и на отдельных ее участках.

Точность измерения повышается за счет того, что небольшая деформация исследуемой детали усиливается с помощью системы рычагов. В сущности, довольно совершенным экстензометром являются обыкновенные магазинные весы. Небольшое (порядка нескольких миллиметров) смещение чашки весов под действием груза преобразуется системой рычагов в поворот стрелки, конец которой смещается на десятки сантиметров. Наиболее совершенные экстен-

зометры позволяют измерять относительные деформации $\sim 10^{-5}$ и обладают коэффициентом усиления $\sim (2 \div 3) \cdot 10^3$. Расстояние между точками, в которых экстензометр крепится к испытуемой детали, так называемая база, составляет для механических тензометров величину порядка сантиметра. Измерения с помощью экстензометра позволяют определить средние напряжения на участке, равном длине базы.

Оптические тензометры. Эти тензометры позволяют уменьшить базу и еще больше увеличить чувствительность. В оптических тензометрах система механических рычагов, усиливающая небольшие смещения, заменена системой зеркал. В качестве индикатора смещения (деформации) используется луч света.

Хорошие оптические тензометры позволяют измерять относительные деформации $\sim 10^{-6}$ при базе $\sim 5-6$ мм. Благодаря тому, что отклоняющие свет зеркальца могут быть выполнены очень легкими, оптический тензометр может применяться для измерения переменных во времени деформаций вплоть до частот 100—150 Гц. Коэффициент усиления самых лучших оптических тензометров равняется $\sim 30\,000$, а база составляет всего 1,5—2 мм. Однако эти дорогие, сложные, уникальные приборы могут быть использованы только в лабораторных условиях.

Электрические тензометры. Электрические тензометры используют изменения электрических параметров, возникающие в тензодатчике под влиянием деформации исследуемой детали.

Таковыми параметрами могут быть емкость, индуктивность и сопротивление.

Емкостный тензодатчик. Такой тензодатчик представляет собой конденсатор, одна пластина которого неподвижна, а вторая, подвижная, прикреплена к исследуемой детали. Деформация детали приводит к изменению емкости конденсатора. Тензодатчик служит элементом колебательного контура генератора (Физика, 10 класс), и изменение емкости тензодатчика меняет частоту генерируемых колебаний. По изменению частоты генерации можно судить о величине деформации в месте прикрепления датчика.

Индуктивный тензодатчик. Индуктивный тензодатчик представляет собой катушку с подвижным железным сердечником. Сердечник (или катушка)

закрепляется на исследуемой детали. При деформации сердечник вдвигается (или выдвигается), и таким образом изменяется индуктивность.

Тензометры, использующие электрические тензодатчики, обладают очень важным преимуществом перед любыми другими тензометрами. С их помощью измерения можно проводить дистанционно, находясь практически на любом расстоянии от исследуемого объекта.

Однако емкостные и индуктивные тензодатчики обладают и очень существенными недостатками. Емкостные тензодатчики очень чувствительны к вибрациям, их трудно крепить к исследуемой детали, а исследование деталей сложной формы часто оказывается просто невозможным. Индуктивным тензодатчикам присущи все отмеченные недостатки емкостных, а кроме того, они имеют относительно большой вес и значительные габариты.

Настоящей революцией в тензометрии явилось изобретение тензодатчиков сопротивления.

Проволочные тензодатчики сопротивления

Проволочные тензодатчики сопротивления были разработаны независимо и практически одновременно в 1938 г. двумя американскими инженерами — Симмонсом и Рюджем.

Принцип работы тензодатчика настолько прост, что при первом знакомстве с изобретением кажется просто невероятным, что оно было сделано так поздно. Тем более невероятным, что уже через несколько лет после появления первых образцов проволочные тензодатчики выпускались в сотнях тысяч экземпляров и нашли самое широкое практическое применение.

Тоненькая проволока приклеивается к испытуемой детали. Если данный участок детали под нагрузкой сжимается, проволока сжимается тоже. Ее длина l уменьшается, а поперечное сечение S слегка увеличивается. Омическое сопротивление проволоки $R = \rho l / S$ уменьшается. При растяжении, напротив, длина проволоки растет, а поперечное сечение уменьшается. В результате сопротивление проволоки увеличивается. По изменению

сопротивления можно судить о величине и знаке деформации.

Сопротивление проволоки меняется при деформации практически мгновенно. Поэтому проволоочные тензодатчики можно использовать для измерения очень быстро меняющихся деформаций.

Проволоку легко прикрепить (наклеить) к детали любой формы.

Проволоочный тензодатчик по размерам не превосходит почтовую марку, и масса его составляет доли грамма (рис. 41). При необходимости такой датчик

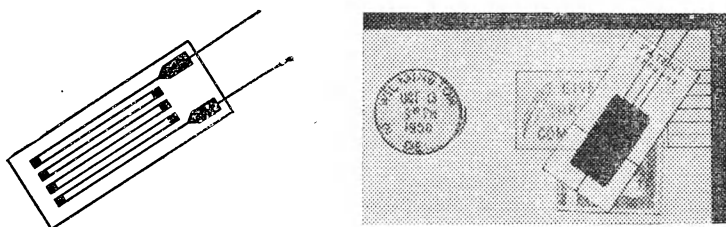


Рис. 41. Проволоочный тензодатчик сопротивления. Датчик выполнен из нескольких витков проволоки с тем, чтобы увеличить сопротивление и чувствительность.

можно изготовить размером в доли миллиметра. Датчик дешев, прост в изготовлении, удобен, надежен и стабилен.

О том, что сопротивление металлической проволоки при растяжении возрастает, еще в 1856 г. сообщил в докладе, сделанном на заседании Лондонского Королевского общества, Уильям Томсон (лорд Кельвин) — создатель абсолютной шкалы температур.

Почему же от открытия Уильяма Томсона до изобретения Симмонса и Рюджа прошло так много времени?

Одна из причин состоит в том, что изменение сопротивления проволоочного датчика при любых реальных деформациях не превосходит десятых долей процента, а в типичном случае может составлять сотые и даже тысячные доли процента. Обнаружить, а тем более точно измерить такое изменение сопротивления не просто.

Получим выражение, связывающее величину относительной деформации проволоки $\varepsilon = \Delta l / l_0$ с относительным изменением ее сопротивления $\Delta R / R_0$.

Исходное сопротивление проволоки равняется

$$R_0 = \rho \frac{l_0}{S_0}. \quad (42)$$

Предположим сначала, что при деформации (растяжении или сжатии) меняется только длина проволоки l , а удельное сопротивление ρ и площадь поперечного сечения S не меняются. Тогда изменение сопротивления проволоки ΔR_l при изменении длины на Δl составит:

$$\Delta R_l = \frac{dR_0}{dl_0} \Delta l = \frac{\rho}{S_0} \Delta l. \quad (43)$$

А относительное изменение $\frac{\Delta R_l}{R_0}$ будет равняться, очевидно,

$$\frac{\Delta R_l}{R_0} = \frac{\Delta l}{l_0} = \varepsilon. \quad (44)$$

В действительности, однако, при деформации изменяется и площадь поперечного сечения S . Если теперь предположить, что длина и удельное сопротивление проволоки остаются при деформации постоянными, то изменение поперечного сечения на величину ΔS приведет к изменению сопротивления ΔR , равному

$$\Delta R_S = \frac{dR_0}{dS_0} \Delta S = -\rho \frac{l_0}{S_0^2} \Delta S. \quad (45)$$

(Знак минус в формуле (45) напоминает о том, что при сжатии площадь увеличивается, а величина R — уменьшается.)

Относительное изменение сопротивления за счет изменения площади составит

$$\frac{\Delta R_S}{R_0} = -\frac{\Delta S}{S_0}. \quad (46)$$

В действительности при деформации меняются, конечно, и длина, и поперечное сечение проволоки. Полное изменение сопротивления будет равняться поэтому

$$\frac{\Delta R}{R_0} = \frac{\Delta R_l + \Delta R_S}{R_0} = \frac{\Delta l}{l_0} - \frac{\Delta S}{S_0}. \quad (47)$$

Пользуясь выражением (47), нужно помнить, что знаки величин Δl и ΔS — противоположны. При сжатии, например, l уменьшается, а S — увеличивается.

Теперь следует вспомнить об одном важном и интересном обстоятельстве. Еще в начале прошлого века замечательный французский физик и математик Симеон Дени Пуассон (1781—1840) установил, что при упругих деформациях, пока выполняется закон Гука, относительное изменение диаметра проволоки $\Delta d/d_0$ пропорционально относительному изменению ее длины:

$$\frac{\Delta d}{d_0} = -\nu \frac{\Delta l}{l_0} = -\nu \varepsilon. \quad (48)$$

Коэффициент ν называется *коэффициентом Пуассона*. Его величина не зависит ни от знака деформации (растяжение или сжатие), ни от величины деформации в пределах упругости. Величина коэффициента Пуассона для металлов лежит в пределах от 0,25 до 0,4.

Легко доказать, что при относительном изменении диаметра проволоки $\Delta d/d_0$ относительное изменение площади ее поперечного сечения составит $\Delta S/S_0 \approx 2\Delta d/d_0$ *).

Таким образом, связь между относительным изменением сопротивления и относительной деформацией окончательно запишется в виде:

$$\frac{\Delta R}{R_0} = (1 + 2\nu) \frac{\Delta l}{l_0} = (1 + 2\nu) \varepsilon. \quad (49)$$

Одним из важнейших параметров тензодатчика является *коэффициент тензочувствительности* F , определяемый как отношение $\Delta R/R_0$ к относительному изменению длины при деформации ε . Выведенная нами формула (49) показывает, что для металлических тензодатчиков сопротивления коэффициент тензочувствительности F равняется

$$F = \frac{\Delta R}{R_0 \varepsilon} = 1 + 2\nu \quad (50)$$

и лежит в пределах $\sim 1,5—1,8$ **).

*) Сделайте это самостоятельно.

**) В действительности, как показал еще в 1881 г. известный русский физик и педагог Орест Данилович Хвольсон, при деформации может меняться и удельное сопротивление металлов ρ . Однако для всех практически используемых при изготовлении проволочных тензодатчиков металлов это изменение очень мало.

Максимальные деформации ϵ , ставящие металлические конструкции на грань разрушения, составляют, как уже упоминалось, $\sim (0,2-0,5)\%$. Таким образом, самые большие изменения сопротивления проволочных тензодатчиков не превышают величины $\Delta R/R_0 = F\epsilon \sim 5 \cdot 10^{-3}$. А во многих практически важных случаях величина $\Delta R/R_0$ составляет $\sim 10^{-4}-10^{-5}$.

Если речь идет об изменении сопротивления, которое быстро меняется во времени, как бывает, когда тензодатчик измеряет динамические нагрузки в быстро вращающихся деталях, в узлах самолета, турбины, вибрационных установках,— для того чтобы точно измерить такое небольшое изменение сопротивления, нужны чувствительные усилители. Эти усилители, так же как и тензодатчики, должны быть надежными, стабильными, сравнительно дешевыми и доступными. А такие усилители (на электронных лампах) появились только в 30-е годы нашего века. В этом, по-видимому, и состоит причина столь позднего открытия Симмонса и Рюджа.

В течение почти 20 лет проволочные тензодатчики сопротивления верой и правдой служили науке и технике. Однако в середине 50-х годов в технике тензометрии произошла «вторая промышленная революция», связанная с появлением тензорезисторов.

Полупроводниковые тензодатчики сопротивления — тензорезисторы

В 1954 г. в одном из ведущих физических журналов мира *Physical Review* была опубликована статья американского исследователя Чарльза Смита. В этой работе Смит описал результаты своих экспериментов по исследованию влияния деформации на сопротивление кристаллов германия и кремния. Результаты были поразительными.

Коэффициент тензочувствительности F в некоторых случаях превышал 100—150! Полупроводники в очередной раз продемонстрировали свои таланты «принцев на горошине»: чувствительность сопротивления к деформации в 100 раз превышала типичную тензочувствительность металлов.

Список сюрпризов, однако, далеко не ограничился только большой величиной тензочувствительности.

Величина и даже знак тензочувствительности зависели от того, *n*- или *p*-типа материал подвергался деформации, от уровня легирования материала донорными или акцепторными примесями.

Наконец, тензочувствительность германия и кремния оказалась резко анизотропной величиной. Ее значение сильно зависело от того, в каком кристаллографическом направлении вырезались образцы из монокристаллов германия и кремния.

Что изменяется при деформации кристалла. Громадная величина тензочувствительности, обнаруженная Смитом, возможность существования и положительной и отрицательной тензочувствительности*), резкая анизотропия — все эти факты ясно показывают, что причина изменения сопротивления германиевых и кремниевых образцов при деформации совсем иная, чем в металлах.

В металлах главную роль в изменении сопротивления играет изменение геометрических размеров при деформации. Как мы видели, этот механизм тензочувствительности приводит к значениям коэффициента $F \sim 1-2$. Значение $F \gtrsim 100$ изменением геометрических размеров не объяснишь. К тому же, изменение геометрических размеров всегда приводит к положительному значению F : образец растягивается — длина увеличивается — сопротивление растёт. Величины Δl и ΔR имеют одинаковые знаки. В германии и кремнии наблюдаются значения $F \sim -100$! Сопротивление образцов при растяжении сильно уменьшается, а при сжатии — возрастает.

Значит? Значит (см. формулу (42)), при деформации полупроводниковых кристаллов сильно меняется удельное сопротивление материала ρ .

Величина ρ определяется концентрацией свободных носителей в материале и их подвижностью. Какой из этих параметров изменяется при деформации?

Вообще говоря, изменяться могут и концентрация, и подвижность.

При деформации меняется расстояние между атомами кристаллической решетки материала, а сле-

*) В соответствии с формулой (50), коэффициент тензочувствительности положителен, когда при растяжении сопротивление растёт, а при сжатии — падает (ΔR и Δl имеют одинаковые знаки).

довательно, и силы взаимодействия между атомами. Это приводит к изменению энергии образования электронно-дырочной пары в материале $\mathcal{E}_g$. В примесном полупроводнике деформация приводит к изменению расстояния между атомом примеси и окружающими его атомами кристаллической решетки. При этом меняется энергия ионизации примеси $\Delta\mathcal{E}$. Поэтому как в собственном, так и в примесном полупроводнике концентрация свободных носителей, вообще говоря, зависит от деформации.

Однако в условиях экспериментов, проделанных Смитом, изменение концентрации свободных носителей, как писал сам исследователь, «хотя и могло быть зарегистрировано, но было совершенно несущественным».

Основную роль в наблюдаемых эффектах играло изменение подвижности свободных носителей, электронов и дырок, обусловленное зависимостью от деформации эффективной массы носителей m^* .

Интересно оценить, какому изменению эффективной массы m^* соответствуют наблюдавшиеся Смитом эффекты. При растяжении кремний разрушается, когда величина относительной деформации ε составляет $\sim (2 \div 4) \cdot 10^{-3}$. В соответствии с формулой (50) при коэффициенте тензочувствительности $F \sim 100$ относительное изменение $\Delta R/R_0$, а следовательно, и относительное изменение эффективной массы носителей m^* составит при этом $\sim 20\text{—}40\%$.

Главное свойство кристаллов. При анализе тензочувствительности первостепенную роль играет кристаллическое строение полупроводниковых материалов.

Главное свойство кристаллов, как известно,—анизотропия, т. е. зависимость физических свойств от направления в кристалле (Физика, 9 класс).

В опытах Смита анизотропия обнаруживается особенно наглядно. В зависимости от того, вдоль какого кристаллографического направления приложено усилие, деформирующее кристалл, коэффициент тензочувствительности для кремниевых образцов с абсолютно одинаковой концентрацией легирующей примеси при одной и той же температуре и т. д. может лежать в пределах от $+200$ до -100 .

Термин «кристаллографическое направление» поясняется рис. 42. На этом рисунке показана эле-

ментарная ячейка кристалла кремния. Монокристалл кремния, как бы велик он ни был (рис. 30), всегда состоит из многочисленных крошечных, размером всего $\sim 5 \cdot 10^{-10}$ м, абсолютно одинаковых элементарных ячеек. Хотя ячейка называется элементарной, видно, что пространственное расположение атомов в ячейке довольно сложно. Восемь атомов (условно перенумерованных на рис. 42 цифрами

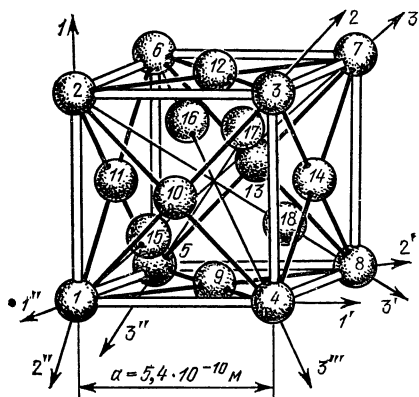


Рис. 42. Элементарная ячейка кристалла кремния. Такой же вид имеет элементарная ячейка германия, арсенида галлия, алмаза, одного из полупроводников.

1—8) образуют куб. В центре каждой из шести граней куба также находится атом (9—14). И наконец, внутри элементарной ячейки, на пространственных диагоналях куба, связывающих атомы 1 и 7, 2 и 8, 3 и 5, 4 и 6, расположены еще четыре атома (15—18).

Представим себе, что героиня книги Льюиса Кэрролла «Алиса в стране чудес», которая однажды ухитрилась так уменьшиться в размерах, что пролезла в кроличью нору,

сумела стать еще в сто миллионов раз меньше. Приобретая атомные размеры, Алиса отправилась путешествовать по монокристаллу кремния.

Путешественница может избрать различные маршруты.

Она может, например, направиться от атома 1 к атому 2 и продолжить свое путешествие в этом направлении сколь угодно далеко (маршрут 1). Вследствие строгой периодичности кристаллической решетки, ландшафты, которые представятся взгляду путешественницы, будут повторяться снова и снова. Атомы кремния будут встречаться Алисе через расстояния, в точности равные a — периоду решетки. Если в дополнение к волшебному росту Алиса об-

ладала бы и волшебным зрением, позволяющим увидеть электроны, вращающиеся на своих орбитах, то и конфигурация электронных орбит, связывающих атомы, также была бы совершенно одинаковой и повторялась бы через расстояния, равные величине a .

Из рис. 42 ясно, что если бы Алиса выбрала вместо маршрута 1 маршрут 1', начинающийся путешествием от атома 1 к атому 4 (или маршрут 1'' — от атома 5 к атому 1), она наблюдала бы абсолютно те же ландшафты, что и на маршруте 1. Маршруты (или, научно выражаясь, кристаллографические направления) 1, 1' и 1'' полностью эквивалентны.

Кристаллографические направления 2, 2' и 2'', проходящие по диагоналям граней куба, полностью эквивалентны друг другу, но сильно отличаются от направлений 1 — 1''. В направлениях 2 — 2'' расстояния между соседними атомами равняются не a , а $a\sqrt{2}$; конфигурация электронных орбит, связывающих атомы, также будет совершенно другой.

А теперь от вымышленной путешественницы Алисы обратимся к свободным носителям заряда, электронам и дыркам, совершающим непрерывные путешествия по кристаллу. Как мы помним, эффективная масса свободных носителей m^* определяется характером периодических электрических полей атомов решетки, вдоль которых пролетает носитель (гл. 5). Мы убедились, рассматривая свойства элементарной ячейки, что структура этих полей в разных кристаллографических направлениях различна. Поэтому и значения эффективной массы электронов и дырок вдоль различных кристаллографических направлений могут быть неодинаковыми. Тем более неодинаковыми могут быть изменения эффективной массы под воздействием деформации, направленной вдоль различных кристаллографических осей кристалла. Просто посмотрев на рис. 42, можно убедиться, что, скажем, растяжение, направленное вдоль оси 1, будет деформировать кристалл совершенно иначе, чем растяжение вдоль диагоналей граней куба (оси 2 — 2'') или вдоль пространственных диагоналей (оси 3 — 3'').

