1966 г. Май *Том 89, вып. 1*

УСПЕХИ ФИЗИЧЕСКИХ НАУК

ФИЗИКА НАШИХ ДНЕЙ

537.311.33

ЭФФЕКТ ГАННА *)

Дж. Ганн

ВВЕДЕНИЕ

Необъясненные эффекты придают особый интерес работе физика, занимающегося твердым телом. Автор, именем которого назван один из таких эффектов, обнаружил, что сильное электрическое поле, приложенное к маленькому кристаллу арсенида галлия, вызывает колебания тока, частота которых превышает 10^9 *гц*. Это достаточно высокая частота, чтобы взволновать воображение людей, занимающихся электроникой; подобные устройства могут оказаться более пригодными для систем связи, чем электронные лампы или транзисторы. Явления такого рода привлекают внимание и теоретиков: в данном случае они хотели бы выяснить причины возникновения колебаний. Теоретический анализ привел к предположению о наличии необычного механизма, напоминающего ударную волну, с помощью которого можно объяснить уменьшение подвижности электронов при изменении их энергии. Такой подход может быть плодотворным не только для создания качественно новых электронных устройств, но и для рассмотрения электронных явлений в кристаллах.

М. Вольф

Приблизительно в 1952 г. меня впервые стали интересовать необычайные явления, возникающие в полупроводниках под действием сильных электрических полей. В это время я, инженер-электротехник, заинтересовался физикой. Я разрабатывал способы изготовления германиевых точечных диодов в Британской компании счетных машин. Из этой работы ничего не вышло, так как ко времени окончания эксперимента цена на диоды упала и не было никакого расчета производить их. Но параллельно с этим я занимался другими работами, и однажды мне надо было сделать электрический контакт между металлом и германием n-типа, отличавшийся от точечных контактов в диодах.

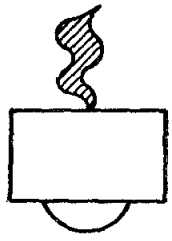
Разница заключалась в том, что этот контакт должен был быть в миллионы раз шире, а его сопротивление не должно было зависеть от направления тока (сопротивление точечного диода много больше для одного из двух направлений тока; это свойство позволяет диодам действовать как выпрямители). Однажды я обнаружил, что и в случае широких контактов сопротивление менялось при изменении направления тока и, что еще более странно, плотность тока была значительно меньше, чем в точечном диоде. С тех пор, как впервые были применены транзисторы, прошло уже 4 года, но действие точечного диода все еще не было полностью объяснено,

---

*) J. B. Gunn, The Gunn Effect, International Science and Technology, № 46, 43 (October 1965). Перевод Е. Д. Михальчи.

В оригинале — резюме: «В этой статье описана история открытия нового явления в физике твердого тела и многообещающие применения этого явления в электронике».

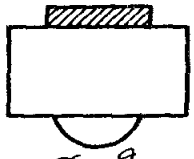
хотя диод легко было изготовить и на его основе разрабатывались первые виды транзисторов. Необходимо было значительно развить существующие теории полупроводников для того, чтобы объяснить квадратичную зависимость плотности тока, текущего через точечный диод в обратном направлении, от приложенного напряжения.



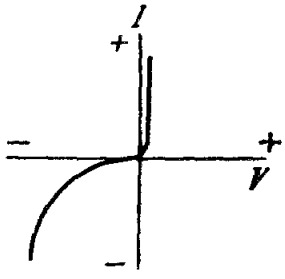
Диод с точечным контактом (точечный диод)

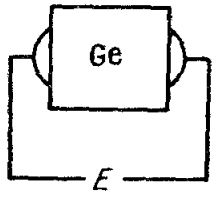
Неожиданно оказалось, что эти теории лучше объясняли вольтамперную характеристику диодов с широкими контактами. Это обстоятельство убедило меня в том, что на самом деле известные до сих пор теории нельзя применять к точечным диодам, в которых происходило нечто, резко отличающееся от того, что мы предполагали. Я решил попытаться выяснить причину этого.

Приблизительно в то же время Райдер и Шокли опубликовали статью, которая, как я думал, прольет



Диод с широким контактом





$v = \mu E$
Скорость дрейфа = подвижность × электрическое поле

свет на эту проблему. Они нашли, что при увеличении электрического поля, действующего поперек образца германия n-типа, плотность тока тоже растет, но ее нарастание постепенно ослабевает, а при напряженностях поля около 3000 *в/см* вообще прекращается. Из плотности тока можно определить скорость дрейфа электронов в приложенном поле. Зависимость дрейфовой скорости от напряженности поля приведена на рис. 1.