Именно этой анизотропией в изменении эффективной массы носителей и объясняются наблюдавшиеся Смитом эффекты. В кремнии p -типа, скажем, при растяжении кристалла вдоль пространственной диаго-

нали куба коэффициент тензочувствительности F , в зависимости от концентрации легирующей примеси, может лежать в пределах от $\sim +80$ до $\sim +200$. А при таком же растяжении вдоль ребра куба коэффициент тензочувствительности близок к единице: эффективная масса дырок при деформации вдоль этого направления практически не меняется. В кремнии n -типа, напротив, эффективная масса электронов сильнее всего изменяется именно при деформации кристалла вдоль ребра куба. Коэффициент тензочувствительности при деформации в этом кристаллографическом направлении в зависимости от уровня легирования лежит в пределах от -60 до -90 . (При сжатии вдоль оси 1 эффективная масса электронов сильно возрастает.) При деформации вдоль пространственной диагонали куба эффективная масса электронов меняется значительно слабее. Коэффициент тензочувствительности $F \approx -5$.

От физического исследования к прибору. Несмотря на то, что с момента опубликования работы Ч. Смита прошло около 30 лет, ссылки на нее и сейчас довольно часто встречаются в литературе. Работа эта стала классической и во многом способствовала выяснению тонких деталей движения электронов и дырок в полупроводниках. Но кроме того, полученные Смитом результаты почти немедленно получили широкое практическое применение.

Громадная тензочувствительность, обнаруженная исследователем в германии и кремнии, для инженеров-тензометристов означала возможность замены проволочных датчиков сопротивления полупроводниковыми тензодатчиками с чувствительностью в 100 раз большей!

Конечно, на пути от физического исследования к практическому прибору предстояло преодолеть немало трудностей. Однако за увеличение чувствительности на два порядка стоило поборотся.

Первая очевидная трудность состояла в том, что кремний, по своим свойствам наиболее подходящий материал для изготовления тензодатчиков, хрупок. А между тем, возможность изгибать тензодатчик является одним из важнейших преимуществ проволочных тензосопротивлений. Именно это свойство делает возможным исследования деформаций в деталях сложной формы.

Чем тоньше кремниевая полоска, тем меньше радиус кольца, в который она может быть согнута, если на поверхности отсутствуют сколы и трещины. Выход заключался в том, чтобы научиться изготавливать очень тонкие полоски кремния с хорошо обработанной поверхностью. На рис. 43 показан кремниевый тензодатчик, толщина которого составляет 170 мкм. Видно, что инженеры сумели успешно справиться со сложной технологической задачей.

Вторая трудность, которую необходимо было преодолеть, заключалась в большой температурной чувствительности сопротивления полупроводников. В том самом свойстве, которое делает полупроводники такими замечательными терморезисторами!

Используя тензорезисторы, мы судим о величине деформации по изменению сопротивления. Но если, скажем, никакой деформации нет, а изменилась температура тензодатчика? Измерительный прибор зафиксирует изменение сопротивления: возникает сигнал так называемой «кажущейся деформации».

Конечно, с самого начала было ясно, что для изготовления тензорезисторов следует использовать полупроводники при температурах, соответствующих примесному истощению (гл. 3). Тогда зависимость удельного сопротивления от температуры относительно слаба и определяется только изменением с температурой подвижности полупроводника μ (гл. 5). Однако даже в области примесного истощения температурный коэффициент сопротивления α_T у полупроводников гораздо выше, чем у специальных высокостабильных металлических сплавов, использовавшихся для изготовления проволочных тензодатчиков.

Существует замечательная схема, названная именем английского физика Чарльза Уитстона, который впервые применил ее на практике. Эта очень простая схема, так называемый «мост Уитстона», уже 150 лет верой и правдой служит науке и технике. Она не толь-



Рис. 43. Один из первых серийных кремниевых тензорезисторов (1960 г.). Тензодатчик можно свернуть в кольцо радиусом 3,2 мм.

ко позволяет полностью скомпенсировать паразитный сигнал, обусловленный нагревом тензодатчика, но и увеличивает чувствительность схем с тензодатчиками.

Мы уже упоминали о мостовой схеме в гл. 6, в том параграфе, где описывалась работа болометров. Обсудим теперь свойства этой замечательной схемы более подробно.

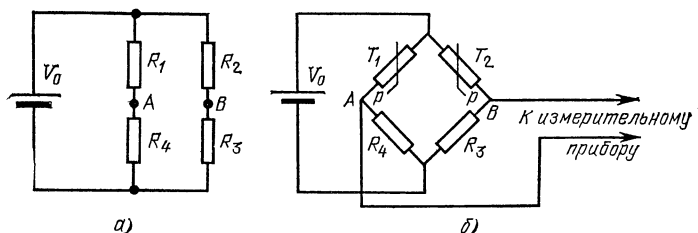


Рис. 44. Мост Уитстона.

а) С четырьмя сопротивлениями (R_1 — R_4); б) с двумя тензорезисторами (T_1 и T_2) и двумя сопротивлениями (R_3 и R_4).

Из рис. 44, а видно, что схема действительно очень проста. Она содержит всего четыре сопротивления, соединенные в две параллельные ветви по два последовательно включенных сопротивления.

Решим сначала простую задачу: рассчитаем разность потенциалов между точками А и В (рис. 44, а). Ток I_1 , текущий по ветви, содержащей сопротивления R_1 и R_4 , равен, очевидно, $I_1 = V_0 / (R_1 + R_4)$. А падение напряжения на сопротивлении R_4 равняется $V_4 = V_0 R_4 / (R_1 + R_4)$. Аналогично, падение напряжения на сопротивлении R_3 равняется $V_3 = V_0 R_3 / (R_2 + R_3)$. Следовательно, разность потенциалов между точками А и В равняется $V_4 - V_3$:

$$V_4 - V_3 = V_0 \left[\frac{R_2 R_4 - R_1 R_3}{(R_1 + R_4)(R_2 + R_3)} \right].$$

Если выполняется условие

$$R_2 R_4 = R_1 R_3, \quad (51)$$

разность потенциалов между точками А и В равняется нулю при любом напряжении источника питания моста V_0 и любых значениях сопротивлений R_1 , R_2 , R_3 , R_4 . При выполнении условия (51) мост Уитстона сбалансирован.

На рис. 44, б изображена часто применяемая на практике схема моста Уитстона, содержащая два тензорезистора *). Один из тензорезисторов, как обычно, крепится к исследуемой детали (чаще всего приклеивается специальным клеем). Второй тензорезистор (его часто называют «холостым») размещается в непосредственной близости от первого, рабочего тензорезистора. Но к испытываемой детали не крепится. Или крепится так, чтобы деформации детали не приводили к деформации вспомогательного тензорезистора (например, через эластичную прокладку). Тензодатчики T_1 и T_2 выбирают по возможности совершенно идентичными — современная технология позволяет изготавливать тензосопротивления с очень хорошо воспроизводимыми параметрами. Таким образом, не только исходные сопротивления тензодатчиков R_1 и R_2 равны, но и любые изменения в сопротивлении, вызванные нагревом или, например, освещением, также одинаковы. Величины сопротивлений R_3 и R_4 также выбираются равными. В исходном состоянии мост сбалансирован.

И тогда ... Как бы сильно ни менялись сопротивления R_1 и R_2 тензодатчиков под влиянием изменения температуры или освещения, величины R_1 и R_2 остаются равными друг другу. А следовательно, условие (51) продолжает выполняться. Мост остается сбалансированным и на вход измерительного прибора напряжение не поступает. Разность потенциалов между точками A и B появится только в том случае, если деталь, а вместе с ней и рабочий тензорезистор подвергнутся деформации.

Использование вспомогательного тензорезистора позволяет скомпенсировать мешающие сигналы кажущейся деформации.

Объединенными усилиями физиков, технологов и инженеров трудности, препятствовавшие широкому применению тензорезисторов, были преодолены довольно быстро. В 1957 г. появились первые лабораторные образцы, а с 1959 г. началось массовое применение тензорезисторов.

*) Обратите внимание на обозначение тензорезистора в схеме. Знак нелинейного сопротивления дополняется буквой p , показывающей, что сопротивление зависит от давления. Сравните обозначения тензорезистора и термистора (рис. 38, 40).

... **Что умеют тензорезисторы.** Прежде всего — измерять деформации и напряжения в деталях машин и строительных конструкциях. Без миниатюрных — длиной в миллиметры и доли миллиметра — тензорезисторов (рис. 45) нельзя обойтись при создании исполинских плотин и противоселевых заграждений, весящих десятки миллионов тонн, корпусов судов, выдерживающих двенадцатибалльный шторм и корпусов самолетов, подвергающихся обледенению и сумасшедшим порывам ветра, рессор многотонных КАМАЗов и

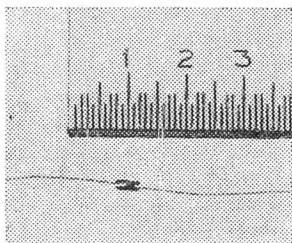


Рис. 45. Миниатюрный кремниевый тензодатчик. Расстояние между контактами 1,6 мм.

антисейсмических фундаментов высотных домов. При разработке и конструировании всех этих и сотен других конструкций тензорезисторы позволяют измерить напряжения, возникающие в узлах и деталях конструкции, выбрать оптимальную форму, материал и необходимый запас прочности.

Помимо такого прямого использования — непосредственно «по специальности», тензорезисторы используют

также для измерения некоторых других важных параметров — физических и биологических.

К числу наиболее важных приборов на основе тензорезисторов относятся акселерометры, датчики давления и приборы медицинского и биологического контроля.

Акселерометры — приборы, измеряющие ускорения, являются важнейшими элементами автопилотов, систем наведения ракет и ряда других устройств. Акселерометр на основе тензорезистора представляет собой легкую, миниатюрную (~ 1 см) деформируемую балку, на конце которой закреплен груз точно известной массы m . Сила f , деформирующая балку, пропорциональна ускорению a ($f = ma$), а показания тензодатчика, как мы знаем, пропорциональны деформирующей силе f . Поэтому показания тензодатчика можно проградуировать прямо в единицах ускорения. В зависимости от конструкции чувствительность акселерометра может лежать в пределах от $\sim 10^{-2}$ до

($\sim 10^2$ Ом/г. А полная шкала прибора — составлять (10—1000)г^{*}).

Устройство датчиков давления, в принципе, очень просто. Тензорезистор крепится к легкой диафрагме, деформация которой, измеряемая тензодатчиком, пропорциональна давлению. Миниатюрные, устойчивые к действию вибрации и других мешающих факторов, такие датчики оказываются очень удобными при работе на испытательных стендах, в вакуумных установках, при измерении давления в потоке жидкости или газа.

Объясняя Фаусту, почему он просит подписать договор о продаже души дьяволу капелькой крови, Мефистофель восклицал: «Кровь — сок совсем особенного свойства!» Тем не менее, тензорезисторы успешно используются и для определения давления крови.

Катетеры с тензорезисторами вводятся непосредственно в кровеносные сосуды. Они настолько миниатюрны, что практически не вносят возмущения в ток крови.

Тензорезисторы используются для контроля объема и скорости дыхания в системах телеметрической информации о земных и космических пациентах, для слежения за ритмом сердечных сокращений и давлением непосредственно в сосудах сердца, для контроля усилий, развиваемых в экстремальных условиях спортсменами и космонавтами, для проведения многих других биологических исследований.

Глава 8

ФОТОРЕЗИСТОРЫ.

ВНУТРЕННИЙ ФОТОЭФФЕКТ

Истина — в оттенках.

Фоторезистор — это полупроводниковый образец с омическими контактами, сопротивление которого меняется под действием света.

*) Непривычная единица измерения (г) довольно часто применяется в технике, когда речь идет об измерении ускорений. Этой буквой обозначается ускорение свободного падения $g = 9,8$ м/с². Чувствительность 10 Ом/г означает, что сопротивление тензодатчика меняется на 10 Ом при воздействии ускорения в 1 г.

Как трудно быть точным

Определение, данное фоторезистору, просто и понятно. Оно аналогично определениям других, уже знакомых нам полупроводниковых приборов,— термистора и тензорезистора. И тем не менее, внимательный читатель имеет серьезные основания остаться недовольным.

В гл. 6, знакомясь с болометрами, мы описывали их как полупроводниковые образцы, сопротивление которых меняется под действием излучения, в том числе света звезд и Солнца. В чем же разница между фотосопротивлением и болометром?

Как важно быть точным

Физические явления, лежащие в основе работы фотосопротивления и болометра, совершенно различны. Однако внешние их проявления на редкость схожи: в обоих случаях сопротивление полупроводникового образца падает под действием света.

Чтобы понять, в чем состоит разница между болометром и фотосопротивлением, сделаем простой эксперимент (рис. 46).

Пластина кремния, снабженная омическими контактами, соединена последовательно с нагрузочным сопротивлением R_n и источником напряжения V_0 . Пластина может освещаться источником света. С помощью фильтров (или спектрографа) из спектра источника света выделяется излучение с желаемой длиной волны λ . Между источником света и пластинкой размещен модулятор. При одном положении модулятора свет попадает на пластину (модулятор открыт), при другом, когда модулятор закрыт, пластина затенена.

Пусть вначале модулятор закрыт и через пластину и нагрузочное сопротивление R_n течет *темновой ток* I_t , определяемый *темновым сопротивлением* пластины. В момент $t = 0$ модулятор открывается и пластина оказывается освещенной. Сопротивление пластины падает, ток в цепи возрастает и пропорционально току возрастает падение напряжения на сопротивлении нагрузки R_n . Изменение напряжения на нагрузке регистрируется на экране электронного осциллографа.

Приготовления закончены. Приступим к экспериментам. Выберем вначале набор фильтров, пропускающий инфракрасное излучение с длиной волны $\lambda \sim 5$ мкм. Откроем модулятор и через некоторое время, в момент t_0 , закроем его. На экране осциллографа наблюдается кривая 1, показывающая, как

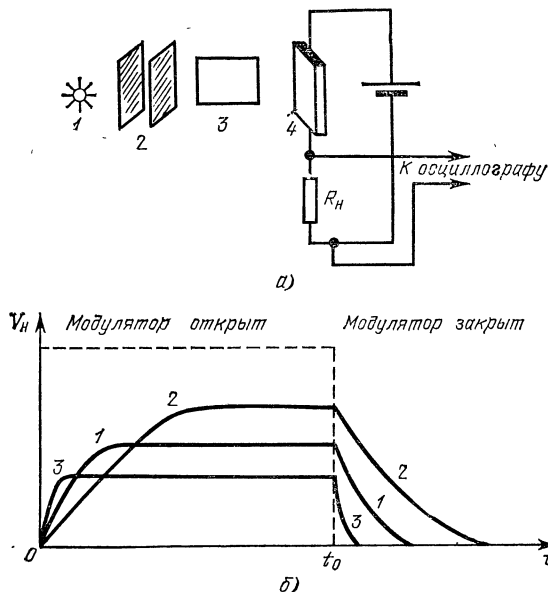


Рис. 46. Эксперимент, демонстрирующий разницу между болометрическим режимом и режимом работы фотосопротивления. а) Схема эксперимента: 1 — источник света, 2 — фильтры, 3 — модулятор, 4 — кремниевая пластинка с омическими контактами; б) зависимость от времени изменения напряжения на нагрузочном сопротивлении. Кривые 1 и 2: $\lambda \approx 5$ мкм. Кривая 3: $\lambda \approx 1$ мкм. Кривая 1 — кремниевая пластинка на воздухе. Кривая 2 — кремниевая пластинка в откачанной вакуумной колбе.

меняется со временем падение напряжения V_H на сопротивлении R_H . Эта кривая отражает изменение во времени сопротивления кремниевой пластинки. Когда модулятор открывается, сопротивление пластинки начинает уменьшаться, а ток в цепи и, следовательно, напряжение V_H — возрастать. Через некоторое время достигается стационарное состояние: сопротивление пластины остается неизменным. После того как в момент t_0 модулятор закрывается, сопротивление пластины начинает возрастать, а напряжение V_H — падать.

Теперь, ничего не меняя в схеме эксперимента, поместим кремниевую пластинку в откачанную колбу и вновь откроем модулятор. Зависимость $V_n(t)$ заметно изменится (кривая 2). Ясно, что сопротивление пластины изменяется теперь медленнее: времена нарастания и спада тока в цепи существенно увеличились. Кроме того, под действием того же самого светового потока сопротивление пластины упало сильнее, чем в случае, когда опыт проводился на воздухе.

А теперь сменим набор фильтров так, чтобы длина волны света, попадающего на образец, λ была ~ 1 мкм. Вновь сделаем те же опыты в воздухе и в вакууме. В обоих случаях на экране осциллографа наблюдается одна и та же кривая 3. Помещение кремниевой пластины в откачанную колбу при $\lambda \sim 1$ мкм абсолютно не скажется ни на временах нарастания и спада тока, ни на величине стационарного сигнала.

Обсудим теперь результаты эксперимента.

Длине волны $\lambda \sim 5$ мкм соответствует энергия фотона $\mathcal{E}_f = hc/\lambda \approx 0,25$ эВ. Эта величина меньше, чем энергия образования электронно-дырочной пары в кремнии $\mathcal{E}_g \approx 1,1$ эВ. Поэтому свет с длиной волны $\lambda \sim 5$ мкм не способен непосредственно генерировать электроны и дырки. Тем не менее, попадая в кристалл кремния, свет с такой длиной волны поглощается. Поглощение происходит на свободных носителях, имеющих в кристалле (электронах в кремнии n -типа или дырках в кремнии p -типа). Энергия фотона передается свободному электрону (или дырке), а затем, в результате столкновения электрона или дырки с решеткой, — решетке кристалла. Кристалл нагревается. За счет увеличения температуры кристаллической решетки увеличивается концентрация носителей и осциллограф в схеме, показанной на рис. 46, фиксирует увеличение тока в цепи.

Пластика кремния работает в режиме болометра: изменение проводимости обусловлено изменением температуры кристалла.

Понятно, что если изменить условия теплоотвода, изменится и тепловой режим болометра. Именно этот эффект наблюдается, когда образец помещается в откачанную вакуумную колбу. Условия для отвода тепла ухудшаются: теперь образец может отдавать избыточное тепло только за счет лучеиспускания. Ох-

лаждение за счет конвекции воздуха исключено. Поэтому, с одной стороны, установление теплового равновесия, охлаждение и нагрев кристалла происходит медленнее. С другой стороны, при том же самом световом потоке, падающем на пластинку, установившаяся температура выше, чем при работе на воздухе, и изменение сопротивления следовательно, больше. Все эти особенности ясно видны при сравнении кривых 1 и 2 (рис. 46).

Длине волны $\lambda = 1$ мкм соответствует энергия фотона $\mathcal{E}_\phi \approx 1,25$ эВ. В данном случае $\mathcal{E}_\phi > \mathcal{E}_g$. При поглощении такого кванта света в кристалле кремния возникает электронно-дырочная пара. Увеличение проводимости кристалла под действием света, обусловленное не изменением температуры кристалла, а возрастанием числа свободных носителей, возникших в результате поглощения квантов света, принято называть *фотопроводимостью*. В рассматриваемом примере образец работает в режиме фотосопротивления. Понятно, что изменение условий теплоотвода не сказывается на параметрах фотопроводимости, — на воздухе и в вакуумной колбе фотосопротивление работает одинаково.

Поглощение кванта, сопровождающееся появлением свободных носителей: электрона, дырки или электронно-дырочной пары, называется *внутренним фотоэффектом* (Физика, 10 класс).

Вот теперь мы можем дать точное определение: фоторезистор — это полупроводниковый резистор, сопротивление которого меняется под действием света за счет внутреннего фотоэффекта.

Внутренний фотоэффект

Не названный по имени, внутренний фотоэффект много раз уже появлялся в этой книге. Во Введении, где обсуждался вопрос о том, как можно разорвать электронные связи и создать электронно-дырочные пары. В гл. 4, где обдумывалась задача, как изобличить глубокие примеси в роли «убийц». В гл. 5, где рассматривалась задача о диффузии неосновных носителей.

В этом параграфе мы обсудим характеристики внутреннего фотоэффекта, определяющие работу фоторезисторов.

Коэффициент поглощения. Пусть на поверхность образца падает световой поток такой интенсивности, что непосредственно у поверхности, вблизи точки $x = 0$, на единицу площади падает N_0 квантов в секунду (рис. 47)*). По мере проникновения в глубь

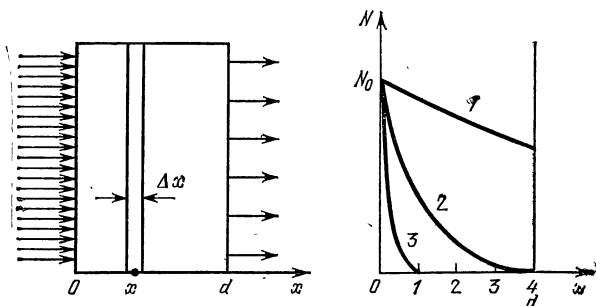


Рис. 47. К определению коэффициента поглощения α . Плотность светового потока (плотность потока фотонов) N экспоненциально спадает от поверхности в глубь полупроводника в соответствии с формулой (52). В правой части рисунка показаны три зависимости $N(x)$, соответствующие трем значениям αd : кривая 1 — $\alpha d = 0,1$; 2 — $\alpha d = 1$; 3 — $\alpha d = 5$.

полупроводника число фотонов будет уменьшаться за счет поглощения. Количественной мерой поглощения служит коэффициент поглощения α . Чем больше величина α , тем резче спадает интенсивность лучистого потока от поверхности в глубь материала.

Закон, по которому плотность фотонов убывает с координатой, имеет вид:

$$N(x) = N_0 e^{-\alpha x}. \quad (52)$$

Плотность потока фотонов экспоненциально спадает вглубь от поверхности полупроводника.

Из формулы (52) ясно, что $[\alpha] = \text{м}^{-1}$. (В книгах и статьях по физике полупроводников чаще предпочитают измерять α в см^{-1} .)

Характер распределения лучистой энергии (плотности фотонов) в образце зависит от величины α и толщины образца d .

*) Иногда предпочитают говорить не о световом, а о лучистом потоке, чтобы подчеркнуть, что речь совсем не обязательно идет о фотонах, энергия которых соответствует видимому свету. Как мы увидим, область спектральной чувствительности фоторезисторов простирается от длин волн, соответствующих γ -лучам, до дальней инфракрасной области спектра.

В соответствии с формулой (52) плотность фотонов на задней грани образца толщиной d равняется $N_0 e^{-\alpha d}$. Эти, достигшие задней грани пластины, фотоны избежали поглощения. Поглотилось в образце

$$N_1 = N_0 - N_0 e^{-\alpha d} = N_0 (1 - e^{-\alpha d}) \quad (53)$$

фотонов (на единицу площади).

Если $\alpha d \ll 1$, т. е. если коэффициент поглощения очень мал или образец очень тонкий, то $e^{-\alpha d} \approx 1$ и при $x = d$ (на задней стенке образца) $N \approx N_0$. Распределение потока фотонов в образце практически однородно. Почти все фотоны, попавшие в образец, вылетают из него, не поглотившись.

При $\alpha d \approx 1$ плотность фотонов на задней стенке приблизительно в e раз ($e = 2,72$) меньше, чем на передней стенке образца. Заметная доля светового потока ($\approx 63\%$) поглощается в образце.

При $\alpha d \gg 1$ световой поток полностью поглощается в образце. Более того, практически все фотоны поглощаются в узкой, приповерхностной области (рис. 47, кривая 3).

Спектральная зависимость коэффициента поглощения. Величина коэффициента поглощения зависит от материала и от длины волны света λ , которым освещается образец. Чтобы подчеркнуть это последнее обстоятельство, коэффициент поглощения часто обозначается как $\alpha(\lambda)$.

Давайте попробуем, используя наши знания в области физики полупроводников, хотя бы качественно предугадать, как выглядит зависимость $\alpha(\lambda)$.

Когда длина волны света мала, энергия фотона $\mathcal{E}_\phi$, обратно пропорциональная длине волны λ ($\mathcal{E}_\phi = hc/\lambda$), велика. При $\mathcal{E}_\phi > \mathcal{E}_g$ фотоны могут разрывать электронные связи атомов кристаллической решетки полупроводника и образовывать электронно-дырочные пары. Так как атомов решетки очень много ($\sim 10^{23} \text{ см}^{-3}$), следует ожидать, что в этой области длин волн поглощение будет велико. По мере увеличения длины волны света и, соответственно, уменьшения энергии фотона до величины $\mathcal{E}_\phi \sim \mathcal{E}_g$ коэффициент поглощения будет падать. Энергии $\mathcal{E}_\phi = \mathcal{E}_g$ соответствует красная граница собственного внутреннего фотоэффекта *). Внутренний фотоэффект при $\mathcal{E}_\phi >$

*) Вспомните о красной границе внешнего фотоэффекта (Физика, 10 класс).

$> \mathcal{E}_g$ называют *собственным* потому, что при условии $\mathcal{E}_\phi > \mathcal{E}_g$ электроны и дырки всегда появляются парами, так же как в собственном полупроводнике под действием тепловых колебаний решетки.

Фотоны с энергией $\mathcal{E}_\phi < \mathcal{E}_g$ образовывать электронно-дырочные пары не способны. Однако в полупроводниках, как мы знаем, всегда присутствуют примеси. И даже при $\mathcal{E}_\phi < \mathcal{E}_g$ энергия кванта может оказаться достаточной для того, чтобы разорвать гораздо более слабую связь между донором и электроном и создать свободный фотоэлектрон. Или способствовать переходу электрона от атома решетки к атому акцепторной примеси с образованием дырки. В этом случае принято говорить о примесном внутреннем фотоэффекте. Так как концентрация примесей в подавляющем большинстве случаев гораздо меньше, чем концентрация атомов полупроводника, следует ожидать, что коэффициент α , соответствующий примесному поглощению, будет гораздо меньше, чем в случае собственного поглощения.

Итак, мы ожидаем, что в области малых значений λ (большие величины $\mathcal{E}_\phi$) поглощение будет велико. При увеличении λ до значения $\lambda_{гр} = hc/\mathcal{E}_g$, соответствующего красной границе собственного поглощения, значение α резко уменьшится и будет оставаться небольшим при дальнейшем увеличении λ .

Угадали ли мы ход зависимости $\alpha(\lambda)$? В основном, да. Однако некоторые важные оттенки оказались неучтенными.

На рис. 48 показана зависимость $\alpha(\lambda)$ для кремния *p*-типа при комнатной температуре. Качественно такой же вид имеют зависимости $\alpha(\lambda)$ для кремния *n*-типа, а также германия, арсенида галлия, фосфида индия *n*- и *p*-типов и многих других полупроводников. Проанализируем эту зависимость.

Собственное поглощение. В кремнии энергия образования электронно-дырочной пары $\mathcal{E}_\phi$ равняется 1,1 эВ. Условие $\mathcal{E}_\phi = \mathcal{E}_g$ выполняется, таким образом, при длине волны $\lambda_{гр} \approx 1,13$ мкм. Из рис. 48 видно, что в области собственного поглощения, т. е. в области коротких длин волн ($\lambda < \lambda_{гр}$), ход зависимости $\alpha(\lambda)$ мы описали правильно. При $\lambda \approx \lambda_{гр}/2 \approx 0,5$ мкм коэффициент поглощения составляет $\sim 10^4$ см⁻¹. С увеличением длины волны коэффициент поглощения резко уменьшается и при $\lambda \approx 1,3$ мкм

величина α даже в не очень хорошо очищенном кремнии ($N_a = 4 \cdot 10^{16} \text{ см}^{-3}$) составляет всего $\approx 0,36 \text{ см}^{-1}$ *).

Однако ход зависимости $\alpha(\lambda)$ с дальнейшим увеличением длины волны λ мы предсказали неверно. Коэффициент поглощения α монотонно возрастает с увеличением λ . Величина α при данной длине волны

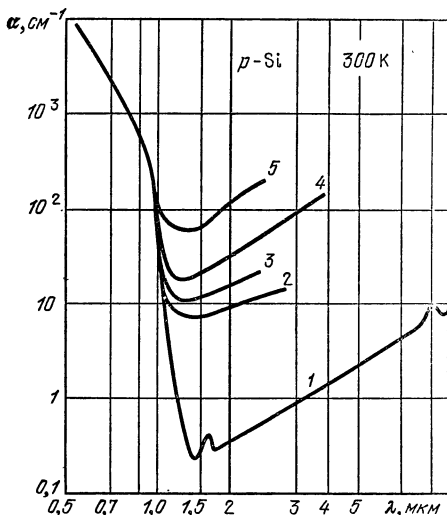


Рис. 48. Типичная зависимость коэффициента поглощения α от длины волны λ при комнатной температуре (на примере кремния p -типа).

Разные кривые соответствуют различным уровням легирования образца мелкой акцепторной примесью (бор): 1 — $N_a = 4 \cdot 10^{16} \text{ см}^{-3}$; 2 — $3 \cdot 10^{17} \text{ см}^{-3}$, 3 — $8 \cdot 10^{17} \text{ см}^{-3}$; 4 — $2,5 \cdot 10^{18} \text{ см}^{-3}$; 5 — $7 \cdot 10^{18} \text{ см}^{-3}$;

пропорциональна концентрации легирующей примеси и при концентрации примесей $\sim 10^{19} \text{ см}^{-3}$ даже минимальное значение α составляет $\approx 10^2 \text{ см}^{-1}$.

Какое же обстоятельство мы не учли?

Поглощение на свободных носителях. Дело в том, что при легировании кристалла мелкой примесью

*) В соответствии с формулой (52) и показанной на рис. 48 зависимостью $\alpha(\lambda)$ кремниевая пластинка толщиной $d = 1 \text{ мм}$ ослабит зеленый свет ($\lambda \approx 0,5 \text{ мкм}$; $\alpha \approx 10^4 \text{ см}^{-1}$) в 10^{430} раз, красный свет ($\lambda \approx 0,7 \text{ мкм}$, $\alpha \approx 2 \cdot 10^3 \text{ см}^{-1}$) в 10^{90} раз, а инфракрасное излучение с длиной волны $\lambda \sim 1,2 \text{ мкм}$ пройдет сквозь такую пластинку, практически не поглотившись! Кремниевые (а также германиевые и арсенид-галлиевые пластины) часто используют в качестве простых и эффективных фильтров, полностью отсекающих коротковолновое излучение.

(донорной или акцепторной) и достаточно высокой температуре все примесные атомы ионизованы за счет тепловых колебаний решетки кристалла (вспомните о примесном истощении, с. 77). Таким образом, с одной стороны, кванты света ионизовать атомы примеси не в состоянии — те уже и так ионизованы. С другой — свободные электроны и дырки в кристалле сами способны поглощать фотоны. Монотонное увеличение α с ростом λ при $\lambda > \lambda_{\text{гр}}$ вызвано поглощением света на свободных носителях заряда.

Этот механизм поглощения заслуживает обсуждения, хотя и не имеет отношения к внутреннему фотоэффекту: поглощение света на свободных носителях *не фотоактивно*. (Исчезновение фотона при этом механизме поглощения не приводит ни к появлению электрона, ни к появлению дырки. А следовательно, и не меняет проводимость кристалла).

Начнем с того, что действительно свободный носитель, например электрон в вакууме, вообще не способен поглотить фотон!

В самом деле, поглотив фотон, электрон приобрел бы энергию фотона $\mathcal{E}_\phi = h\nu$, а следовательно, и скорость v , определяемую очевидным равенством: $\mathcal{E}_\phi = = m_0 v^2/2$ (m_0 — масса свободного электрона в вакууме). Электрон, движущийся со скоростью v , обладает импульсом p_0 :

$$p_0 = m_0 v = \sqrt{2\mathcal{E}_\phi m_0} = \sqrt{2hm_0\nu}. \quad (54)$$

Откуда же взялся у электрона этот импульс? От поглощенного фотона? Посмотрим.

Импульс фотона p_ϕ равен (Физика, 10 класс):

$$p_\phi = \frac{\mathcal{E}_\phi}{c} = (\mathcal{E}_\phi/c^2) c = h\nu/c = h/\lambda. \quad (55)$$

Отношение импульса электрона, поглотившего фотон с энергией $\mathcal{E}_\phi = h\nu$, к импульсу такого фотона равняется:

$$p_0/p_\phi = c \sqrt{2m_0/h\nu}. \quad (56)$$

Подставив в формулу (56) значения скорости света в вакууме c , постоянной Планка h и массы свободного электрона m_0 , можно эту формулу переписать в виде:

$$p_0/p_\phi \approx 1,6 \cdot 10^{10} / \sqrt{\nu}. \quad (57)$$

Частота электромагнитной волны, соответствующей, например, зеленому свету ($\lambda = 0,5$ мкм), рав-

няется, как нетрудно подсчитать, $\nu = c/\lambda = 6 \cdot 10^{14}$ Гц. Таким образом, электрон, поглотивший квант зеленого света, обладает, в соответствии с формулой (57), импульсом в 650 раз большим, чем этот поглощенный фотон! Ясно, что позаимствовать такой импульс у фотона электрон не в состоянии. Откуда же электрон может взять необходимый импульс? В вакууме такой импульс электрону взять неоткуда. Между тем, при поглощении электроном фотона, как и при любом физическом процессе, должны выполняться основные законы сохранения — законы сохранения импульса и энергии. Как мы убедились, в акте поглощения фотона свободным электроном в вакууме законы сохранения импульса и энергии одновременно выполняться не могут. Следовательно? Следовательно, такой процесс невозможен *).

Другое дело в кристалле. Электрон, поглотивший квант, может получить необходимый импульс, благодаря тепловым колебаниям решетки (фононам) или благодаря взаимодействию с примесными центрами. Ясно, что чем меньший импульс требуется позаимствовать электрону, чтобы поглотить фотон, тем больше вероятность, что он его получит. Требуемый импульс уменьшается с уменьшением энергии фотона (см. формулу (54)). Вот почему поглощение на свободных носителях растет с длиной волны света, т. е. с уменьшением энергии фотона (рис. 48)!

Примесное поглощение. На кривой I рис. 48 в области длины волны $\lambda \approx 9$ мкм виден небольшой всплеск коэффициента поглощения. Это, пробиваясь, сквозь фон поглощения на дырках, напоминает о себе третий, уже упоминавшийся механизм поглощения, — примесное поглощение. В данном случае поглощение

*) Вдумчивый читатель может обратить внимание на то обстоятельство, что для электромагнитной волны с частотой $\nu \approx 2,6 \cdot 10^{20}$ Гц отношение p_0/p_f , в соответствии с формулой (57), равняется единице. И, казалось бы, квант с энергией $\mathcal{E}_f = h\nu \approx 10^6$ эВ может поглотиться электроном. Такой вывод, однако, был бы неправильным. Скорость электрона, поглотившего квант с такой большой энергией, уже нельзя подсчитывать по простой формуле $\mathcal{E}_f = m_0 v^2/2$. Подсчет дал бы значение v большее, чем скорость света в вакууме c . При таких больших энергиях для подсчета энергии и импульса электрона следует пользоваться релятивистскими формулами теории относительности (Физика, 10 класс). Убедитесь сами, используя релятивистские соотношения, что поглощение электроном в вакууме кванта с любой энергией невозможно.

на кислороде — примеси, всегда присутствующей в кремнии. Пик примесного поглощения можно исследовать гораздо более детально, если убрать фон — поглощение на свободных носителях. Для этого необходимо охладить кристалл до температуры значительно меньшей, чем температура примесного истощения. Электроны, не возмущаемые тепловыми колебаниями решетки, будут прочно удерживаться атомами доноров. А акцепторы утратят способность отбирать электроны у атомов решетки и образовывать дырки.

Свободные носители в кристалле исчезнут. Исчезнет, естественно, и поглощение на свободных носителях.

В таких условиях появляется возможность исследовать не только примесное поглощение на глубоких

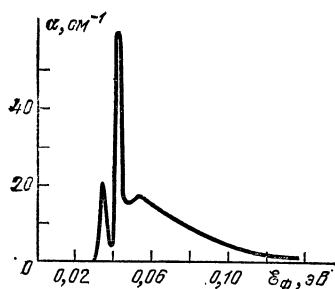


Рис. 49. Зависимость коэффициента поглощения α в кремнии, легированном бором, от энергии фотонов $\mathcal{E}_\phi$ ($T=4\text{K}$).

центрах, для ионизации которых требуется квант света с относительно большой энергией, но и поглощение на мелких донорах и акцепторах с очень небольшой энергией ионизации.

На рис. 49 показана спектральная зависимость коэффициента поглощения в кремнии, охлажденном до температуры жидкого гелия ($\approx 4\text{K}$) в области длин волн

$\lambda \approx 10\text{--}60\text{ мкм}$ (энергия фотонов $0,02\text{--}0,12\text{ эВ}$). Максимальное поглощение в области энергии фотонов $\sim 0,04\text{ эВ}$ соответствует энергии ионизации мелкого акцептора — бора в кремнии.

И тут вполне уместен вопрос «почему?»

Почему, в отличие от случая собственного поглощения (рис. 48), зависимость α от энергии фотона при примесном поглощении имеет максимум?

То обстоятельство, что поглощение исчезает при энергии фотона $\mathcal{E}_\phi$, меньшей, чем энергия ионизации примеси $\Delta\mathcal{E}$, вопросов не вызывает. А вот почему падает поглощение при $\mathcal{E}_\phi > \Delta\mathcal{E}$?

Причина заключается в тех же неумолимых требованиях законов сохранения энергии и импульса, которые, как мы только что видели, мешают свобод-

ному электрону поглотить фотон. При $\mathcal{E}_\phi > \Delta\mathcal{E}$ избыточная энергия фотона $\mathcal{E}_\phi - \Delta\mathcal{E}$ должна быть передана свободному электрону. Однако это возможно только в том случае, если одновременно электрон получит и соответствующий импульс $p = [2(\mathcal{E}_\phi - \Delta\mathcal{E})/m^*]^{1/2}$ (ср. с формулой (54) и вспомните, что движение электронов в кристалле характеризуется эффективной массой m^*). Получить нужный импульс от фотона электрон не может (см. формулы (56) и (57)). Он должен позаимствовать его у тепловых колебаний решетки (фононов) или у примесных центров. Чем больше требуемый импульс, тем труднее его получить. Именно поэтому коэффициент примесного поглощения и падает с увеличением избыточной энергии фотона $\mathcal{E}_\phi - \Delta\mathcal{E}$.

Почему же в таком случае коэффициент собственного поглощения монотонно растет с увеличением энергии фотона (рис. 48)? Дело в том, что при собственном поглощении исчезновение фотона приводит к появлению не одной, а двух частиц: электрона и дырки. При этом удовлетворить требованию закона сохранения импульса гораздо легче: например, если эффективные массы электрона и дырки равны и, появившись, они с равными скоростями разлетаются в противоположные стороны, суммарный импульс, требуемый для генерации электронно-дырочной пары, просто равен нулю.

Примесное поглощение часто бывает *фотоактивным*. В результате поглощения фотона появляется свободный электрон (или дырка). Фотоактивное примесное поглощение широко используется для создания фоторезисторов; чувствительных в инфракрасной области спектра, где энергия фотонов невелика.

А теперь используем приобретенные знания, чтобы решить следующую задачу. Пусть в нашем распоряжении имеются все полупроводники, указанные в табл. II. Вырежем из каждого из них пластинку и посмотрим через нее на свет. Что мы увидим?

Самый длинноволновый фотон, который еще способен вызвать ощущение цвета у человека, — это квант красного цвета, которому соответствует длина волны $\lambda \approx 0,75$ мкм. Энергия такого кванта $\mathcal{E}_\phi = 1,65$ эВ. Из табл. II видно, что эта величина гораздо больше, чем энергия образования электронно-дырочной пары $\mathcal{E}_g$ в InSb, Ge, Si, InP и GaAs. Следовательно, через

пластины, изготовленные из этих полупроводников, мы не увидим ничего. Любой квант, соответствующий видимому свету, будет поглощен даже в относительно тонком слое полупроводника — за счет сильнеешего собственного поглощения.