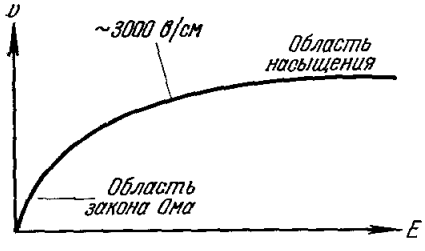


Рис. 1.

Мне казалось, что насыщение скорости дрейфа, которое ясно видно на рисунке, могло бы объяснить поведение плотности тока в точечном диоде. Кривая показывает, что для данного числа электронов существует некоторая предельная плотность тока, которую нельзя превысить. Эта предельная плотность тока, по-видимому, как раз и достигается в точечном диоде, где точечный контакт так мал (возможно, менее микрона в диаметре), что плотность тока около него могла стать весьма большой. При этих условиях около контактов создается очень сильное электрическое поле, которое в свою очередь увеличивает число свободных электронов. Конечный результат сводится к тому, что ток должен нарастать с увеличением напряжения; это явление и наблюдалось многими экспериментаторами.

ЭЛЕКТРОНЫ ПОЛУЧАЮТ ТЕПЛО

Данные Шокли и Райдера, относящиеся к дрейфовой скорости, были весьма расплывчаты, и поэтому, чтобы понять имеет ли моя теория какое-нибудь отношение к действительности, мне необходимы были новые экспериментальные данные. К этому времени я поступил в Королевское радиолокационное управление на трехгодичную стажировку. Меня это вполне устраивало, так как я мог заниматься интересующими меня вопросами. Я потратил много времени, повторяя измерения, и попутно заинтересовался проблемой горячих электронов.

Кривую, приведенную на рис. 1, очень грубо можно объяснить следующим образом. Когда электрическое поле E увеличивается, электрон приобретает от поля кинетическую энергию, пропорциональную vE. Для сохранения установившегося равновесия электрон должен передать часть энергии кристаллической решетке, поэтому его средняя энергия должна быть больше, чем энергия решетки, подобно тому, как любой нагретый предмет отдает энергию только в том случае, если его температура хотя бы немного выше температуры окружающей среды. Если мы считаем, что у электронов есть температура, а это в силу ряда причин вполне допустимо, можно описать этот процесс, сказав, что температура электронов превышает температуру решетки, и именно эта разность температур позволяет энергии рассеиваться. Назовем электроны с более высокой средней энергией «горячими», чтобы противопоставить их «холодным» электронам проводимости в обычном транзисторе. При нагревании электроны чаще сталкиваются с фононами решетки до тех пор, пока в области около 3000 *в/см* их дрейфовая скорость не насыщается,— отсюда кривая Райдера — Шокли. Последующие годы я изучал различные свойства горячих электронов не только для подтверждения моей теории диода, но и потому, что меня всегда интересовала техника высокоскоростных импульсных измерений сама по себе, а проведение измерений с горячими электронами — великолепный повод для работы с импульсным оборудованием. Характер зависимости дрейфовой скорости от E определяется концентрацией примесей, и я проделал массу утомительных измерений, чтобы найти и точно определить эту зависимость. Становилось все труднее и труднее выделить «температуру», и поэтому надо было переходить к другим способам ее измерения. Одним из способов определения температуры проводников электричества является измерение уровня мощности электрических шумов, возникающих в проводнике. При этом способе измеряются очень незначительные флуктуации тока, возникающие в результате того, что электроны не только перемещаются в заданном направлении с определенной средней скоростью, но и движутся по всем направлениям совершенно произвольным образом. Путем измерения флуктуаций тока мы определяем степень беспорядочности движения электронов, которая определяется температурой. Я работал около двух лет над экспериментами по температурным шумам вместе с Эрихом Эрлбахом, после чего решил заняться исследованием других веществ, кроме германия n-типа.

Германий — элемент 4-й группы и называется «неполярным» полупроводником, потому что все атомы кристаллической решетки электрически нейтральны. Если вы соедините элемент 3-й группы, например галлий, с элементом 5-й группы таким, как мышьяк, то полученное соединение ведет себя как вещество, состоящее из атомов 4-й группы. Однако так как электроны распределяются между двумя атомами различной валентности, каждый атом будет иметь небольшой электрический заряд. Смещение атомов из-за тепловых колебаний решетки приводит к появле-

нию электрических полей, которые в свою очередь влияют на условия прохождения электронов через решетку и, следовательно, на поведение горячих электронов. Именно по этой причине было интересно сравнить поведение полярного соединения из элементов 3—5 групп и германия.