В GaP энергия образования электронно-дырочной пары ($\mathcal{E}_g \approx 2,3$ эВ) больше, чем энергия квантов красного, оранжевого и даже желтого ($\lambda_{ж} \approx 0,55$ мкм, $\mathcal{E}_ф \approx 2,25$ эВ) цвета. Но меньше, чем энергия квантов зеленого ($\mathcal{E}_ф \approx 2,5$ эВ) цвета. Следует поэтому ожидать, что пластина GaP поглотит кванты зеленого (а тем более, еще более энергичные кванты синего и фиолетового) цвета и пропустит красный, оранжевый и желтый. Действительно, кристаллы GaP имеют на просвет очень красивый красно-оранжевый оттенок. А как будет окрашен на просвет кристалл SiC с энергией $\mathcal{E}_g = 3,2$ эВ? Эта величина больше, чем энергия даже самого коротковолнового, фиолетового кванта ($\lambda \approx 0,4$ мкм; $\mathcal{E}_ф \approx 3,1$ эВ). Следовательно, такой кристалл пропустит весь спектр видимого цвета и будет выглядеть совершенно прозрачным? Да, такие кристаллы встречаются, но исключительно редко. Большинство кристаллов SiC на просвет выглядят зелеными. Их окрашивает (за счет примесного поглощения) примесь азота, практически всегда присутствующая в карбиде кремния. (Как всегда: чтобы ответ был правильным, нужно учесть множество оттенков. Иногда — самых неожиданных.)

Квантовая эффективность. Пусть на переднюю грань полупроводниковой пластины в единицу времени попадает N_0 квантов света с длиной волны λ . Коэффициент поглощения $\alpha(\lambda)$ и толщина пластины d известны. Сколько фотоэлектронов (и фотодырок) возникает в образце?

Число квантов, поглощающихся в образце, нам известно. Оно определяется формулой (53). Не познакомься мы с различными процессами поглощения, можно было бы подумать, что число возникших фотоносителей равно числу поглощенных квантов.

Но теперь мы знаем, что процессы поглощения могут быть фотоактивными и нефотоактивными.

Отношение числа возникших под действием света электронно-дырочных пар (при собственном поглощении) или числа электронов (дырок) при примесном поглощении к общему числу поглощенных квантов

называется *квантовой эффективностью* или *внутренним квантовым выходом*. Обозначается эта величина обычно греческой буквой β . Если все поглотившиеся в образце кванты исчезли за счет нефотоактивного процесса поглощения, например, на свободных носителях, внутренний квантовый выход $\beta = 0$. Взгляните на рис. 48. Квантовая эффективность поглощения в кремнии для длины волны $\lambda \approx 3$ мкм практически равна нулю

Напротив, для длины волны $\lambda \approx 0,7$ мкм величина β практически равняется единице — каждый поглощенный квант в области сильного собственного поглощения генерирует электронно-дырочную пару.

Рассмотрим более сложную задачу. В чистом кремнии при комнатной температуре коэффициент поглощения света с $\lambda = 1,06$ мкм равен приблизительно $\alpha_1 = 10$ см⁻¹ (рис. 48, кривая 1). При этом все поглощение фотоактивно ($\beta = 1$). В том же кремнии, легированном бором до концентрации $\sim 7 \cdot 10^{18}$ см⁻³, коэффициент поглощения α на той же длине волны $\lambda = 1,06$ мкм равен ≈ 100 см⁻¹ (рис. 48, кривая 5). Какова величина β в таком легированном кремнии?

При комнатной температуре все атомы медьного акцептора — бора полностью ионизованы. А значит, увеличение α произошло не за счет фотоактивного примесного поглощения, а за счет нефотоактивного поглощения на свободных носителях.

Выделим в образце узкую полоску толщиной Δx (рис. 47). Пусть плотность потока фотонов в этой полоске равна N . Число фотонов, которое поглотится внутри полоски в единицу времени за счет фотоактивного поглощения с созданием электронно-дырочных пар, равняется, очевидно:

$$-\Delta N_1 = \alpha_1 N \Delta x. \quad (58)$$

А общее число поглощенных внутри полоски квантов

$$-\Delta N = \alpha N \Delta x *). \quad (59)$$

По определению, квантовая эффективность $\beta = \Delta N_1 / \Delta N = \alpha_1 / \alpha$. Таким образом, в кремнии, легированном бором до концентрации $N_a \sim 7 \cdot 10^{18}$ см⁻³,

*) Формулы (52) и (59) описывают одно и то же явление: уменьшение плотности потока фотонов по мере проникновения в глубь полупроводника. Постарайтесь (это — не простая задача) установить связь между выражениями (52) и (59).

внутренний квантовый выход на длине волны $\lambda \approx 1,06$ мкм равен приблизительно 0,1.

Если энергия кванта очень велика и значительно превышает энергию создания электронно-дырочной пары $\mathcal{E}_g$, то один такой квант может создать в полупроводниковом кристалле тысячи и даже миллионы электронно-дырочных пар. Так, один квант γ -излучения с энергией $\mathcal{E}_\phi \approx 10^6$ эВ создает в кремнии приблизительно $4 \cdot 10^5$ электронно-дырочных пар ($\beta \approx 4 \cdot 10^5$!).

В фотосопротивлениях величина β обычно очень близка к единице. Это обстоятельство, в частности, дало повод научной сотруднице Ленинградского Физико-технического института М. П. Михайловой к созданию следующих шуточных строк:

«Знает верблюд, и кочевник, и птица,
Что квантовый выход — всегда единица».

Исключения составляют фотосопротивления, предназначенные для регистрации рентгеновских и γ -лучей. В этих приборах $\beta \gg 1$.

Фотопроводимость

С понятием проводимости мы обстоятельно познакомились в гл. 5. В случае электронного полупроводника проводимость описывается формулой (17, а), в случае дырочного — формулой (17, б).

Когда за счет внутреннего фотоэффекта в полупроводнике под действием света возникают дырки, электроны или электронно-дырочные пары, проводимость увеличивается.

Разница $\Delta\sigma$ между проводимостью освещенного полупроводника и темновой проводимостью называется фотопроводимостью.

Под действием света в полупроводнике возникают избыточные по отношению к равновесным носители тока. Благодаря процессу рекомбинации избыточные носители гибнут (рекомбинируют). В условиях, когда процессы генерации и рекомбинации уравниваются друг друга, в образце существует стационарная (не зависящая от времени) фотопроводимость.

Чтобы определить значение стационарной фотопроводимости $\Delta\sigma_{ст}$, следует определить, сколько носителей возникает в полупроводнике под действием света (скорость генерации) и сколько их гибнет за счет рекомбинации. Вспомните, совершенно такую же

программу расчета мы осуществляли, когда вычисляли равновесную концентрацию электронно-дырочных пар в собственном полупроводнике.

Скорость генерации. Снова рассмотрим пластинку полупроводника, на поверхность которой падает лучистый поток (рис. 47). Далеко от этой пластинки нам в данной главе уйти не удастся просто потому, что фоторезистор и представляет собой пластинку полупроводника, на которую падает лучистый поток.

Пусть толщина пластины d мала, так что $\alpha d \ll 1$. Тогда, в соответствии с формулой (52), число фотонов, падающих на переднюю стенку образца, практически равно числу фотонов, достигающих задней. Поглощение мало, и плотность фотонов в образце практически однородна. Ну а если так, то мы знаем, какая доля фотонов поглощается в образце: она определяется формулой (59)*). Только Δx в этой формуле следует заменить на толщину пластины d .

Пусть каждую секунду на единицу поверхности передней грани образца падает N_0 квантов света. Тогда за время Δt в пластине единичной площади поглотится $\alpha N_0 d \Delta t$ квантов. Примем, что внутренний квантовый выход равен β и предположим, для простоты, что под действием света возникают только электроны (примесное поглощение). Тогда число электронов, возникших на единицу площади пластины за время Δt , будет, очевидно, $\Delta N = \alpha \beta N_0 d \Delta t$. Или в единице объема: $\Delta n = \alpha \beta N_0 \Delta t$.

Скорость генерации g , т. е. увеличение концентрации носителей в единицу времени под действием света, равняется

$$g = \Delta n / \Delta t = \alpha \beta N_0. \quad (60)$$

Скорость рекомбинации. Скорость, с которой гибнут неравновесные электроны, возникшие под действием света, зависит от того, при каких обстоятельствах они погибают. В фотосопротивлениях избыточные носители в подавляющем большинстве случаев гибнут под «ножом убийц» — глубоких центров (см.

*) Конечно, долю поглощенных в образце фотонов можно вычислить и по формуле (53). Убедитесь сами, что вычисления по формулам (53) и (59) дают практически одинаковые результаты при $\alpha d \ll 1$ (тонкий образец). При $\alpha d \sim 1$, а тем более при $\alpha d \gg 1$ результаты вычислений по этим формулам совершенно различны. Постарайтесь понять, в чем тут дело.

гл. 4). В этом случае число неравновесных носителей, рекомбинирующих в единице объема в единицу времени, — *скорость рекомбинации* r , обратно пропорциональна времени жизни избыточных электронов τ_n *)

$$r = \Delta n / \tau_n. \quad (61)$$

Стационарная фотопроводимость. В стационарных условиях генерация избыточных фотоносителей уравновешивается их рекомбинацией: $g = r$. Или

$$\alpha \beta N_0 = \Delta n_{\text{ст}} / \tau_n.$$

Избыточная стационарная концентрация фотоэлектронов $\Delta n_{\text{ст}}$:

$$\Delta n_{\text{ст}} = \alpha \beta N_0 \tau_n. \quad (62)$$

Эти избыточные фотоэлектроны обеспечивают избыточную (по отношению к темновой) проводимость проводника $\Delta \sigma_{\text{ст}}$ — стационарную фотопроводимость. В соответствии с формулой (17, а), в рассматриваемом случае величина стационарной фотопроводимости $\Delta \sigma_{\text{ст}}$ равняется

$$\Delta \sigma_{\text{ст}} = e \Delta n_{\text{ст}} \mu_n = e \alpha \beta N_0 \tau_n \mu_n. \quad (63)$$

Формулу (63) полезно переписать в виде:

$$\Delta \sigma_{\text{ст}} = e \alpha \beta \frac{P}{h\nu} \tau_n \mu_n, \quad (64)$$

где P — световая энергия, падающая на единицу площади образца в секунду (Дж/(см²·с)); $h\nu$ — энергия кванта света (в джоулях).

Если поглощение собственное и в образце возникают одновременно и электроны и дырки, то

$$\Delta \sigma_{\text{ст}} = e \alpha \beta \frac{P}{h\nu} (\tau_n \mu_n + \tau_p \mu_p). \quad (65)$$

А теперь для наглядности — конкретный пример. Пусть на кремниевую пластинку падает красный свет ($h\nu = 1,65 \text{ эВ} \approx 2,65 \cdot 10^{-19} \text{ Дж}$). Мощность источника света такова, что на единицу площади попадает $P = 0,1 \text{ Дж/(см}^2 \cdot \text{с)}$. Каково значение стационарной фотопроводимости?

Времена жизни носителей в кремнии τ_n и τ_p можно менять в очень широких пределах: от $\sim 10^{-9} \text{ с}$ до

*) Значение τ_n , конечно, зависит от характера глубоких центров и от их концентрации.

$\tau \sim 10^{-3}$ с введением примесей, дающих глубокие рекомбинационные уровни. Примем $\tau_n = \tau_p = 10^{-5}$ с (это достаточно типичная величина). Все остальные данные в нашем распоряжении есть. По длине волны света λ можно (из рис. 48) определить величину α . Значение β в области сильного собственного поглощения следует считать равным единице. Подвижности μ_n и μ_p для кремния указаны в табл. III.

Если вы внимательно проделали расчет, то получили величину $\Delta\sigma_{ст} \approx 1 \text{ Ом}^{-1} \cdot \text{см}^{-1}$. Велика она или мала?

Как всегда, это зависит от того, с чем сравнивать.

Если освещалась пластинка кремния с проводимостью, близкой к собственной ($n_i = p_i \approx 10^{10} \text{ см}^{-3}$ при комнатной температуре), то темновая проводимость кремния σ_t равнялась:

$$\sigma_t = e(n_i \mu_n + p_i \mu_p) = e n_i (\mu_n + \mu_p) \approx 3 \cdot 10^{-6} \text{ Ом}^{-1} \cdot \text{см}^{-1}.$$

Для такой пластины возникновение стационарной фотопроводимости $\Delta\sigma_{ст} \approx 1 \text{ Ом}^{-1} \cdot \text{см}^{-1}$ означает увеличение проводимости под действием света в 300 000 раз!

С другой стороны, пластинка *p*-кремния, легированная до концентрации 10^{18} см^{-3} , обладает проводимостью $\sigma \approx 30 \text{ Ом}^{-1} \cdot \text{см}^{-1}$. (Подвижность дырок с увеличением уровня легирования уменьшается — рис. 27.) Для такой пластины $\Delta\sigma_{ст} = 1 \text{ Ом}^{-1} \cdot \text{см}^{-1}$ означает изменение проводимости под действием света такой же интенсивности всего на 3%.

Как устанавливается фотопроводимость. Итак, если свет падает на пластину полупроводника достаточно долго, то в образце возникает стационарная фотопроводимость, определяемая формулами (65) и (66). Неплохо бы теперь ответить на вопрос, что значит «достаточно долго»? За какое время и по какому закону в образце устанавливается под действием света избыточная проводимость $\Delta\sigma_{ст}$?

Скорость генерации носителей под действием света g , определяемая формулой (60), не зависит от концентрации избыточных носителей Δn . А скорость рекомбинации r , в соответствии с формулой (61), пропорциональна Δn . Пусть в момент $t=0$ включился свет. В первый момент избыточные носители отсутствуют. Скорость рекомбинации равна нулю. А скорость генерации, не зависящая от Δn , равняется по-

стоянной величине (формула (60)). Генерация преобладает над рекомбинацией, и концентрация Δn растет. Однако по мере роста Δn увеличивается скорость рекомбинации (формула (61)). Когда скорость рекомбинации сравнивается со скоростью генерации, в образце устанавливается стационарная концентрация $\Delta n_{\text{ст}}$, определяемая выражением (62).

Зависимость, определяющая увеличение концентрации фотоносителей со временем, имеет вид:

$$\Delta n = \Delta n_{\text{ст}} (1 - e^{-t/\tau_n}). \quad (66)$$

В момент $t=0$ величина Δn , как видно из формулы (66), равна нулю. При $t \rightarrow \infty$ концентрация избыточных электронов стремится к стационарной величине $\Delta n_{\text{ст}}$. Если с момента включения света прошло время, равное времени жизни $\tau_n (t = \tau_n)$, то в соответствии с формулой (66) $\Delta n \approx 0,63 \Delta n_{\text{ст}}$. Через $t = 2\tau_n$ концентрация составит $\approx 86\%$ от стационарной, а через $t = 3\tau_n$ отношение $\Delta n / \Delta n_{\text{ст}} = 0,95$ — стационарное состояние практически установилось *).

Характерное время нарастания концентрации неравновесных носителей после включения света определяется временем жизни τ_n .

Когда вычислена концентрация избыточных носителей Δn , то с определением фотопроводимости все просто. Она, естественно, равна произведению Δn на заряд электрона и подвижность:

$$\Delta \sigma = e \Delta n \mu_n = e \Delta n_{\text{ст}} \mu_n (1 - e^{-t/\tau_n}) = \Delta \sigma_{\text{ст}} (1 - e^{-t/\tau_n}). \quad (67)$$

После включения света фотопроводимость возрастает от нуля до установившегося значения $\Delta \sigma_{\text{ст}}$ за время порядка $3\tau_n$.

*) Физический смысл формулы (66) можно проанализировать глубже, если взять производную от этого выражения по времени и найти таким образом скорость изменения концентрации:

$$\frac{d}{dt} (\Delta n) = \frac{\Delta n_{\text{ст}}}{\tau_n} e^{-t/\tau_n}.$$

В момент $t = 0$ величина $d(\Delta n)/dt$ равна $\Delta n_{\text{ст}}/\tau_n = \alpha \beta N_0 = g$ (см. формулы (62) и (60)). С течением времени скорость изменения концентрации монотонно убывает и обращается в нуль при $t \rightarrow \infty$. Если вы справились с предыдущей задачей (с. 193), попробуйте, исходя из формул (60) и (61), получить формулу (66).

Как спадает фотопроводимость. Если после того, как в образце под действием света установилась стационарная фотопроводимость, выключить свет, концентрация избыточных носителей начнет уменьшаться. Ясно, что характерное время, за которое избыточные фотоэлектроны вымрут и исчезнут, есть время жизни τ_n . Мы говорили об этом уже много раз. Да, собственно, это и следует из самого смысла понятия времени жизни.

Закон, описывающий изменение концентрации со временем после выключения света, имеет вид:

$$\Delta n = \Delta n_{\text{ст}} e^{-t/\tau_n}. \quad (68)$$

Фотопроводимость $\Delta\sigma$, следовательно, спадает после выключения света со временем, как

$$\Delta\sigma = \Delta\sigma_{\text{ст}} e^{-t/\tau_n}. \quad (69)$$

За время $t = \tau_n$ избыточная проводимость уменьшится в e раз ($e \approx 2,72$), за время $2\tau_n$ — в e^2 раз ($e^2 \approx 7,4$), за $3\tau_n$ — в e^3 раз ($e^3 \approx 20$) и т. д.

Вспомним решенный нами пример на вычисление стационарной фотопроводимости. Значение $\Delta\sigma_{\text{ст}}$ в случае сильно легированного кремния составляло всего 3% от темновой проводимости σ_t . Через время $3\tau_n$ величина $\Delta\sigma \approx \Delta\sigma_{\text{ст}}/20$. Таким образом, значение σ через $t = 3\tau_n$ отличается от σ_t всего на 0,15%. В подавляющем большинстве задач можно принять, что затухание проводимости при такой разности $\sigma - \sigma_t$ практически закончено.

А вот в случае, когда освещалась пластинка, изготовленная из собственного кремния, величина $\Delta\sigma_{\text{ст}}$ превосходила значение σ_t в $3 \cdot 10^5$ раз. В этом случае через время $t = 3\tau_n$ величина фотопроводимости, уменьшившись в 20 раз, все еще будет превышать значение σ_t в 15 000 раз! Вряд ли можно в этом случае признать, что спад фотопроводимости закончен. Если мы согласимся считать процесс закончившимся в случае, когда проводимость будет отличаться от темновой, скажем, на 1%, то из формулы (69) получим, что для этого с момента выключения света должно пройти время $t \approx 17\tau_n$.

Да, характерное время спада фотопроводимости есть время жизни неравновесных носителей.

Да, за время $t = 3\tau_n$ фотопроводимость значительно, в 20 раз, спадает.

И однако это, как мы видим, не освобождает нас от необходимости думать над оттенками решаемой задачи.

Спектральная зависимость фотопроводимости. Из величин, определяющих в соответствии с полученными нами формулами (64) и (65) стационарную фотопроводимость, от длины волны зависят две: коэффициент поглощения α и квантовая эффективность β . Ну, и чего же тут думать? Если зависимости $\alpha(\lambda)$ и $\beta(\lambda)$ известны, то зависимость фотопроводимости от длины волны будет пропорциональна произведению $\alpha(\lambda) \times \beta(\lambda)$.

Нет. Не будет.

Давайте думать. Пусть (для простоты) величина $\alpha(\lambda)$ остаётся постоянной, а внутренний квантовый выход β меняется с изменением длины волны. Будет ли $\Delta\sigma_{ст}$ пропорциональна β ? По-видимому, да. Если α от λ не зависит, то распределение носителей, возникших в пластинке полупроводника под действием света, также не зависит от длины волны λ . В таких условиях все просто — чем больше β , тем большая доля поглощенных квантов генерирует электроны и дырки.

Ну что ж. И подумав, мы готовы уверенно повторить: величина фотопроводимости пропорциональна квантовой эффективности β !

Теперь обсудим другую ситуацию. Пусть коэффициент β от λ не зависит, а коэффициент поглощения α сильно меняется с изменением длины волны. Такую ситуацию представить себе очень просто. Обратимся к рис. 48. В области длин волн $\lambda \leq 1,1$ мкм коэффициент β в чистом кремнии практически равен единице. А величина α изменяется на три-четыре порядка при изменении λ всего в 2 раза.

Как будет зависеть фотопроводимость кремниевой пластинки от длины волны света λ , если менять λ в диапазоне от 0,5 до 1 мкм? Если, не думая, довериться формуле (65), то ответ состоит в том, что зависимость $\Delta\sigma_{ст}(\lambda)$ просто повторит в соответствующем масштабе зависимость $\alpha(\lambda)$.

В действительности дело обстоит совершенно иначе. Если экспериментально исследовать зависимость $\Delta\sigma_{ст}(\lambda)$, то результат эксперимента будет выглядеть

так, как показано на рис. 50. Возникает впечатление, что у внутреннего фотоэффекта имеется не только красная, но и фиолетовая граница. Вместо того чтобы монотонно возрастать с уменьшением длины волны, фотопроводимость, пройдя через максимум, резко падает при дальнейшем уменьшении λ .

Формула (65) ничего подобного не предусматривает. Поэтому возникают сразу два вопроса. Во-первых, как объяснить результаты эксперимента, во-вторых, почему «не работает» формула (65)?

Прежде всего, вспомним, что формула (65) получена нами для тонкой пластинки. Такой тонкой, что произведение $\alpha d \ll 1$ (d — толщина пластины). При $\alpha d \ll 1$, как мы установили, обсуждая формулу (53), во-первых, фотоны распределены в образце однородно, во-вторых, лишь небольшая их часть поглощается в образце. В таких условиях, чем больше α , тем больше фотонов, поглощаясь в образце, генерируют электронно-дырочные пары (формула (60)), и тем больше, следовательно, фотопроводимость.

По мере увеличения α , когда произведение αd становится больше единицы, практически все фотоны, проникшие в образец, поглощаются в нем (если на этот счет возникают какие-либо сомнения, следует снова внимательно посмотреть на формулу (53)). Уже теперь формула (65) перестает быть справедливой. Из (65) следует, что фотопроводимость пропорциональна величине α . Но при $\alpha d \gtrsim 1$ для этого нет никаких оснований: если и так все фотоны поглощаются в образце, то ясно, что с дальнейшим ростом α ожидать увеличения фотопроводимости не приходится.

Тогда, казалось бы, следует ожидать, что с ростом α величина $\Delta\sigma_{ст}$ насытится при $\alpha d \gtrsim 1$. Почему же

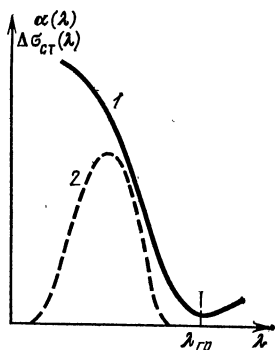


Рис. 50. Качественный ход спектральной зависимости коэффициента поглощения $\alpha(\lambda)$ (кривая 1) и фотопроводимости $\Delta\sigma_{ст}(\lambda)$ (кривая 2) в области собственной проводимости, $\lambda_{гр} = hc/\mathcal{E}_f$ соответствует красной границе собственного внутреннего фотоэффекта.

фотопроводимость падает с дальнейшим увеличением α ?

Потому, что при $\alpha d \gg 1$, как мы знаем, все кванты света поглощаются в узкой области вблизи поверхности образца (рис. 47, кривая 3). Время жизни

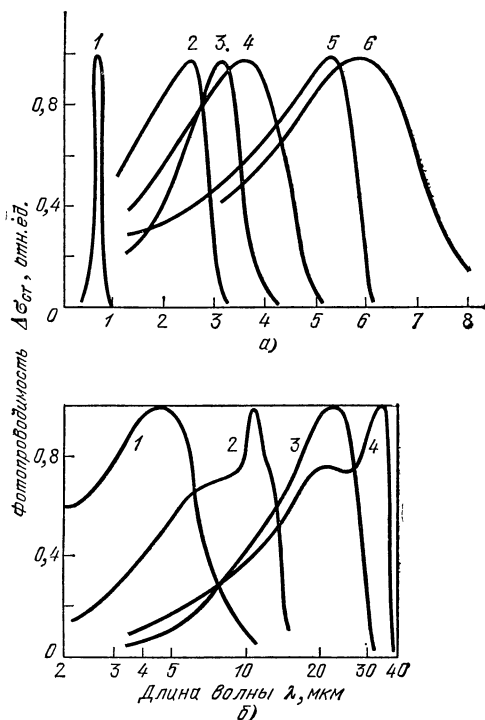


Рис. 51. Качественные зависимости спектральной чувствительности фотопроводимости различных полупроводниковых материалов.

а) В области собственной проводимости: 1 — CdS (300 K), 2 — PbS (300 K), 3 — PbS (77 K), 4 — PbSe (300 K), 5 — InSb (77 K), 6 — PbSe (77 K); б) при примесной проводимости. Показаны спектральные чувствительности фоторезисторов на основе германия, легированного различными примесями: 1 — германий с примесью золота, Ge: Au, (60 K); 2 — германий с ртутью, Ge: Hg, (27 K); 3 — германий с медью, Ge: Cu, (15 K); 4 — германий с цинком, Ge: Zn, (4 K). (Из книги: Аксененко М. Д., Красовский Е. А. «Фоторезисторы», — М.: Сов. радио, 1973.)

неравновесных носителей τ вблизи поверхности может быть в сотни и тысячи раз меньше, чем время жизни электронов и дырок в объеме полупроводника. Поверхность кристалла неизбежно содержит следы обработки: шлифовки, полировки, химического травления

и т. д. В результате концентрация дефектов вблизи поверхности значительно выше, чем в объеме. А многие дефекты действуют точно так же, как глубокие уровни — «профессиональные убийцы» неравновесных носителей (см. гл. 4). Чем больше величина a (при $ad \gg 1$), тем ближе к поверхности оказываются рожденные светом электроны и дырки. Кроме того, когда электроны и дырки рождаются в очень узком слое, концентрация их становится большой и они начинают сталкиваться друг с другом. В этих условиях, помимо гибели на глубоких центрах, заметный вклад в процесс исчезновения неравновесных носителей может давать непосредственная рекомбинация электронов и дырок. При этом время жизни начнет еще резче падать с дальнейшим ростом a .

Таким образом, при $ad > 1$ число носителей, генерируемых в образце, перестает меняться с ростом a , а скорость рекомбинации возрастает. В результате фотопроводимость уменьшается.

На рис. 51 показаны кривые спектральной чувствительности фотопроводимости различных полупроводниковых материалов, используемых при изготовлении фоторезисторов. Конечно, следует рассматривать эти кривые как качественные. Ясно, что точный ход зависимостей $\Delta\sigma_{\text{ст}}(\lambda)$ определяется толщиной образца d , характером обработки поверхности, уровнем легирования материала, объемным временем жизни и т. д. Видно, однако, что все без исключения материалы характеризуются избирательной (селективной) спектральной чувствительностью фотопроводимости.

Фоторезисторы

Мы долго трудились, вскапывая каменистую почву внутреннего фотоэффекта и ухаживая за капризными деревцами фотопроводимости. Зато теперь спелые плоды в виде фоторезисторов сами упали нам в руки.

На рис. 52 показаны некоторые из серийно выпускаемых промышленностью фоторезисторов. Они отличаются друг от друга формой, размерами, материалом, назначением. Тем не менее, изучив физику внутреннего фотоэффекта и фотопроводимости, мы многое о них уже знаем и, что, пожалуй, не менее важно, знаем, какие нужно задать вопросы, чтобы выяснить то, что еще неясно.

Основу любого фоторезистора составляет полупроводниковая пластинка или пленка.

Если фотосопротивление рассчитано на работу в области собственной проводимости, максимум спектральной чувствительности фоторезистора будет соответствовать энергиям фотонов, примерно равным энергии образования электронно-дырочной пары в данном полупроводнике $\mathcal{E}_g$ (рис. 50, 51, а).

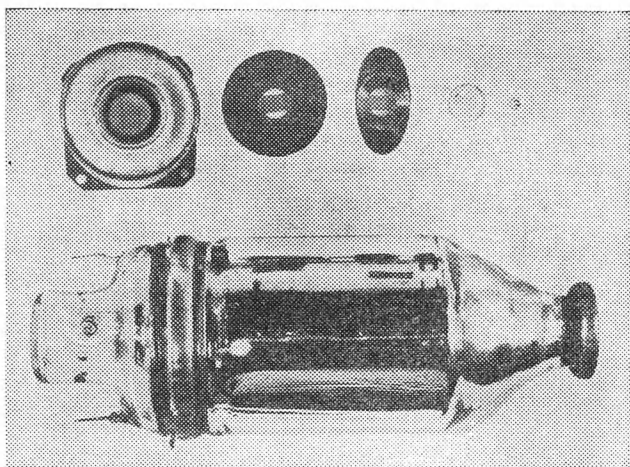


Рис. 52. Различные типы серийно выпускаемых фоторезисторов.

Если фоторезистор рассчитан на работу в области примесной проводимости, максимум спектральной чувствительности будет соответствовать энергии фотонов, приблизительно равной энергии ионизации примеси $\Delta\mathcal{E}$. Чем меньше энергия ионизации примеси, тем более длинноволновое излучение способен зарегистрировать фоторезистор. И тем ниже должна быть рабочая температура фоторезистора, чтобы избежать ионизации примеси за счет тепловых колебаний решетки. Фоторезистор, предназначенный для регистрации длинноволнового излучения, необходимо охлаждать.

Чем меньше время жизни неравновесных носителей заряда, тем быстрее будет реагировать фоторезистор на изменение светового потока (формулы (67))

и (69)). Зато тем меньше, при прочих равных условиях, будет чувствительность фотосопротивления (формулы (64) и (65)).

Что умеют фоторезисторы. В середине 70-х годов в одном из американских журналов была напечатана красочная фотография, сделанная со спутника. Однако это была не просто фотография, снятая на цветную пленку, а скорее изображение, которое можно назвать «теплофотографией». Фоторезистор, изготовленный из полупроводникового материала (CdHgTe), с ничтожно малой энергией образования электронно-дырочной пары $\mathcal{E}_g$, охлажденный до температуры жидкого гелия, с высокой точностью регистрировал температуру земных объектов по испускаемому ими тепловому инфракрасному излучению. Электронная система преобразования изображения по сигналам с фоторезистора от точки к точке строила изображение. Чем большую температуру имела точка земной поверхности, на которую был направлен фоторезистор, тем большей интенсивности голубой цвет соответствовал этой точке на цветной фотографии.

Снимок был сделан глубокой осенью, когда на изучаемой территории лежал снег. Чуть-чуть голубоватые, почти белые равнины занимали большую часть фотографии. Ослепительно ярко голубели поселки и небольшие городки. Через весь снимок протянулась широкая голубая лента громадной, еще не застывшей, а потому излучавшей тепло реки. Река выглядела странно. До некоторой точки она была ровного голубого цвета. А ниже по течению от этой точки — заметно более насыщенного, более темного, постепенно вновь переходившего в прежний ровный голубой оттенок. К той точке, где цвет реки менялся, была поставлена стрелка. И у стрелки написано: «В этой точке вода становится теплее на 0,5 градуса. Нагреть такую могучую реку на полградуса могут тепловые отходы только очень энергоемкого производства. Скорее всего, в этом месте расположен секретный завод по производству обогащенного плутония».

Еще один впечатляющий пример. В июне 1982 г., когда английские десантники захватили Порт-Стенли — административный центр Фолклэндских островов, в Южном полушарии была зима. Как писала газета «Известия», немалую роль в развитии событий сыграло то обстоятельство, что любая попытка аргентин-

ских солдат обогреться, развести костер, немедленно приводила к накрытию костра (а с ним, как правило, минометной или артиллерийской позиции) английской ракетой с головкой самонаведения. Фоторезистор, по сигналам которого наводилась ракета, реагировал на инфракрасное излучение костра.

Фоторезисторы, установленные на спутниках, следят за запуском чужих ракет. Каждый такой запуск сопровождается характерной и очень мощной вспышкой оптического и теплового излучения. Анализ характера вспышки позволяет судить о том, какого типа ракета запущена.

Способность фоторезисторов реагировать на тепловое излучение позволяет использовать их и в сугубо мирных целях — для измерения температуры расплавленной стали и чугуна в металлургической промышленности, раскаленной массы материала в керамической, цементной и многих других отраслях промышленности. Приборы для измерения температуры нагретых тел по интенсивности и спектральному составу теплового излучения называются *пирометрами*. Пирометры, в которых использованы фоторезисторы, способны измерять температуры, приблизительно в 10 раз более низкие, чем оптические пирометры.

Фоторезисторы на основе полупроводников, чья спектральная характеристика фотопроводимости имеет максимум в области видимого света (CdS , CdSe), используются в приборах, измеряющих уровень искусственной и естественной освещенности.

Фоторезисторы широко используются в системах автоматической охраны территорий и помещений. Световой луч, проходящий по периметру охраняемой территории, падает на фоторезистор. При пересечении луча сопротивление фоторезистора резко возрастает и на вход исполнительной системы поступает импульс, вырабатывающий сигнал тревоги. Чтобы обеспечить скрытность охранной системы, используется инфракрасный луч. А чтобы нарушитель не мог обмануть систему, подсвечивая фоторезистор своим собственным инфракрасным источником в то время, когда он пересекает охранный луч, применяется модуляция охранного луча: интенсивность луча быстро изменяется во времени по определенному закону. Шифр, в соответствии с которым производится модуляция, может меняться.

Сходный принцип лежит в основе охранной системы, скрывать которую нет необходимости. При проходе мимо стойки автоматического контроля в метро ваша нога пересекает световой луч, и фоторезистор вырабатывает сигнал «тревога». Если вы не отключили предварительно исполнительный механизм, опустив 5 копеек в щель монетоприемника, перед вами угрожающе хлопают металлические дверцы и к вам спешит дежурный контролер с выражением лица, не сулящим вам никаких житейских радостей.

Изменение сопротивления фоторезисторов при пререзании подсвечивающего луча используется в многочисленных счетчиках изделий на конвейерах, частотомерах, контролирующих и автоматически регулирующих частоту вращения двигателей, в защитных устройствах, ограждающих травмоопасные зоны станков и механизмов, в устройствах чтения перфокарт в электронных вычислительных машинах.

Фотосопротивления служат чувствительными элементами *нефелометров* — приборов для определения степени мутности жидкостей, суспензий, коллоидных растворов (греческое слово *perhelé* означает облако). Принцип работы нефелометра очень прост. Световой поток от калиброванного источника света, пройдя через мутную среду, попадает на фоторезистор. Чем выше степень замутненности среды, тем меньше света попадает на фоторезистор, тем выше, следовательно, сопротивление резистора. В нефелометрах, а также в сходных по принципу действия измерителях степени загрязненности воздуха в шахтах, подготовительных участках литейных цехов и т. д. часто используется схема моста Уитстона с двумя фоторезисторами. На один фоторезистор (сравнительный) свет попадает, пройдя через эталонную кювету с заданной степенью мутности. Если мутность контролируемой среды такая же, свет, попадающий на рабочий фоторезистор, имеет ту же интенсивность, что и свет, попадающий на сравнительное фотосопротивление. В результате оба фоторезистора имеют одинаковое сопротивление и выходной сигнал моста равен нулю (см. гл. 7). При отклонении прозрачности контролируемой среды от эталонной мост разбалансируется. По уровню выходного сигнала можно судить о степени замутненности среды.

Список профессий фоторезистора по разнообразию и универсальности, пожалуй, не уступает перечню

профессий термистора. Иногда эти приборы даже выступают в роли конкурентов. Мы уже обсуждали в гл. 6, как используется термистор в схемах пожарной сигнализации. Иногда, однако, благоразумнее бывает положиться на народную поговорку «нет дыма без огня» и не ждать, пока температура в помещении поднимется так высоко, что сработает термисторная тепловая сигнализация. Дым, появившийся в помещении, ослабляет сигнал, попадающий на фоторезистор от специальной осветительной лампочки. При определенной степени задымленности пожарная сигнализация вырабатывает сигнал тревоги.

Фоторезисторы используются для измерения частоты и амплитуды вибраций, скоростей и геометрических размеров объектов, измерения перемещений, в том числе и очень небольших: фоторезисторное устройство способно зафиксировать перемещение объекта на десятые доли микрометра. О применениях фоторезисторов см. книгу Э. О. Богданова «Фоторезисторы и их применение». — Л.: Энергия, 1978.

Основные параметры фоторезисторов. Некоторые из важнейших параметров фоторезисторов определяются, в основном, свойствами полупроводникового материала, из которого они изготовлены. С этими параметрами мы практически уже знакомы. К ним относятся *темновое сопротивление* фоторезистора R_T , диапазон спектральной чувствительности и быстродействие.

Темновое сопротивление R_T определяется темновой проводимостью полупроводникового материала и его геометрическими размерами. Для фоторезисторов различных типов оно лежит в пределах от $\sim 10^2$ до $\sim 10^8$ Ом. Небольшие значения R_T характерны для фотосопротивлений с высоким быстродействием. Большие величины R_T типичны для относительно медленных, но чувствительных фоторезисторов.

Диапазон спектральной чувствительности, т. е. диапазон длин волн, к которым чувствительно фотосопротивление и длина волны, соответствующая максимальной чувствительности, λ_m , определяются, в основном, фундаментальными свойствами полупроводникового материала (рис. 51): величиной энергии образования электронно-дырочной пары $\mathcal{E}_g$ для фоторезистора, работающего в области собственного поглощения, и энергией ионизации примеси $\Delta\mathcal{E}$ для фоторезистора, в котором используется примесное поглощение.

Быстродействие фоторезисторов характеризуется постоянной времени спада и нарастания фототока τ . Величины τ для фоторезисторов различных типов лежат в пределах от $\sim 10^{-2}$ до $\sim 10^{-10}$ с. Самые быстрые фоторезисторы с $\tau \sim 10^{-10}$ с используются для проведения физических экспериментов, в которых необходимо фиксировать изменение во времени очень быстро меняющихся световых потоков. Казалось бы, изготовить материал с малым временем жизни τ совсем не сложно. Достаточно ввести в него много глубоких примесей. Однако, когда речь идет об очень малых значениях τ , которые необходимо получить в данном полупроводниковом материале, задача может оказаться совсем не простой. Каждая примесь имеет свой предел растворимости в данном материале. Введение очень большого количества примеси может привести к резкому ухудшению других свойств полупроводника (например, уменьшению подвижности, появлению структурных дефектов и т. д.). Так что для получения очень малых значений τ иногда приходится прибегать к особым технологическим приемам. Так, летом 1982 г. американские исследователи опубликовали работу, в которой сообщали, что им удалось изготовить фоторезистор из фосфида индия (InP) с $\tau \leq 10^{-10}$ с с помощью бомбардировки InP протонами высоких энергий.

Важность других параметров определяется характером применений фоторезистора. Несмотря на громадное разнообразие схем и устройств, в которых используются фоторезисторы, можно условно выделить два главных типа фотосопротивлений.

К первому типу относятся фоторезисторы, которые используются в схемах со специальной подсветкой (счетчики, фотореле, измерители угловых ускорений и перемещений, охранные системы и т. д.). Для таких фоторезисторов, работающих при больших световых потоках, определяющим параметром является *кратность изменения сопротивления K* (при определенном уровне засветки). Величина K равняется отношению темнового сопротивления R_T к сопротивлению освещенного фоторезистора. Типичное значение K лежит в пределах от $\sim 10^2$ до $\sim 10^3$ (при освещенности 200—300 лк).

Характер применения фоторезисторов второго типа требует способности обнаруживать минимально возможные уровни световых (лучистых) потоков. Для та-

ких фоторезисторов главный параметр получил очень естественное название *обнаружительной способности*. Этот параметр обозначается обычно символом D^* .