ПОЯВЛЯЕТСЯ АРСЕНИД ГАЛЛИЯ

Чтобы проводить опыты с горячими электронами, необходимо иметь высококачественные кристаллы, так как эксперименты с кристаллами низкого качества либо вообще невозможно проделать, либо они бессмысленны. В то время арсенид галлия был единственным из соединений 3—5 групп, производство которого было достаточно освоено, чтобы его можно было использовать для сравнения с германием. В этот момент проведение опытов с арсенидом галлия не казалось мне очень важным. По грубым оценкам эффекты, обусловленные горячими электронами, были бы менее заметны в арсениде галлия, чем в германии. Поэтому я не ожидал увидеть что-либо весьма интересное — я думал, что получу сходную, но менее четко выраженную картину. Затруднения при наладке оборудования задержали эксперимент на несколько месяцев, но эта отсрочка меня мало волновала. Однако необходимо было сделать еще одну попытку, и в конце концов я решил, что надо или кончать эксперимент, или забыть о нем. Потратив около двух недель на необходимую постановку эксперимента, я вскоре обнаружил, что при повышении электрического поля до 1000 или 2000 *в/см* происходит нечто совершенно непредвиденное. Вместо обычного изменения тока с напряжением начиналось что-то невообразимое — ток начинал совершенно беспорядочно скакать вверх и вниз и это очень напоминало электрические шумы; однако амплитуда этих шумов была так велика, что она не могла быть объяснена ни одним из известных мне механизмов шума. Изменения тока были порядка ампер, а не наноампер, которые обычно наблюдаются. Сталкиваясь с чем-то неожиданным, каждый экспериментатор, конечно, прежде всего начинает подозревать, что у него не в порядке оборудование. Поэтому я тут же заменил арсенид галлия германием и обнаружил, что ток оставался постоянным. Это меня озадачило. Я не имел никакого представления о том, что́ же может вызывать эти флуктуации, но они обладали одной характерной чертой, которая была совершенно очевидна: если вы делали вид, что не замечаете быстрых флуктуаций, и снимали зависимость среднего тока, ток увеличивался почти линейно с увеличением напряжения до некоторой критической величины, после достижения которой он начинал падать. Прибор, ток в котором падает с ростом напряжения, во многом похож на отрицательное сопротивление и в принципе может быть использован для усиления и других интересных вещей. Классическим примером современного прибора с отрицательным сопротивлением является туннельный диод Есаки. Казалось, что и арсенид галлия обладает отрицательным сопротивлением.

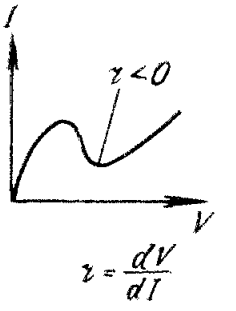
Другим обстоятельством, говорившим в пользу того, что эти флуктуации не были просто наваждением, было то, что они происходили за время порядка наносекунды или меньше. Это означало, что здесь существуют процессы, частота которых порядка 10^9 *гц*, а любой механизм, дающий ток 1 *а* при частоте 10^9 *гц*, можно использовать при условии, конечно, что эти флуктуации не случайны. Если же они случайны, то самое большее, что вы можете из этого извлечь, — это мешать работе радарных установок.

Нетрудно было провести опыты, которые убедили меня, что этот весьма заметный эффект отрицательного сопротивления не был обусловлен тем,

что происходило на контактах (как это имело, например, место в диоде Есаки). Изменение тока вызывалось каким-то явлением, происходившим если не во всем полупроводнике, то в достаточно обширной области его. Эти так называемые объемные эффекты представляют особый интерес с практической точки зрения, так как приборы, использующие объемный эффект, выпускать значительно легче. Когда эффекты проявляются в области соединения или контакта, вы почти всегда сталкиваетесь с трудностями при изготовлении прибора, дающего большую мощность на высоких частотах. При объемных явлениях положение облегчается: без особого ущерба, например, можно удвоить выходную мощность, увеличив площадь сечения кристалла в два раза. Этого вы не можете сделать в транзисторе.

ИСТОЧНИКИ ОБЪЕМНЫХ ЭФФЕКТОВ

Разговоры об объемных эффектах велись и раньше. 10 лет назад Г. Кремер, который работает ныне с Вэрионом, предположил — по причинам, которые мы не будем рассматривать здесь,— что при достаточном увеличении энергии электронов их подвижность может измениться таким образом, что электроны будут двигаться медленнее в более сильном электрическом поле. Тогда возникнет отрицательное сопротивление, т. е. ток, который переносится этими электронами, будет уменьшаться при увеличении электрического поля. К сожалению, идея Кремера не была воплощена в жизнь, так как не удалось достичь достаточно высоких энергий: эти энергии должны были быть значительно выше максимальной энергии горячих электронов. Однако это предложение Кремера сыграло важную роль, так как в нем впервые говорилось о возможности существования объемного отрицательного сопротивления. Шокли подсчитал, что если бы существовал эффект Кремера, нельзя было бы обнаружить область с отрицательным сопротивлением, измеряя зависимость тока от напряжения, как тогда думали. Это объяснялось тем, что как только напряжение



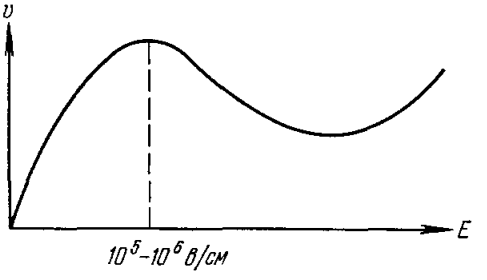
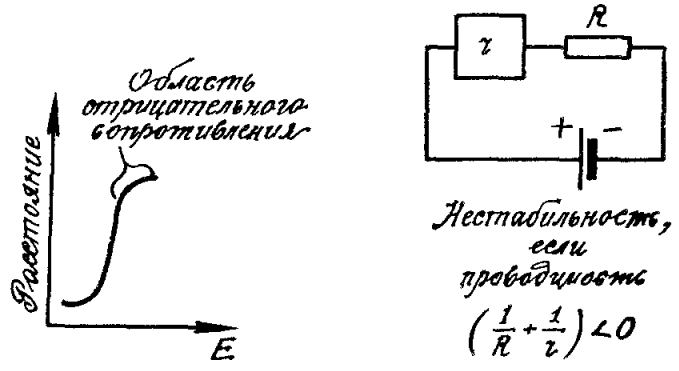


Рис. 2.

становилось достаточно высоким (его средняя величина превышала напряжение, соответствующее максимуму кривой на рис. 2), электрическое поле в кристалле становилось нестабильным и однородное распределение поля по образцу нарушалось. Эта область отрицательного сопротивления составляла очень незначительную часть пластины полупроводника, остальная часть которого действовала как обычное сопротивление.

Тогда обо всем этом читали, принимали к сведению и откладывали в сторону. Хотя сделанные мной предположения казались не очень убеди-

тельными, я допускал, во-первых, что наблюдаемое резкое уменьшение тока можно объяснить появлением области отрицательного сопротивления, которая в отличие от всех ранее известных случаев двигалась по образцу, и, во-вторых, что можно добиться стабильной картины, если образец сделать более коротким. Последнее предположение было основано на аналогии с электрической цепью, состоящей из двух последовательно включенных сопротивлений: одним (R) положительным, другим (r) отрицательным.



Такая цепь является нестабильной, если $\left(\frac{1}{R}+\frac{1}{r}\right)$ меньше нуля. Поэтому, сокращая длину образца, мы просто уменьшаем величину положительного сопротивления в цепи до тех пор, пока $1/R$ не станет больше абсолютной величины $1/r$.

ЧТО ПРОИСХОДИТ ПРИ ДЛИНЕ ОБРАЗЦА 0,2 *мм*

Именно поэтому я попросил инженера Дж. Стейнлса, который был моей опорой в течение всего эксперимента, изготовить несколько образцов, длина которых была бы меньше обычных 5 *мм*. Ничего не произошло, пока мы не достигли 0,2 *мм* (а сделать такой образец в то время было весьма трудно). В этом образце картина была иная: вместо случайных осцилляций ток начал изменяться более регулярным образом (рис. 3). Ток изменялся, как в обычной радиоволне, частота колебаний оставалась высокой — от 10^9 до $6\cdot10^9$ *гц*. Эти результаты впервые были опубликованы в июле 1963 г. Однако это было не то, что я ожидал. Я думал, что при уменьшении длины образца ток уменьшится, но будет оставаться стабильным при увеличении напряжения. Надежда на новый прибор с отрицательным сопротивлением не сбылась, однако был получен более интересный результат — генератор, дающий ток в несколько ампер на частоте 10^9 *гц* и более. Причина, по которой это было так интересно не только для нас, но и для всего полупроводникового производства, заключалась в том, что тогда невозможно было получать достаточно большую мощность от транзисторов на высоких частотах. Для этих целей применялись хрупкие электронные лампы, потребляющие большую мощность. Но прежде чем сказать, будет ли этот прибор действительно полезным, надо было ответить на множество вопросов. Наблюдение колебаний с частотой 10^9 *гц* еще не означает, что удастся построить прибор с полосой частот в 10^9 *гц*. Ценность прибора зависит от объема передаваемой информации, который определяется полосой частот.