По смыслу величина D^* обратно пропорциональна той минимальной мощности, которую способен зарегистрировать фотоприемник.

Обнаружительная способность D^* — сложный критерий, в котором учтен целый ряд параметров фоторезистора. В том числе и собственный шум фотосопротивления, мешающий регистрировать очень слабые лучистые потоки. (Подробный вывод формулы, связывающей обнаружительную способность фоторезистора с другими его параметрами, содержится в книге М. Д. Аксененко и Е. А. Красовского «Фоторезисторы», — М.: Сов. радио, 1973.) Мы ограничимся здесь тем, что сравним обнаружительную способность реальных фоторезисторов и идеального фоторезистора, который никакими собственными шумами не обладает.

Казалось бы, такой фоторезистор может регистрировать любой, даже самый слабый лучистый поток. Это заключение, однако, неправильно. Дело в том, что при конечной температуре окружающей среды на фоторезистор воздействуют кванты теплового излучения (фона). Даже в отсутствие специального источника, который должен быть обнаружен фоторезистором, в фотосопротивлении под действием фонового

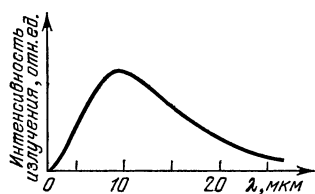


Рис. 53. Зависимость интенсивности излучения от длины волны для 300 К.

излучения непрерывно происходит процесс генерации свободных носителей, меняющий проводимость кристалла. Излучение фона хаотически меняется во времени (флуктуирует) и эти флуктуации (шум) мешают обнаруживать слабое излучение источника.

Распределение энергии в спектре излучения (Физика, 10 класс) зависит от температуры фона или, попросту говоря, от температуры окружающей фоторезистор среды. Спектр излучения тел, соответствующий комнатной температуре (300 К), показан на рис. 53. Из рисунка видно, что максимум энергии излучения лежит в инфракрасной области спектра ($\lambda \sim 10$ мкм). Фотонов с длинами волн, соответствующих видимой

области ($\lambda \sim 0,4\text{--}0,7$ мкм), в спектре практически нет. Впрочем, мы это хорошо знаем и сами. Ночью, когда тела не светят отраженным светом Солнца, очень легко набить себе шишку в темноте, несмотря на то, что окружающие нас тела очень интенсивно «светятся» — в инфракрасном диапазоне.

Представим себе, что в нашем распоряжении находится фоторезистор, изготовленный из полупроводника с большой энергией образования электронно-дырочной пары, например, из фосфида галлия ($\mathcal{E}_g = 2,3$ эВ). Длина волны, соответствующая красной границе внутреннего фотоэффекта, равняется $\lambda_{гр} \approx \approx 1,24/\mathcal{E}_g \approx 0,54$ мкм. Фотоны с энергией $\mathcal{E}_\phi < < \mathcal{E}_g$ генерировать в таком фоторезисторе электронно-дырочные пары не могут. Но в спектре фона при $T = 300$ К практически нет квантов света с энергией $\mathcal{E}_\phi > \mathcal{E}_g$. Поэтому фоновое излучение, соответствующее комнатной температуре, будет создавать в таком фоторезисторе ничтожные шумы. Следовательно, обнаружительная способность D^* у идеального фоторезистора, изготовленного из материала с большим значением $\mathcal{E}_g$, будет очень высокой.

Если выбрать материал для фоторезистора с меньшим значением $\mathcal{E}_g$, то бóльшая доля квантов, излучаемых окружающей фоторезистор средой, будет вызывать генерацию электронно-дырочных пар в полупроводнике. Уровень шума, создаваемого фоном, повысится, а обнаружительная способность — понизится. С уменьшением $\mathcal{E}_g$ полупроводника (с увеличением $\lambda_{гр}$) обнаружительная способность D^* уменьшается.

Однако, если $\lambda_{гр}$ полупроводника становится заметно больше, чем длина волны, соответствующая максимуму теплового излучения фона ($\lambda \sim 10$ мкм, см. рис. 53), то уже практически все фотоны, излучаемые окружающей средой, генерируют в полупроводнике свободные носители заряда и дают вклад в фоновый шум. Дальнейшее увеличение $\lambda_{гр}$ полупроводника практически не будет увеличивать уровень шума. В таких условиях обнаружительная способность перестанет падать с увеличением $\lambda_{гр}$.

На рис. 54 показана теоретически рассчитанная зависимость обнаружительной способности D^* идеального фоторезистора от значения $\lambda_{гр}$ полупроводникового материала. На этом же рисунке точками показаны максимальные величины D^* для реальных

фоторезисторов, изготовленных из различных полупроводниковых материалов. Как видно из рисунка, обнаружительная способность реальных приборов,

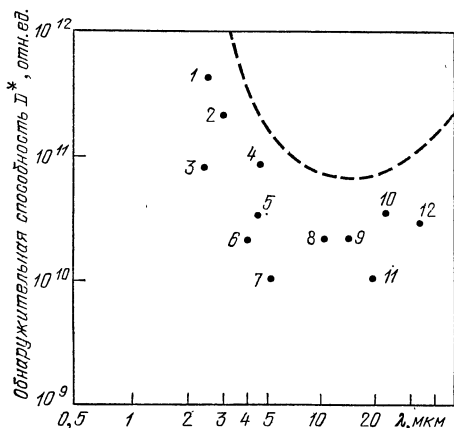


Рис. 54. Обнаружительная способность D^* реальных фоторезисторов в сравнении с обнаружительной способностью идеального фоторезистора, в котором собственные шумы полностью отсутствуют.

Пунктирная кривая — расчетная зависимость $D^*(\lambda_{гр})$ для идеального фоторезистора. Точки — максимальные величины обнаружительной способности у реальных фоторезисторов: 1 — PbS (при температуре рабочего элемента $T_p = 200$ K); 2 — PbS ($T_p = 77$ K); 3 — PbS (300 K); 4 — InSb (77 K); 5 — PbSe (77 K); 6 — PbSe (195 K); 7 — Ge: Au (60 K); 8 — Ge: Hg (27 K); 9 — Ge: Cd (4,2 K); 10 — Ge: Cu (4,2 K); 11 — Si: Sb (4,2 K); 12 — Ge: Zn (4,2 K).

особенно охлаждаемых, не так уж далека от предельных теоретических значений в очень широком диапазоне длин волн.

Глава 9 ЭФФЕКТ ХОЛЛА

Что польза магнита очень велика и прямо изумительна — это настолько хорошо известно, что нет необходимости произносить по этому поводу длинную речь.

Из предисловия Э. Райта к книге В. Гильберта «О магните...» (около 1600 г.)

Не правда ли, образец, показанный на рис. 55, и схема его включения в цепь что-то напоминают? Сравним рис. 55 со схемой, показанной на

рис. 31. Сходство бросается в глаза. В обоих случаях образец имеет форму прямоугольного параллелепипеда. Он снабжен омическими токовыми контактами (1 и 2 на рис. 55), через которые пропускается измеряемый амперметром ток I . В обоих случаях имеются также и потенциальные контакты, на которых измеряется разность потенциалов. Только на рис. 31 их два — A и B , а на рис. 55 — четыре (3, 4, 5 и 6). Используя контакты 5 и 6, мы сможем, следуя примеру, который подал более ста лет назад американский физик Э. Холл, обнаружить и исследовать удивительные явления.

Пропустив через образец ток I и измерив вольтметром напряжение V_{3-4} между потенциальными электродами 3 и 4, можно, как мы уже знаем, определить проводимость полупроводника σ (см. формулу на с. 119). Никто не мешает нам измерить и величину V_{5-6} — напряжение между электродами 5 и 6. Если исследуемый образец однороден, а контакты 3 и 5, 4 и 6 расположены точно друг под другом, то естественно, что $V_{3-4} = V_{5-6}$. Однако, вообще говоря, это равенство выполняется далеко не всегда.

Нередко бывает, что даже при строго одинаковом расстоянии между контактами 3—4 и 5—6 величины V_{3-4} и V_{5-6} не равны. Значит, проводимость в разных частях образца разная! Перемещая потенциальные электроды вдоль поверхности, можно исследовать распределение неоднородностей в образце и (иногда) подсказать технологам, что нужно сделать, чтобы образцы получались однородными.

Но допустим, что измеренные значения V_{3-4} и V_{5-6} равны, образец однороден и контакты 3—5 и 4—6 расположены точно друг под другом. Измерим напряжения V_{3-5} и V_{4-6} . Результат $V_{3-5} = V_{4-6} = 0$ не кажется удивительным. Действительно, глядя на рис. 55, легко сообразить, что контакты 3 и 5, точно так же, как и контакты 4 и 6, расположены на

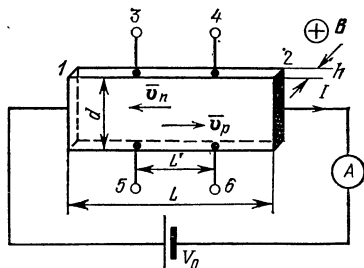


Рис. 55. В перпендикулярном к плоскости рисунка магнитном поле $\vec{B}$ между контактами 3 и 5 и между контактами 4 и 6 возникает э. д. с. Холла.

поверхностях равного потенциала (на эквипотенциальных поверхностях).

А теперь поместим образец в магнитное поле, направленное перпендикулярно рисунку и снова измерим напряжение V_{3-5} (или V_{4-6}). Мы обнаружим, что между этими электродами появилась разность потенциалов! Величины V_{3-5} и V_{4-6} пропорциональны току I , протекающему через образец, и значению индукции магнитного поля B . В этом и состоит эффект, открытый в 1879 г. Холлом и названный его именем.

Почему возникает эффект Холла

Давайте сначала качественно обсудим причину появления холловского напряжения между контактами 3 и 5 (или 4 и 6) при приложении магнитного поля.

Пусть, для определенности, исследуемый образец n -типа и протеканию тока соответствует дрейф заряженных носителей (электронов) в направлении, противоположном направлению тока. На рис. 55 ток через образец течет слева направо. Таким образом, электроны в образце движутся справа налево, от контакта 2 к контакту 1.

В магнитном поле на движущиеся носители действует сила Лоренца. Направление силы можно найти, воспользовавшись правилом левой руки (Физика, 9 класс). Если вы правильно применили это правило, то отогнутый большой палец покажет, что сила Лоренца будет направлена вверх. Электроны под действием силы Лоренца будут натекать на верхнюю грань образца, заряжая ее отрицательно, и оттекать от нижней грани, заряжая ее положительно. Между верхней и нижней гранями образца, между контактами 3—5 и 4—6, возникает холловская разность потенциалов (э. д. с. Холла).

Если образец изготовлен из полупроводника p -типа и ток переносится дрейфом дырок, то при том же направлении тока и магнитного поля сила Лоренца будет направлена в том же направлении — вверх. Однако движению дырок в направлении верхней грани образца соответствует положительный потенциал верхней грани относительно нижней. Холловское напряжение (э. д. с. Холла) будет иметь знак, проти-

воположный знаку холловского напряжения в электронном образце.

По закону холловского напряжения можно судить о типе проводимости образца.

Напряженность холловского поля

Холловское напряжение, возникшее под действием силы Лоренца, препятствует дальнейшему натеканию носителей на верхнюю (при выбранном направлении тока и магнитного поля) грань образца. В установившемся стационарном состоянии сила Лоренца, действующая на носители со стороны магнитного поля $F_L = e\bar{v}B$, будет уравновешена силой $F_H = eE_H$, действующей на носители со стороны поля Холла E_H . Носители перестанут нтекать на верхнюю и оттекать от нижней грани образца и, как и до приложения магнитного поля, будут двигаться параллельно длинным граням параллелепипеда (рис. 55). Таким образом, напряженность холловского поля E_H определится из уравнения

$$e\bar{v}B = eE_H. \quad (70)$$

Здесь $\bar{v} = \mu E_0$ — дрейфовая скорость носителей в поле $E_0 = V_0/L$, V_0 — приложенное к образцу напряжение, L — длина образца (рис. 55).

Отсюда

$$E_H = \mu B E_0. \quad (71)$$

Уравнение (71) показывает, что при заданной величине поля, создаваемого батареей E_0 (поле E_0 часто называют тянущим полем), и заданной величине индукции магнитного поля B величина холловского поля E_H пропорциональна подвижности носителей заряда μ .

Вспомните то место гл. 2, где рассказывалось о драматической истории установления характера электропроводности в Ag_2S . Теперь понятно, почему обнаружение эффекта Холла в сернистом серебре стало сильнейшим аргументом в пользу электронной, а не ионной проводимости Ag_2S . Подвижность ионов обычно в тысячи раз меньше, чем электронная или дырочная подвижность. При чувствительности установок для измерения эффекта Холла (холловских установок),

существовавших во времена Холла и Стрейнца, э. д. с. Холла, соответствующая подвижности ионов, просто не могла быть измерена.

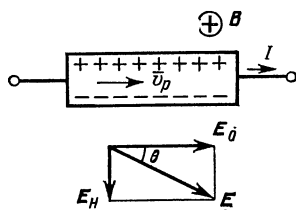


Рис. 56. Напряженность электрического поля E в образце направлена под холловским углом θ к тянущему полю E_0 . Сила, действующая на носители со стороны холловского поля E_H , уравнивается силой Лоренца. (Знак холловского поля на рисунке соответствует дырочной проводимости образца.)

Обратите внимание на необычную ситуацию, которая иллюстрируется рис. 56. Суммарная напряженность электрического поля в образце E равна векторной сумме тянущего поля E_0 и холловского поля E_H . Но носители движутся параллельно полю E_0 , а не суммарному полю E . Почему?

Сила, действующая на электрон со стороны холловского поля, уравнивается силой Лоренца, равной ему по величине и противоположной по направлению. Угол θ между тяну-

щим и полным полем называется холловским углом. Из выражения (71) ясно, что $\operatorname{tg} \theta = \mu B$.

Что измеряется экспериментально

Экспериментатор, изучающий эффект Холла в полупроводнике, измеряет ток I , текущий через образец, индукцию магнитного поля B и напряжения между электродами 3, 4, 5 и 6 (рис. 55). Его (и наша) цель состоит в том, чтобы рассчитать, используя измеренные величины, концентрацию и подвижность носителей в полупроводнике. Понимая теперь физическую картину эффекта Холла, мы без труда сделаем соответствующий расчет.

Сначала, не включая магнитное поле, измерим проводимость образца σ . Для этого следует, как мы знаем, измерить напряжение между электродами 3 и 4 (V_{3-4}) при заданном токе I . Величина σ равняется

$$\sigma = \frac{IL'}{V_{3-4}hd}, \quad (72)$$

(Аналогичные измерения и расчет по формуле (72) проводятся и для контактов 5—6. Это позволяет контролировать однородность образца.)

Включим теперь магнитное поле и измерим при том же токе I через образец величину магнитной индукции B и значение э. д. с. Холла V_H между контактами 3 и 5 (V_{3-5}). А теперь — простой расчет. Холловское напряжение V_H равняется, разумеется, холловскому полю E_H , умноженному на ширину образца d :

$$V_{3-5} = V_H = E_H d = \mu B E_0 d. \quad (73)$$

Значение тянущего поля E_0 известно. Оно просто равняется напряжению между контактами 3 и 4, деленному на расстояние между ними: $E_0 = V_{3-4}/L'$. Используя формулу (72), получаем:

$$E_0 = V_{3-4}/L' = I/h d \sigma. \quad (74)$$

Подставим теперь (74) в (73):

$$V_{3-5} = IB\mu/h\sigma. \quad (75)$$

Или, вспомнив, что $\sigma = en_0\mu$ (см. формулу 17, а),

$$V_{3-5} = IB/en_0h^* \quad (76)$$

Величины V_{3-5} , I , B — измерены. Концентрация носителей в образце, таким образом, определяется как

$$n_0 = IB/ehV_{3-5}. \quad (77)$$

При известных значениях проводимости σ и концентрации n_0 величина подвижности μ определится просто как $\mu = \sigma/en_0$.

Мы уже упоминали в гл. 5, как важно уметь определять величины подвижности и концентрации (а не только проводимости). Использование эффекта Холла представляет собой самый распространенный способ измерения этих величин в полупроводниках.

Применения эффекта Холла

Одно из самых важных применений эффекта Холла состоит в исследовании свойств полупроводников, металлов и некоторых диэлектриков.

*) Может показаться, что между формулами (73) и (76) есть противоречие. В одном случае холловское напряжение зависит только от подвижности носителей, в другом — только от концентрации. В действительности никакого противоречия, конечно, нет. Для того чтобы пропустить через образец заданный ток I , нужно приложить тем большее тянущее поле E_0 , чем меньше концентрация n_0 и подвижность μ (см. формулу (74)).

С его помощью исследуются зависимости концентрации и подвижности носителей в различных материалах от температуры, давления, характера и концентрации введенных примесей. Зависимости, показанные на рис. 19, 20 и 27, получены именно с помощью исследования эффекта Холла.

Современные холловские установки для исследования свойств полупроводников представляют собой сложные инженерные сооружения. Температура в рабочей камере, в которой помещается образец, может автоматически, в соответствии с заданной программой изменяться от нескольких единиц до нескольких сотен кельвинов. Для создания сильного магнитного поля применяются охлаждаемые жидким гелием сверхпроводящие соленоиды, позволяющие получать значения магнитной индукции 8—10 Тл. Чтобы иметь возможность измерять малые значения подвижности ($\leq 10^{-6} \text{ м}^2/(\text{В} \cdot \text{с})$), используют сложные радиотехнические схемы, позволяющие зарегистрировать очень маленькое холловское напряжение на фоне сильных шумов. Результаты измерений непосредственно передаются в ЭВМ, где они обрабатываются по заданным программам.

Таких сложных установок, конечно, немного. Однако общее число более простых холловских установок составляет тысячи и десятки тысяч: эффект Холла используется в тысячах лабораторий для текущего контроля качества полупроводниковых материалов и приборов. Кроме того, и относительно простые установки для измерения эффекта Холла дают возможность изучать многие интересные физические явления.

Помимо обслуживания породившей его физики твердого тела, послужной список эффекта Холла украшают десятки технических применений. Мы обсудим здесь только самые важные и интересные из них *).

Измерение магнитного поля. В современной технике и физике используются магнитные поля в диапазоне значений индукции от $\sim 10^{-6}$ до $\sim 10^3$ Тл. Необходимо уметь измерять значения индукции при тем-

*) Подробнее о технических применениях эффекта Холла можно прочитать в книге: Кобус А., Тушинский Я. Датчики Холла и магниторезисторы. — М.: Энергия, 1971.

пературах от тысячных долей до тысяч кельвинов, в диапазоне частот от постоянного поля до частот $\sim 10^{10}—10^{11}$ Гц, соответствующих СВЧ диапазону.

Датчики Холла (так называются полупроводниковые приборы, действие которых основано на эффекте Холла) широко используются для измерения магнитной индукции. Очень большой диапазон измеряемых магнитных полей и необходимость измерений в широком температурном интервале обуславливают использование для изготовления датчиков Холла самых различных полупроводниковых материалов: Ge, Si, InSb, InAs, HgTe, CdTe, GaAs. Для измерений слабых магнитных полей при низких температурах применяют датчики Холла, изготовленные из материалов с высокой подвижностью и небольшой энергией образования электронно-дырочной пары $\mathcal{E}_g$. Для измерений в области высоких температур — материалы с большим значением $\mathcal{E}_g$.

Принцип измерения магнитного поля с помощью эффекта Холла очень прост: холловское напряжение на образце при заданном токе через датчик прямо пропорционально величине магнитной индукции (см. формулу (76)). Высокая степень линейности (точность, с которой выходной сигнал датчика пропорционален измеряемой величине) является одним из важных преимуществ датчиков Холла перед другими датчиками, измеряющими магнитную индукцию.