С этой точки зрения прибор подавал надежды, так как изменения тока обычно начинались за время, меньшее периода колебаний приложенного напряжения. Тот факт, что ток нарастает почти мгновенно, давал основание предполагать, что полоса частот была того же порядка, что и при генерации. Это было весьма необычно и позволяло надеяться, что, каков бы ни был механизм этого странного явления, оно может быть использовано для радиоинтерференционных методов локации и для микроволнового телефона.

Однако в первую очередь надо было решить, может ли прибор работать в непрерывном режиме. До сих пор для всех экспериментов с горячими электронами применялась импульсная методика; электрическое

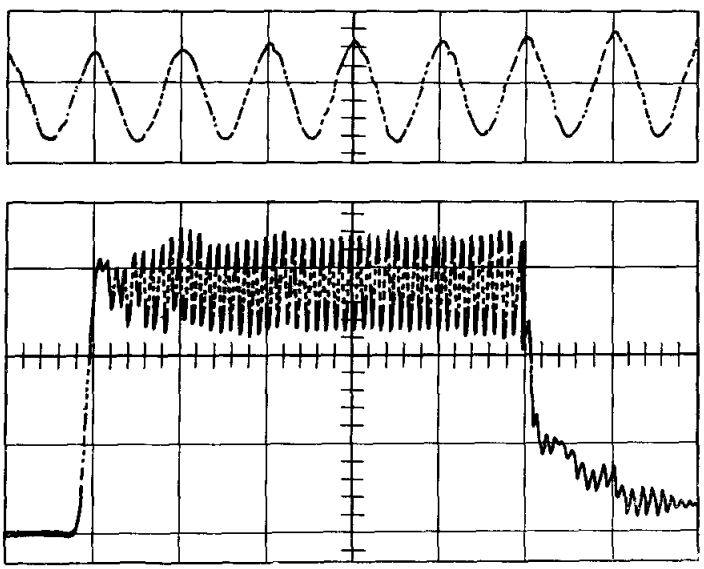


Рис. 3. Колебание тока в образце Ga — As n-типа длиной $2{,}5\cdot10^{-3}$ *см* под действием импульсного напряжения с амплитудой 16 *в* и длительностью 10^{-8} *сек*.

Частота колебаний тока $4{,}5\cdot10^{9}$ *гц*; на нижней осциллограмме масштаб по горизонтали $2\cdot10^{9}$ *сек/см*, по вертикали 0,23 *а/см*; верхняя кривая — увеличенное изображение нижней. (Текст подписи — из статьи Ганна в Solid State Communications 1(4), 88 (1963).— *Перев.*)

поле включалось только на время измерений (на микросекунду, а то и меньше). В противном случае решетка нагревалась так же, как электроны, и образец взрывался. Создание прибора, генерирующего непрерывные колебания, имело бóльшую ценность, чем разработка импульсного генератора. Если бы был найден способ уменьшения плотности электронов в арсениде галлия, можно было бы надеяться запустить генератор непрерывных колебаний. На словах уменьшить плотность электронов очень легко. Надо только попросить вашего друга, выращивающего кристаллы, сделать его более чистым. На самом деле это оказалось невозможным, так как в то время технология выращивания кристаллов не была совершенной. Тогда мои коллеги Н. Бреслау, Дж. Стейнлс и я попробовали нечто другое — мы сделали пластину меньше, чтобы быстрее отводить тепло от контакта. Из имевшегося в нашем распоряжении арсенида галлия пластину пришлось сделать такой тонкой, что работать с ней было почти невозможно. Однажды мы получили пластину, которая по каким-то причинам имела меньшую плотность электронов, чем мы ожидали. Я думаю,

что это произошло из-за загрязнения, и атомы примеси поглотили часть свободных электронов. Это обстоятельство позволило создать генератор, дающий непрерывные колебания. Этот генератор функционировал целую неделю, что послужило доказательством стабильности его работы. Об этом было сообщено в ноябре 1964 г. Почти одновременно работа о генерации непрерывных колебаний была представлена учеными из лаборатории Бэлла и из Линкольнской лаборатории МТИ. С тех пор технология очистки кристаллов усовершенствовалась, и я считаю, что проблема непрерывных колебаний окончательно решена.