Измерение тока. На первый взгляд, использование датчиков Холла для измерения силы тока может показаться занятием совершенно ненужным. Или, по крайней мере, экзотическим. В самом деле, существуют десятки, если не сотни приборов, устройств и схем для измерения тока. В первую очередь, конечно, амперметры самых разнообразных систем, рассчитанные на токи от пикоампер (10^{-12} А) до килоампер (10^3 А). Затем шунты — небольшие сопротивления, включаемые последовательно в ту цепь, ток в которой желательно измерить. Напряжение на шунте пропорционально протекающему току.

Однако в современной технике существует потребность в измерении постоянных токов в десятки и сотни тысяч ампер. Именно такие токи текут, например, в промышленных электролитических установках. Импульсные токи, используемые в современном физическом эксперименте, могут составлять миллионы и де-

сятки миллионов ампер. Да к тому же и иметь сложную зависимость от времени, которую также необходимо изучать. В цепь, по которой текут такие чудовищные токи, никакой посторонней, дополнительной нагрузки в виде шунта или амперметра включать нельзя. Мало того, что такая нагрузка будет потреблять мощность в сотни и тысячи киловатт, она может исказить электрические характеристики изучаемой цепи. В некоторых случаях разрывать цепь, чтобы последовательно включить какое-нибудь приспособление для измерения тока, нельзя по условиям эксплуатации.

Как же поступать в этих случаях?

Измерять ток по величине создаваемого им магнитного поля! (Физика, 9 класс.) А там, где речь идет об измерении магнитного поля, датчики Холла, как мы видели только что, почти всегда очень удобны, а иногда и просто незаменимы.

Время установления холловского напряжения составляет обычно $\sim 10^{-11}$ — 10^{-12} с; это позволяет исследовать даже очень короткие импульсы тока. Высокая линейность холловских датчиков позволяет непосредственно измерять величину тока и зависимость его от времени. Маленькие размеры датчиков позволяют в случае необходимости измерять и распределение тока в пространстве (например, внутри электролитических ванн).

Измерение мощности сверхвысокочастотных электромагнитных волн. Измерение СВЧ мощности с помощью датчиков Холла представляет собой пример очень изящного и простого решения сложной технической задачи. Датчики Холла для измерения СВЧ мощности не нуждаются ни в батарее для создания тянущего поля, ни в магните. Источником и тянущего поля, и магнитного поля служит та же самая электромагнитная волна, мощность которой измеряется!

На рис. 57 показана электромагнитная волна, распространяющаяся в направлении k . Вектор напряженности электрического поля e колеблется в вертикальной плоскости; вектор индукции магнитного поля b — в горизонтальной. На боковых гранях холловского датчика, помещенного в поле электромагнитной волны, возникает э. д. с. Холла

$$u_H = \mu deb. \quad (78)$$

Формула (78) отличается от формулы (73) тем, что для обозначения холловского напряжения u_H , индукции магнитного поля b и напряженности электрического поля e использованы не большие, а маленькие буквы. Сделано это для того, чтобы лишний раз напомнить, что в поле электромагнитной волны эти величины зависят от времени:

$$e = E_m \sin \omega t,$$

$$b = B_m \sin \omega t.$$

Э. д. с. Холла u_H меняется во времени с удвоенной по отношению к электрическому и магнитному полю частотой:

$$u_H \sim E_m B_m \sin^2 \omega t = \frac{E_m B_m}{2} (1 - \cos 2\omega t). \quad (79)$$

По сравнению с другими приборами для измерения СВЧ мощности датчик Холла имеет два важных преимущества. По знаку холловской э. д. с. на датчике можно судить о направлении распространения волны. С изменением распространения волны на противоположное меняется на противоположный и знак холловского напряжения, поскольку когда электромагнитная волна меняет направление, то при неизменном направлении вектора e направление вектора b меняется на противоположное. Второе преимущество холловских датчиков перед термисторными, болометрическими и многими другими типами измерителей мощности СВЧ колебаний состоит в том, что э. д. с. Холла чувствительна к разности фаз между электрическим и магнитным полем в электромагнитной волне. При распространении волны в воздухе разность

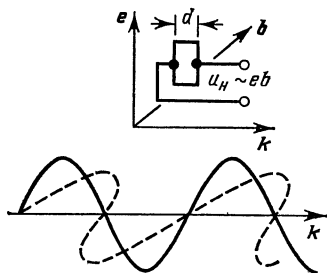


Рис. 57. В электромагнитной волне, распространяющейся в направлении вектора k , вектор напряженности электрического поля e колеблется в вертикальной плоскости (сплошная кривая), а вектор магнитной индукции b — в горизонтальной (пунктир). На боковой поверхности образца, помещенного в поле электромагнитной волны, возникает э. д. с. Холла u_H (вставка).

фаз всегда равна нулю. Электрическое и магнитное поле меняются во времени *синфазно*, т. е. достигают максимума и проходят через нуль в одни и те же моменты времени (как показано на рис. 57). Но в волноводах и резонаторах, где обычно и используются СВЧ холловские датчики, ситуация часто бывает иной. Между электрическим и магнитным полем возникает разность фаз:

$$e = E_m \sin \omega t, \quad b = B_m \sin(\omega t + \varphi),$$

аналогично тому, как возникает разность фаз между током и напряжением при работе на нагрузку, содержащую реактивные элементы — емкость или индуктивность (Физика, 10 класс). Среднее за период холловское напряжение зависит от сдвига фаз между магнитным и электрическим полем. Можно показать (сделайте это сами), что величина среднего за период холловского напряжения пропорциональна $\cos \varphi$, где φ — разность фаз между магнитным и электрическим полями. Мощность, поглощаемая в нагрузке, также пропорциональна $\cos \varphi$. Поэтому холловские датчики СВЧ мощности позволяют непосредственно контролировать мощность, поглощаемую нагрузкой.

Используя полупроводники с высокой подвижностью μ , удастся сконструировать достаточно чувствительные СВЧ холловские приемники. С помощью антимоноид-индиевых приемников можно добиться чувствительности, соответствующей 0,1 мкВ холловского напряжения на 1 мВт поглощаемой в нагрузке СВЧ мощности.

Глава 10

ЭФФЕКТ ГАННА

Вот вам бог — а вот порог.

Около двадцати лет назад, в начале 60-х годов, американский физик Джон Ганн решил исследовать вольтамперные характеристики сравнительно новых тогда полупроводников: арсенида галлия (GaAs) и фосфида индия (InP) в сильных электрических полях. Он обнаружил удивительный эффект, названный его именем.

В чем состоит эффект Ганна

Картина, которую наблюдал Ганн, показана на рис. 58. К образцам, представлявшим собой просто кусочки GaAs или InP с двумя омическими контактами, прикладывались импульсы напряжения. Если амплитуда импульса V_n была меньше, чем некоторая пороговая величина V_t — разная для разных образцов, то все было понятно. Ток через образец нарастал, пока нарастало напряжение, оставался постоянным, когда напряжение не менялось, и уменьшался, когда напряжение падало. Этот случай показан пунктиром на рис. 58, а, б. Меняя амплитуду импульса напряжения и измеряя амплитуду соответствующего этому напряжению тока, можно было построить вольтамперную характеристику образца. Она приблизительно соответствовала закону Ома.

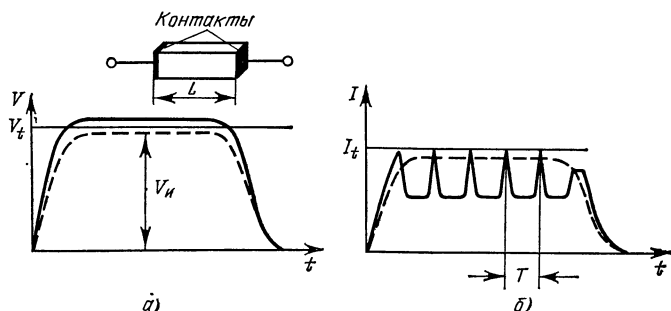


Рис. 58. Зависимость напряжения на образце от времени (а) и тока через образец от времени (б) в экспериментах Ганна. Пунктиром показан случай, когда амплитуда импульса напряжения V_n меньше, чем критическое значение V_t . Сплошной кривой — случай, когда $V_n > V_t$. На вставке показана форма образца.

Но стоило напряжению превысить значение V_t (индекс t — от английского слова threshold — порог), как ток при постоянном напряжении начинал периодически изменяться со временем. Этот случай показан на рис. 58 сплошными линиями. Проведя измерения на образцах различной длины L , Ганн установил, что пороговое напряжение пропорционально длине образцов: $V_t \approx E_t L$, где E_t — приблизительно одинаковое для всех образцов значение напряженности поля. Для GaAs величина E_t составляла примерно $3,2 \cdot 10^5$ В/м.

Период колебаний T был также пропорционален длине образца L . Он равнялся приблизительно $T \approx \approx 2L/v_t$, где v_t — скорость электронов при пороговом поле E_t . Поскольку закон Ома выполнялся приближенно вплоть до порогового поля E_t , пороговая скорость $v_t \approx \mu E_t \approx 2 \cdot 10^5$ м/с (подвижность электронов μ в образцах GaAs, имевшихся в распоряжении Ганна, составляла при комнатной температуре $0,5—0,7$ м²/(В·с)).

В 1963 г. Ганн описал результаты своих наблюдений в короткой заметке, которая привлекла к себе внимание десятков физиков и инженеров во всем мире. Физикам представлялось интересным объяснить это удивительное явление. Инженеры сразу поняли, что открытие эффекта Ганна означает появление нового полупроводникового генератора, который может работать в области сверхвысоких частот.

В самом деле, при длине $L \sim 2—10$ мкм частота генерируемых диодом колебаний f составляет

$$f = \frac{1}{T} \approx \frac{v_t}{2L} \sim (1 \div 5) \cdot 10^{10} \text{ Гц.} \quad (80)$$

Такой диапазон частот представляет большой интерес для радиоэлектроники, в частности, для радиолокации и многочисленных систем связи — военных и гражданских.

Хотя десятки квалифицированных специалистов размышляли над природой открытого Ганном явления, прошел целый год, пока оно получило объяснение.

Что происходит в образце

В течение этого года Ганн проделал замечательное по уровню экспериментального мастерства исследование открытого им эффекта. Он установил, что когда напряжение на образце превышает пороговое значение V_t , то характер протекания тока и распределение напряженности электрического поля в образце совершенно изменяется. При $V < V_t$ ток в образце представляет собой поток электронов, текущих от того контакта образца, к которому приложен минус источника напряжения (катода), к контакту, к которому приложен плюс (аноду). Электрическое поле распределено вдоль образца однородно (рис. 59, пункт

тир). Как только напряжение на образце превышает пороговое значение V_t , у катода возникает узкая область очень сильного поля — «домен» *). Хотя ширина этой области невелика, но поле в ней настолько сильное, что на домene падает значительная часть приложенного к образцу напряжения (рис. 59). Домен движется от катода к аноду со скоростью, приблизительно равной $v_t/2$. Достигнув анодного контакта, домен разрушается. Немедленно после этого у катода зарождается новый домен и все повторяется сначала.

Легко понять, как связано периодическое движение домена по образцу с колебаниями тока, которые Ганн наблюдал в своих первых экспериментах.

Представим себе, что к образцу приложено напряжение, несколько меньшее, чем пороговое значение V_t (рис. 58, а, пунктир). Тогда по образцу течет ток, чуть меньший I_t , напряженность поля распределена вдоль образца однородно и по величине меньше, чем пороговое значение $E_t = V_t/L$ (рис. 59, пунктир). Выше уже упоминалось, что пока напряжение V меньше V_t , приближенно выполняется закон Ома. А это значит, что скорость электронов в образце чуть меньше, чем значение $v_t \approx \mu E_t$.

Пусть мы немного увеличили приложенное к образцу напряжение, и оно стало равным V_t . В образце возник домен, и значительная часть напряжения упала на домene. Но это значит, что на долю остальной

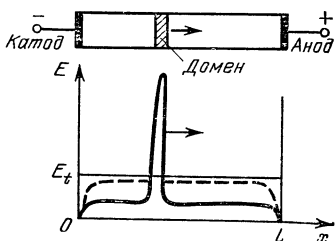


Рис. 59. Распределение электрического поля в образце. Пунктиром показан случай, когда $V_n < V_t$. Сплошной кривой — $V_n > V_t$. Стрелка указывает направление распространения домена: от катода к аноду. В верхней части рисунка схематически показано движение домена по образцу.

*) Слово «домен» происходит от латинского dominicus (хозяинский, господский). В средние века доменом называлась часть личного владения крупного феодала. В физике слово «домен» используется для обозначения области, свойства которой резко отличаются от свойств окружающего ее материала. Существуют ферромагнитные, сегнетоэлектрические домены, домены сильного и слабого поля ... На рис. 60 показан домен сильного поля.

части образца, где домена нет, приходится теперь меньшее напряжение. Поле вне домена упало (ср. пунктирную и сплошную кривые на рис. 59). Упало и значение скорости электронов, пропорциональное полю, а следовательно, и ток, текущий через образец.

Пока домен движется вдоль образца, ток не меняется. Когда домен, достигнув анода, начинает разрушаться, падение напряжения на нем уменьшается. При этом поле в остальной части образца растет и, следовательно, растет ток. Но, как только поле в районе катода образца станет равным пороговому значению E_t , в образце начнет формироваться новый домен и ток снова начнет уменьшаться. Пока напряжение, приложенное к образцу, превышает пороговое значение V_t , этот процесс периодически повторяется (рис. 58, б, сплошная кривая).

Удивительная зависимость скорости электронов от электрического поля

Замечательные эксперименты Ганна позволили увидеть, что происходит в образце, когда в нем возникают колебания тока. Однако они не отвечали на вопрос, почему в образце появляется

домен; в чем физическая причина образования домена.

На этот вопрос в 1964 г. сумел ответить американский физик Герберт Кремер. Он напомнил, что незадолго до открытия Ганна на основании теоретических расчетов было показано, что в арсениде галлия, фосфиде индия и некоторых других полупроводниках зависимость дрейфовой скорости электронов от напряженности поля должна иметь вид, показанный на рис. 60.

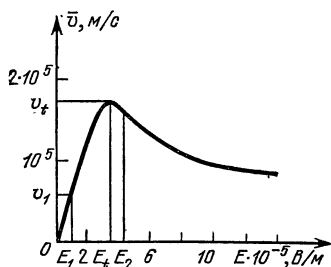


Рис. 60. Зависимость дрейфовой скорости электронов $\bar{v}$ от напряженности электрического поля E для GaAs. Аналогичная зависимость характерна для многих полупроводников (InP, GaInSb, InGaAsP и др.).

В определенном диапазоне значений поля E дрейфовая скорость электронов $\bar{v}$ падает с увеличением напряженности поля. Как мы убедимся в следующем

параграфе, такой вид зависимости $\bar{v}(E)$ действительно позволяет объяснить все наблюдавшиеся Ганном явления. Но как такая характеристика $\bar{v}(E)$ возникает?

Два важных обстоятельства. С такой удивительной ситуацией нам еще не приходилось сталкиваться. Изучая термисторы (гл. 6), мы встречались, правда, с очень необычными, S -образными вольтамперными характеристиками. Однако такие необычные зависимости тока от напряжения в термисторах возникали за счет джоулева разогрева полупроводника протекающим током. При этом менялась концентрация свободных носителей в образце.

Обдумывая ход зависимости $\bar{v}(E)$, показанный на рис. 60, важно помнить о двух обстоятельствах. Во-первых, эта зависимость является изотермической. Она рассчитывается (и измеряется!) в условиях, когда температура образца постоянна. Во-вторых, показанная на рис. 60 кривая является не зависимостью тока от напряжения, а зависимостью дрейфовой скорости электронов $\bar{v}$ от напряженности электрического поля E . Как мы хорошо знаем, $\bar{v} = \mu E$; концентрация носителей здесь ни при чем — в зависимость, связывающую $\bar{v}$ и E , она просто не входит. Если дрейфовая скорость электронов не пропорциональна напряженности поля, то это может означать только одно: с ростом напряженности поля от величины E начинает зависеть подвижность носителей μ . О том, что такая ситуация возможна, мы уже упоминали вкратце в гл. 5 (см. раздел «Горячие электроны»).

Оба эти обстоятельства заслуживают того, чтобы остановиться на них поподробней.

Спасительные импульсы. Если рассчитать, каким значениям напряженности поля E соответствуют напряжения V , при которых вольтамперная характеристика термистора начинает за счет джоулева разогрева заметно отличаться от линейного закона (рис. 37), то окажется, что в типичном случае величина E составит $\sim 10^3 - 10^4$ В/м. Мощность P , выделяющаяся в образце, пропорциональна квадрату напряженности поля E . Действительно, $P = IV = jESL$, где S — площадь поперечного сечения образца, L — его длина, j — плотность тока. Произведение SL равняется объему образца. Поделив мощность P на объем образца, получим, что мощность, выделяющаяся в еди-

нице объема, или плотность мощности $P_0 \equiv jE = \sigma E^2$ (см. формулу (15)). Из рис. 60 видно, что величины E , с которыми нам предстоит иметь дело, составляют $\sim 3 \cdot 10^5 - 10^6$ В/м. Выделяющаяся в образцах мощность в сотни тысяч, в миллионы раз больше, чем мощность, вызывающая заметный разогрев полупроводниковых материалов в термисторах!

Каким же образом можно ухитриться изучать характеристики полупроводников в столь сильных полях в изотермических условиях, не разогревая образец?

Очень просто. Следует прикладывать напряжение к образцу в виде достаточно коротких импульсов!

Количество тепла, выделяющееся в образце, пропорционально длительности импульса. Если использовать короткие импульсы и прикладывать их к образцу редко, так, чтобы в промежутке между импульсами выделившееся в образце тепло успело полностью рассеяться, можно достичь очень сильных электрических полей, не нагревая образец *).

Если исследуются образцы с большой концентрацией электронов, а напряженность поля, при которой возникает изучаемый эффект, велика, то иногда возникает необходимость в использовании импульсов наносекундной ($\sim 10^{-9}$ с) и даже субнаносекундной ($10^{-10} - 10^{-11}$ с) длительности. Для работы с такими короткими импульсами необходимо уникальное оборудование и большое экспериментальное мастерство.

А все-таки они горячие! Теперь мы знаем, как достичь больших значений поля, не изменив температуры кристалла. Если импульс напряжения, пусть даже очень большой амплитуды, достаточно короток — кристаллическая решетка полупроводника не успеет нагреться.

*) Рассчитайте, импульсы напряжения какой длительности необходимо использовать, чтобы, не нагревая образец больше чем на 5 К, достичь полей $\sim 10^6$ В/м. При расчете принять, что вплоть до значений $E = 10^6$ В/м выполняется закон Ома. Выделившееся за время импульса тепло рассеяться не успевает и полностью идет на нагрев образца. Исследованию подлежит образец GaAs с концентрацией носителей $n_0 = 10^{16}$ см $^{-3}$. Для определения подвижности μ воспользуйтесь табл. II. Удельную теплоемкость C_V принять равной $1,5 \cdot 10^6$ Дж/(м $^3 \cdot$ К). Ответ: $t \sim 5 \cdot 10^{-9}$ с.

А как обстоит дело с температурой носителей заряда — электронов (или дырок)?

Если подвижность носителей и амплитуда приложенного поля достаточно велики, температура электронов может составить тысячи и десятки тысяч градусов, несмотря на то, что температура кристалла не увеличилась и на градус за время импульса!

На первый взгляд это кажется непостижимым. Но первый взгляд уже не один раз нас подводил. В действительности, такая ситуация естественна, и понять в чем дело совсем не трудно.

Рассмотрим простую аналогию. Мы зашли в большую холодную комнату. В комнате есть печка; мы закладываем в нее дрова, разжигаем. Через полчаса дверцы печки светятся ровным малиновым светом, от нее дышит тепло. И в этот момент... мы заливаем огонь. Стало ли в комнате теплее? Практически нет. Разве что на долю градуса. А ведь внутренность печки была раскалена. Ее температура составляла сотни градусов!

В этом житейском примере легко разберется каждый. Количество тепла, которое необходимо, чтобы раскалить печку, относительно невелико. Теплостойкость печки мала. Поэтому и прогревается печка быстро. Теплостойкость комнаты гораздо больше. Вот если бы печка топилась несколько часов...