С самого начала, думая о механизмах процесса, я пришел к двум выводам. Во-первых, процесс зависит от внутренних свойств материала — на протяжении всей этой работы я испробовал несколько других материалов и обнаружил эффект только в фосфиде индия, очень похожем на арсенид галлия,— полупроводнике, образованном элементами 3—5 групп. Во-вторых, у меня было неясное чувство, что должно распространяться электрическое возмущение от одного конца кристалла к другому. Действительно, когда мы сделали образец короче 0,2 *мм*, то нашли, что частота увеличивается обратно пропорционально длине. А постоянная пропорциональности оказалась почти точно такой же, как и дрейфовая скорость электронов (10^7 *см/сек*). Другими словами, период колебаний тока был равен времени, которое требуется электрону, чтобы пройти от одного конца кристалла до другого под действием приложенного электрического поля. Поэтому возникло неясное предположение, что должно быть нечто, движущееся через кристалл с той же скоростью, что и электроны. Это нечто нельзя было пощупать пальцами, но позднее я покажу, что мои предположения в конце концов подтвердились. Прежде всего необходимо было ответить на вопрос, почему уменьшается ток. Было много гипотез, и в первую очередь я проверил ту из них, которая вызвала в свое время большую сенсацию в научном мире: это был механизм усиления звуковых волн в полупроводниках, которое наблюдали Хатсон, Мак-Фи и Уайт из лаборатории Бэлла. Обычно звуковые волны затухают в твердом теле, но когда скорость электронов, движущихся под действием электрического поля через пьезоэлектрический полупроводник, превышает скорость звуковой волны в том же направлении, некоторая часть импульса, полученного электроном от электрического поля, может быть отдана фононам звуковой волны. Так как скорость звуковой волны определяется механическими свойствами решетки, скорость звука близка к постоянной. Таким образом, за счет такой передачи импульса должна увеличиваться амплитуда звуковых волн, а электроны должны замедляться, в результате чего ток падает.

Было рассчитано максимальное усиление звуковых волн электронами с дрейфовой скоростью 10^7 *см/сек*, которое вызывало уменьшение тока. По сравнению с опытными данными это усиление оказалось в 10^4 раз меньше; поэтому для нашего случая механизм передачи энергии электронов фононам звуковой волны был неприменим.

Когда я обсуждал этот вопрос с П. Прайсом, одним из моих коллег теоретиков, он заметил: «Если это не акустические колебания, то, может быть, оптические?». Оптические переходы являются вибрационными колебаниями, а при определенных условиях их частота соответствует частоте видимого спектра.

Хотя оптические волны подобны до некоторой степени звуковым, взаимодействие с электронами усиливает оптические колебания в значительно большей степени, чем акустические. Я рассчитал величину усиления и нашел, что для согласия с наблюдаемым значением изменения тока необходима скорость электронов в три раза бóльшая, чем измеренная

нами. Расхождение в три раза для таких грубых оценок не так уж велико, поэтому казалось вполне возможным, что это был механизм, объясняющий уменьшение тока. Третий возможный теоретический подход для объяснения отрицательной проводимости предложили Ридли и Уоткинс из Исследовательской лаборатории Мулларда и Хилсам из Службы электронных исследований в Англии. Суть их работы заключается в том, что при воздействии сильного электрического поля на полупроводники с зонной структурой обычного типа электроны могут переходить из области, качественная характеристика которой, называемая эффективной массой, мала, в область, где эффективная масса много больше. Другими словами, электроны могут перейти из состояния, где их подвижность была велика, в состояние, где их подвижность много меньше. Этим может объясняться эффект отрицательного сопротивления. Чтобы лучше понять эти важные выводы, надо рассмотреть некоторые основные понятия волновой механики.

МИР ЭЛЕКТРОННЫХ ВОЛН

Прежде всего, вспомним, что происходит с классической частицей в свободном пространстве. Кинетическая часть полной энергии частицы ε есть $\frac{1}{2}mv^2$, или $\frac{1}{2}\frac{p^2}{m}$, где p — импульс частицы. Таким образом, энергия частицы является параболической функцией импульса, как показано на рис. 4. В квантовомеханическом мире электронов и атомов электрон не является такой же частицей, как биллиардный шар, служивший в давние времена моделью атомов у химиков. Волновая механика учит нас, что электрон является волновым пакетом с длиной волны, равной h/p, где h — постоянная Планка. При анализе поведения электронов в твердых телах удобно ввести так называемый волновой вектор $\mathbf{k}$, равный $2\pi p/h$. Если электрон свободный, то ε будет пропорционально k^2; это справедливо и для свободной классической частицы, и мы можем нарисовать красивую параболу для зависимости ε от k. Но все дело в том, что движение электрона в твердом теле не будет свободным — он взаимодействует с громадным числом атомов решетки, и вследствие этого взаимодействия потенциальная энергия атомов будет периодически меняться в пространстве, что оказывает сложное воздействие на поведение электрона и искажает простую параболу почти до неузнаваемости. По оси энергии существуют области, называемые разрешенными энергетическими зонами, разделенные друг от друга щелями, где действительных значений $\mathbf{k}$ вообще нет. В полупроводниках почти все энергетические зоны или целиком заполнены, или абсолютно пусты. Интересные электрические свойства возникают из-за присутствия нескольких свободных электронов в пустой зоне («зоне проводимости») или из-за их отсутствия в почти полной зоне («валентной зоне»). Работа большинства устройств, основанных на контакте полупроводников, таких, как транзисторы или инжекционные лазеры,