Но ведь то же самое происходит и с электронами в кристалле. Роль дров, поставляющих тепло, играет здесь электрическое поле, роль быстро раскаляющейся печки — электроны. Роль большой выстуженной комнаты — кристалл. Только времена, с которыми приходится иметь дело, совсем другие.

В типичных полупроводниках температура электронов устанавливается после приложения электрического поля очень быстро, через $\sim 10^{-12}$ с. Разогретые полем электроны, сталкиваясь с кристаллической решеткой, передают ей энергию, и кристалл нагревается. Не «погаси мы печку», не оборви импульс напряжения — кристалл мог бы нагреться до очень высокой температуры. (Даже если учесть, что как только температура кристалла поднимется, — он начнет охлаждаться за счет отвода тепла в окружающую среду. Комната ведь тоже охлаждается за счет теплопроводности стен, крыши и потолка.) Но... мы используем короткие импульсы и получаем возмож-

ность изучать горячие электроны в холодном кристалле.

Почему уменьшается подвижность. В разделе «Горячие электроны» уже упоминалось, что у горячих электронов параметры, определяющие подвижность носителей, — время рассеяния τ_0 и эффективная масса m^* , зависят от напряженности электрического поля. Насколько сильны эти зависимости и какой вид они имеют, определяется типом полупроводника, температурой кристаллической решетки, содержанием примесей и температурой свободных носителей, зависящей от напряженности приложенного электрического поля.

В германии и кремнии, например, при комнатной температуре основной вклад в изменение подвижности носителей с ростом электрического поля дает зависимость от напряженности поля E времени рассеяния τ_0 . Дело в том, что в этих материалах электрон или дырка, получившие от электрического поля определенную, достаточно большую энергию $\mathcal{E}_0$, отдают ее решетке полностью при каждом столкновении. Убедимся в том, что в таких условиях время рассеяния τ_0 будет обратно пропорционально напряженности поля E .

Требуемая энергия $\mathcal{E}_0$ определяется эффективной массой носителя m^* и его скоростью v :

$$\mathcal{E}_0 = m^* v^2 / 2. \quad (81)$$

В свою очередь, скорость определяется ускорением $a = eE/m^*$ и временем между столкновениями τ_0

$$v = eE\tau_0/m^*. \quad (82)$$

Подставляя (82) в (81), получим:

$$\tau_0 = \sqrt{\frac{2m^*\mathcal{E}_0}{e^2}} \cdot \frac{1}{E}. \quad (83)$$

Рис. 61. Зависимости дрейфовой скорости электронов $\bar{v}$ от напряженности электрического поля E для Ge (кривая 1) и Si (кривая 2).

также будет обратно пропорциональна напряженности поля E . А дрейфовая скорость электронов $v_d =$

$= \mu E$ будет величиной постоянной, не зависящей от напряженности поля.

Именно это и наблюдается в действительности.

На рис. 61 показаны зависимости дрейфовой скорости электронов от напряженности электрического поля для германия и кремния при температуре решетки 300 К. В достаточно сильном электрическом поле дрейфовая скорость насыщается, перестает зависеть от поля. Качественно такой же вид имеют в Ge и Si и зависимости дрейфовой скорости дырок от напряженности поля.

В GaAs, как видно из рис. 60, дрейфовая скорость электронов падает с ростом напряженности поля. Для того чтобы это стало возможным, необходимо, чтобы подвижность носителей с возрастанием поля уменьшалась еще быстрее, чем в Ge и Si. В GaAs, InP и некоторых других полупроводниках именно такая ситуация и реализуется. В этих материалах не только время рассеяния горячих электронов уменьшается с ростом напряженности поля, но кроме того, с увеличением поля растет эффективная масса электронов m^* .

С изменением эффективной массы носителей при воздействии внешних условий мы уже встречались. В тензорезисторах (гл. 7) эффективная масса носителей изменяется с давлением из-за изменения взаимного расположения атомов решетки. При этом меняются поля, создаваемые ионами решетки и валентными электронами атомов, между которыми пролетают свободные носители.

В сильном внешнем электрическом поле, где в GaAs наблюдается падение дрейфовой скорости с ростом поля, взаимное расположение атомов решетки не меняется. Природа изменения эффективной массы с полем иная. Дело в том, что взаимодействие между полями атомов решетки и горячими носителями может существенно отличаться от взаимодействия между атомами решетки и равновесными, не разогретыми полем электронами и дырками. Ничего удивительного в этом нет. Вспомним, например, как меняется характер взаимодействия между ионом примеси и свободным носителем в зависимости от скорости электрона или дырки (рис. 28). Хотя взаимодействие между периодическим полем кристаллической решетки и свободными носителями имеет иную природу, чем

примесное рассеяние, неудивительно, что поля, создаваемые ионами кристаллической решетки, будут несколько иначе действовать на горячие носители. А ведь эффективная масса электрона или дырки (напомним об этом еще раз!) вовсе не равна массе свободного электрона в вакууме m_0 . Эффективная масса электрона в кристалле m^* может в десятки раз отличаться от величины m_0 , и значение m^* целиком определяется взаимодействием между электроном (дыркой) и периодическим полем кристаллической решетки.

В GaAs, InP и многих других полупроводниках эффективная масса горячих электронов возрастает с увеличением напряженности электрического поля E так резко, что подвижность уменьшается с ростом напряженности электрического поля быстрее, чем по закону $1/E$. А это значит, что при определенных значениях поля дрейфовая скорость $\bar{v}$ будет уменьшаться с ростом напряженности поля E . На зависимости $\bar{v}(E)$ появится падающий участок (рис. 60).

В сущности, все это довольно просто

Как же связаны явления, которые наблюдал Ганн, с зависимостью $\bar{v}(E)$, показанной на рис. 60?

Для того чтобы ответить на этот вопрос, рассмотрим сначала очень простую задачу, на первый взгляд не имеющую никакого отношения ни к эффекту Ганна, ни к необычным свойствам горячих электронов.

Флуктуация поля на омическом участке зависимости $\bar{v}(E)$ рассасывается. Пусть к образцу приложено небольшое напряжение $V_1 < V_t$. При этом поле вдоль образца распределено однородно и напряженность $E_1 = V_1/L$. Дрейфовая скорость электронов $\bar{v}_1$ равняется μE_1 , и ток представляет собой поток электронов, текущих со скоростью v_1 от катода к аноду.

Предположим теперь, что где-то в образце случайно образовался небольшой участок, где напряженность поля чуть-чуть больше, чем значение E_1 . За счет хаотического характера движения электронов в кристалле маленькие отклонения напряженности поля от среднего значения E_1 (флуктуации поля) возникают непрерывно, и наша задача — проследить за тем, что произойдет с такой флуктуацией, когда она возникнет.

Если на каком-то участке образца поле стало несколько больше, чем в остальном образце, то это значит, что на границах этого участка скопились электрические заряды: отрицательный со стороны катода и положительный со стороны анода (рис. 62). Поле этого объемного избыточного заряда, наложившись на поле внешнего источника, немного усилило поле внутри выделенного участка.

Наличие объемного заряда означает, что слева от выделенного участка скопился небольшой избыток, а справа — образовался небольшой дефицит электронов. Что произойдет с такой флуктуацией поля дальше?

Поскольку мы рассматриваем случай, когда поле E_1 меньше, чем E_t , то скорость электронов пропорциональна полю (рис. 60). Следовательно, электроны внутри флуктуации движутся быстрее, чем электроны в остальной части образца. Это значит, что флуктуация будет «убегать» от избыточных электронов слева и «догонять» тот участок справа, где электронов не хватает. В результате флуктуация поля рассосется.

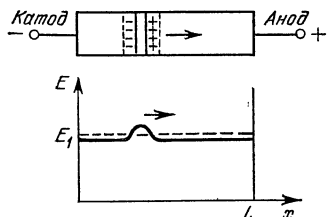


Рис. 62. Распределение электрического поля в образце: без флуктуации (пунктир), с флуктуацией (сплошная кривая). В верхней части рисунка схематически показан объемный заряд на границах флуктуации.

Флуктуация поля на падающем участке зависит от $\bar{v}(E)$ нарастает и превращается в домен. Вернемся теперь к эффекту Ганна и посмотрим, что произойдет с флуктуацией поля, когда к образцу приложено поле E_2 , несколько большее, чем пороговое поле E_t (рис. 60). Пусть вначале поле также распределено вдоль образца однородно. Пусть снова на границах какого-либо участка возник объемный заряд, отрицательный со стороны катода и положительный со стороны анода (флуктуация поля). Поле внутри этого участка больше, чем в остальном образце. Но теперь на падающем участке зависимости $\bar{v}(E)$ большему полю E соответствует меньшая дрейфовая скорость! Флуктуация движется медленнее, чем электроны слева и справа от нее.

В результате слева к уже имевшимся в избытке электронам притекут, догоняя флуктуацию, новые электроны. Справа, где наблюдался дефицит электронов, их станет еще меньше: электроны «убегут» от замедлившейся флуктуации и объемный заряд станет больше. Увеличится поле внутри флуктуации. Но увеличению поля при $E > E_t$ соответствует уменьшение дрейфовой скорости — в результате флуктуация станет двигаться еще медленнее. Еще больше электронов накопится слева от флуктуации и еще меньше их станет справа. Поле внутри флуктуации станет еще больше и т. д.

Итак, в условиях, когда дрейфовая скорость электронов уменьшается с ростом поля, однородное распределение поля неустойчиво по отношению к малой флуктуации поля: раз возникнув, такая флуктуация начинает неудержимо нарастать.

Когда прекратится рост объемного заряда на границах флуктуации и, следовательно, ее рост? Для этого, очевидно, необходимо, чтобы скорость электронов внутри флуктуации сравнялась со скоростью электронов вне ее.

Пока избыточное поле внутри флуктуации мало, на ней падает маленькое напряжение и ее присутствие в образце практически не сказывается на скорости электронов вне флуктуации (рис. 62). Однако, по мере нарастания поля внутри флуктуации, на ней начинает падать все большее и большее напряжение. Так как полное приложенное к образцу напряжение постоянно, то на долю остальной части образца напряжения остается меньше. Поле вне «флуктуации» падает. В последней фразе слово «флуктуация» взято в кавычки. Действительно, как-то неудобно называть флуктуацией — словом, по смыслу означающим «небольшое случайное отклонение», такое образование, которое поглощает заметную долю приложенного к образцу напряжения. Такая сильно выросшая флуктуация и есть домен. Вне домена (в отличие от ситуации внутри домена) скорость электронов падает с уменьшением поля, так как поле вне домена всегда меньше E_t . Таким образом, когда падение напряжения на домене велико, уменьшается скорость электронов вне домена. Когда скорость электронов вне домена сравнивается со скоростью домена, его нарастание прекращается. По образцу от катода

к аноду «бежит» неизменный по форме (стабильный) участок сильного поля — домен, который и наблюдался Ганном в его экспериментах.

Максимальное поле в ганновском домене E_m может быть очень большим. При среднем поле на образце $E_t = V_t/L = 3,2$ кВ/см величина E_m может составлять 20—300 кВ/см.

Применения эффекта Ганна

За годы, прошедшие со времени открытия Ганном эффекта возникновения колебаний тока в GaAs и InP, в научных журналах опубликовано более 2,5 тысяч работ, посвященных исследованию эффекта Ганна и различных связанных с ним физических явлений.

Эффект Ганна наблюдался в настоящему времени в более чем двадцати полупроводниковых соединениях: антимониде индия (InSb), антимониде мышьяка (InAs), в германии, подвергнутом сжатию, в кремнии, охлажденном до низких температур, в тройных полупроводниковых соединениях: GaAlAs, InGaSb, InAsP, в четверном соединении InGaAsP и многих других полупроводниках.

В каждом из этих полупроводников исследователи стараются теоретически рассчитать и экспериментально измерить зависимость дрейфовой скорости электронов от напряженности электрического поля $\bar{v}(E)$. (Как мы видели, эта зависимость играет в эффекте Ганна определяющую роль.) Зная $\bar{v}(E)$, можно рассчитать, как зависят параметры домена: поле в домене, размер и скорость движения домена от длины образца, концентрации электронов в нем и приложенного напряжения.

Быстро движущийся домен сильного поля может быть использован для модуляции световых пучков, для генерации довольно мощных акустических потоков очень высокой частоты (ультразвука), для кодирования и декодирования информации в системах связи и для многих других применений.

Однако основное применение ганновских диодов — генераторы сверхвысокой частоты *).

*) Диодом в современной радиоэлектронике принято называть любой элемент, служащий для генерации или усиления сигналов и имеющий два вывода.

Когда отраженный от цели импульс радиолокатора возвращается к антенне радарной установки (Физика, 10 класс), его необходимо усилить. Схема чувствительного современного усилителя обязательно включает в себя сравнительно маломощный, но очень стабильный и малошумящий генератор (гетеродин). Ганновские диоды успешно работают в качестве гетеродинов во многих современных радиолокаторах (рис. 63, а).

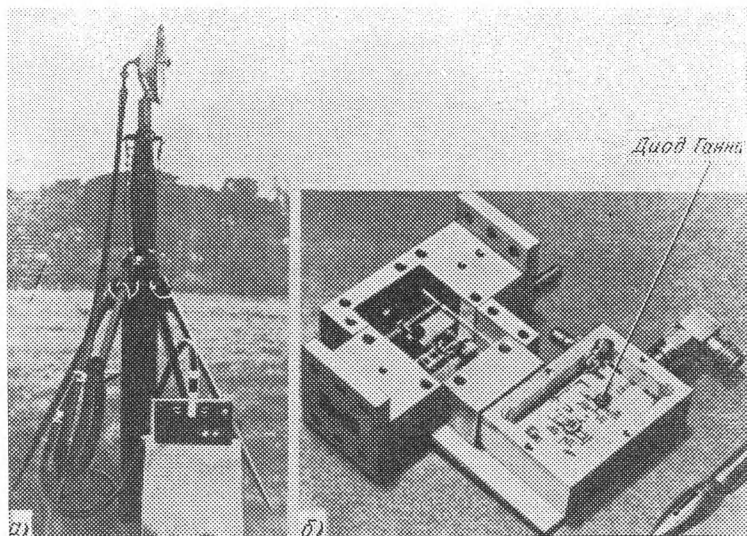


Рис. 63. Применения диодов Ганна.

а) Переносная система радиотелефонной связи между передовыми позициями и штабом батальона (Англия). В качестве усилителя и генератора используется ганновский диод. Система работает на частоте 8 ГГц. Дальность действия — до 40 км; б) элемент портативного узла системы автоматического контроля уличного движения (Япония). Система работает на частоте 20 ГГц. В качестве приемника и передатчика используются диоды Ганна.

Каждый год в мире в результате автомобильных катастроф гибнут десятки тысяч людей. Одна из основных причин аварий — превышение дозированной скорости. В городе за соблюдением предельной скорости следят системы автоматического контроля уличного движения, снабженные специальным радиолокатором, автоматически фиксирующим превышение скорости. Передатчиками в таких радиолока-

торах часто служат диоды Ганна (рис. 63, б). В Швеции с использованием ганновских диодов разработана специальная дорожная система «ложный полицейский». Вдоль шоссе расставлены тысячи небольших металлических будок так, чтобы из каждой будки была видна следующая. В одной из нескольких десятков будок установлен автоматический радиолокатор. Если в пределах видимости локатора машина превысит дозволенную скорость, по сигналу локатора срабатывает фотокамера и зафиксирует номер автомобиля. Затем водителю будет вручено требование уплатить штраф.

В какой именно будке установлен локатор, водитель не знает (локатор переносится с места на место дорожной полицией). Поэтому он вынужден соблюдать дозволенную скорость на всем пути.

С каждым годом все более широкое распространение получает телевизионное вещание через искусственные спутники Земли. Для того чтобы телевизионный сигнал с Земли достиг спутника, этот сигнал преобразовывают, повышая частоту сигнала в десятки и сотни раз. Затем такой же сверхвысокочастотный сигнал транслируется со спутника, принимается наземными трансляционными станциями, снова преобразуется в обычный телевизионный сигнал и в таком виде передается на антенны наших телевизоров.

Расчеты показывают, что гораздо удобнее и выгоднее принимать телевизионные передачи со спутников прямо на антенны индивидуальных телевизоров, а преобразование СВЧ сигнала в обычный телевизионный осуществлять с помощью небольшой приставки в каждом телевизоре. Идеальный элемент для такого преобразователя — простой, стабильный и малолшумящий диод Ганна.

Так что, быть может, недалек день, когда диод Ганна появится, практически, в каждом доме.

ЗАКЛЮЧЕНИЕ

В Мировом океане науки один из самых обширных и бурных океанов — океан Физики. Море Физики Твердого Тела, принадлежащее этому океану, довольно хорошо исследовано и подробно изучено. Но не нужно обольщаться: как свидетельствует беспристрастная и точная статистика, больше всего крушений происходит не в таинственном Бермудском треугольнике и не у грозных ледовых стен Антарктиды, а у обжитых и густонаселенных берегов Ла-Манша. Море Физики Твердого Тела богато подводными течениями, рифами, мелями, а нередко в нем открываются совершенно неожиданные глубины.

Мы посетили один из заливов Физики Твердого Тела — залив Полупроводников. В какой мере наше знакомство с заливом может считаться исчерпывающим?

Мы прошли ряд наиболее популярных туристских маршрутов, проложенных вдоль берегов залива, совершили облет залива на вертолете и сделали несколько прогулочных рейсов на катере, выбрав для круизов хорошую солнечную погоду.

Для первого знакомства этого, может быть, достаточно. Однако, если путешествие оставило в памяти приятное впечатление, стоит вспомнить, что на очень интересную и обширную часть залива мы только мельком взглянули с вертолета. Пройти по берегу к одной из самых важных и интересных областей — Диодам и Транзисторам нам помешал в спортивном отношении очень интересный, а потому нелегкий Барьер $p-n$ -перехода. Другой барьер — Барьер Металл — Полупроводник преградил нам путь к Диодам Шоттки, Полевым Транзисторам, Приборам с зарядовой связью и другим важным областям.

О многих сказочно красивых, интересных и загадочных уголках мы даже не упомянули. Физика полупроводников сделалась наукой такой обширной, а при-

ложения этой науки — такими разнообразными, что целой жизни может не хватить на подробное знакомство с заливом Полупроводников.

С другой стороны, в других заливах, морях и океанах ничуть не меньше красивых мест, подводных течений, острых коралловых рифов и отлогих песчаных пляжей. Вполне возможно, что путешествие по ним привлечет вас больше.

Нужно помнить только, что Мировой океан науки — един. И бутылка, брошенная в заливе Полупроводников, может быть найдена в океанах Биология, Химия и даже в океане Общественные Науки. А в залив Полупроводников заплывает не только планктон, но и довольно крупные рыбы и морские звери из самых разных морей и океанов.

Какое бы море вы ни выбрали — счастливого вам плавания и семь футов под килем!

Михаил Ефимович Левинштейн
Григорий Соломонович Симин

ЗНАКОМСТВО С ПОЛУПРОВОДНИКАМИ

(Серия: Библиотечка «Квант»)

Редактор *Л. А. Панюшкина*
Техн. редактор *Л. В. Лихачева*
Корректоры *Т. С. Вайсберг, Л. С. Сомова*

ИБ № 12526

Сдано в набор 10.08.83. Подписано к печати 11.03.84. Т-07278. Формат 84×118¹/₃₂.
Бумага книжно-журнальная. Литературная гарнитура. Высокая печать.
Усл. печ. л. 12,6. Усл. кр.-отт. 13,02. Уч.-изд. л. 12,98. Тираж 150 000 экз.
Заказ № 759. Цена 40 коп

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ленинградская типография № 2 головное предприятие ордена Трудового
Красного Знамени Ленинградского объединения «Техническая книга»
им. Евгении Соколовой Союзполиграфпрома при Государственном комитете
СССР по делам издательств, полиграфии и книжной торговли. 198052,
г. Ленинград, Л-52, Измайловский проспект, 29

40 коп.