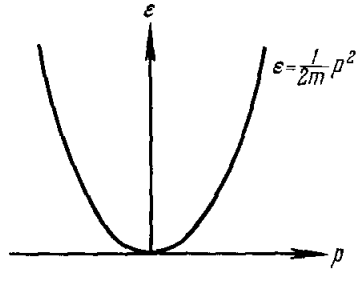
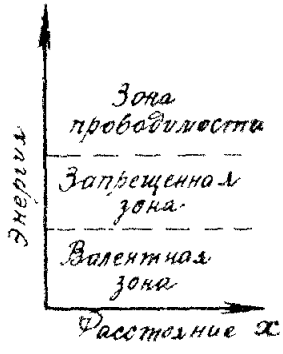


Рис. 4.



зависит от того, что в одной зоне есть избыток электронов, а в другой их не хватает. Но нас интересует не это — нам нужно рассмотреть влияние примесных атомов, находящихся в арсениде галлия, на электроны, впрыснутые в зону проводимости. Для этой цели удобно мысленно представить энергетические зоны как изменение ε с k и изобразить это диаграммой в так называемом k-пространстве (рис. 5). Возможно более употребительно изображение ε как функции расстояния (конфигурационное пространство). Фактически ε зависит от k и x, но для одновременного изображения зависимости от обеих переменных требуется пространство более высокого порядка, в котором нарисовать диаграммы невозможно. Для нас важно изменение ε в зависимости от волнового вектора, а не от расстояния. Поэтому ниже будет использоваться диаграмма ε (k). Несмотря на довольно сложный вид этой диаграммы, которая основана на волновой трактовке электронов, оказывается возможным сохранить некоторые из наших интуитивных представлений об электронах как частицах, если мы допустим, что электроны обладают некоторыми довольно странными свойствами. В классической механике ускорение dv/dt получается из формулы

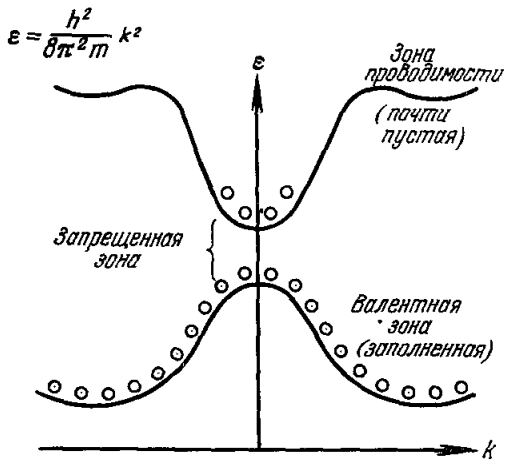


Рис. 5.

$$\frac{dv}{dt} = \frac{F}{m},$$

тогда как, согласно волновой механике, соответствующее уравнение для электронов в твердом теле есть

$$\frac{dv}{dt} = \frac{4\pi^2}{h^2}\left(\frac{d^2\varepsilon}{dk^2}\right)F$$

Величина

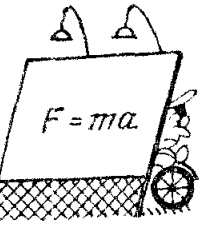
$$\frac{4\pi^2}{h^2}\left(\frac{d^2\varepsilon}{dk^2}\right)$$

играет в уравнениях волновой механики такую же роль, что и $1/m$ в классической. Таким образом, для того чтобы все-таки не терять связи с ньютоновской механикой, мы допустим, что электрон ведет себя так, словно он обладает эффективной массой

$$m^* = \frac{h^2}{4\pi^2}\left(\frac{d^2\varepsilon}{dk^2}\right)^{-1}.$$

Эта масса совершенно отлична от массы свободного электрона. Она является функцией $\mathbf{k}$, в отличие от классической массы, величина которой постоянна и не зависит от импульса. Это означает, что, настаивая на рассмотрении электрона как частицы, мы должны допускать изменение его массы, а это в высшей степени усложняет уравнения движения. Но есть одна ситуация, где эффективная масса может рассматриваться как постоянная, по крайней мере в точке (см. точку k_0 на рис. 6), где энергия имеет минимум. Вблизи k_0 энергия равна константе плюс $\frac{h^2}{8\pi^2 m^*}(\mathbf{k} - \mathbf{k}_0)^2$, где m^* не зависит от k. Если имеется несколько таких минимумов, то

значения m^*, соответствующие каждому из них, могут быть различны. Для каждого вещества форма зоны проводимости неповторима (она определяется видом зависимости ε от k), ее искажают многочисленные минимумы и максимумы. Приведенное соотношение с заданным m^* справедливо только для некоторого определенного минимума. В теории Ридли — Уоткинса рассматривается диаграмма, подобная изображенной на рис. 6 и состоящая из двух минимумов. Нижний минимум обычно заполнен электронами, когда вещество находится в температурном равновесии. Энергетический интервал между следующим минимумом и основным несколько больше энергии, которую получают электроны от электрического поля. Радиус кривизны минимумов приблизительно пропорционален $1/m^*$, а из теоретических положений следует, что если более высокий минимум имеет меньшую кривизну, то электроны, попавшие в эту зону, будут замедляться, так как их эффективная масса будет больше. Такой механизм уменьшения подвижности электронов может иметь место в случае полупроводников из элементов 3—5 групп.



$$\frac{dv}{dt} = \frac{1}{m} F$$

$$\frac{dv}{dt} = \frac{4\pi^2}{h^2} \frac{d^2\varepsilon}{dk^2} F$$

$$\frac{dv}{dt} = \frac{1}{m^*} F$$

$$\frac{1}{m^*} = \frac{4\pi^2}{h^2} \frac{d^2\varepsilon}{dk^2}$$

На эту тему имелись только теоретические работы, и, насколько мне известно, никто никогда не наблюдал подобного явления. Мне, естественно, хотелось проследить, не могут ли эти теории иметь отношение к нашим результатам. С этой целью мой коллега Си Брайон проделал некоторые измерения температурных шумов арсенида галлия и нашел, что средняя энергия электронов была много меньше 0,36 *эв*— энергетического интервала между минимумами в арсениде галлия, что поколебало мою веру в механизм Ридли — Уоткинса. Величине 0,36 *эв* соответствует температура 4000° К, а на опыте получалась температура около 400° К. На той стадии исследований можно было попытаться проверить и ряд других возможных механизмов этого явления, но мне показалось, что механизм, обусловленный оптическими эффектами, будет лучше других. Однако многие сомневались в этом, и однажды Кремер указал, что я пропустил в своих расчетах множитель, учет которого приводил к тому, что вычисленные значения температур шумов оказывались равными всего лишь 600° К, а не 4000° К, а это практически означало согласие с теорией Ридли — Уоткинса. Точность теории была не такой, чтобы кто-нибудь беспокоился об ошибке в 50%, и поэтому казалось, что часть полученных результатов не находится в серьезном противоречии с теорией Ридли — Уоткинса о переходе электронов из одного минимума в другой. Не так давно группа из лаборатории Бэлла провела эксперимент с арсе-

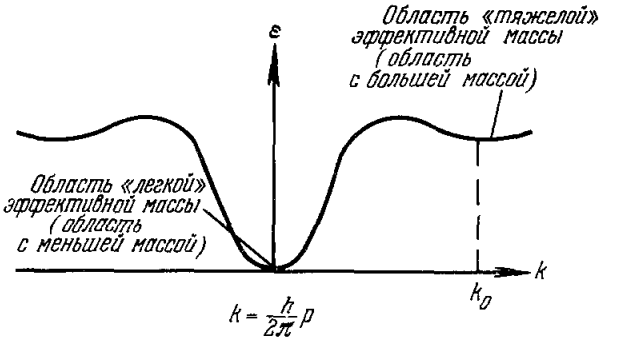


Рис. 6.

хотя бы небольшое количество электронов, проводимость понижается из-за уменьшения подвижности электронов в более высоком минимуме. Это вызывает дальнейшее возрастание электрического поля в этой области, что в свою очередь вызывает перемещение еще большего количества электронов. Таким образом, скорость переноса электронов непрерывно возрастает, и происходит лавинообразное нарастание поля. Мы сделали киносъемку распределения потенциала в арсениде галлия и нашли, что такое описание соответствует действительности. Как только электроны начинают переходить из одного минимума в другой, все происходит очень быстро. Если появилась ударная волна, то поле и ток пропорционально уменьшаются, а волна начинает перемещаться. Можно предполагать, что ядро этой волны состоит из очень энергичных электронов, переходящих в состояние с более высокой энергией. Когда волна проходит через кристалл, она оставляет за собой след в виде возбужденных электронов. Можно показать, что эти замедляющиеся электроны отстают от ударной волны, теряют энергию, пока не вернутся в свое первоначальное состояние, после чего их скорость возрастает настолько, что они вновь догонят ударную волну. Вы можете мысленно представить, что электрон совершает некий цикл (рис. 9), то догоняя ударную волну, то отставая от нее. На основе этой модели можно понять, почему длина кристалла является критической характеристикой для когерентных колебаний. В большом образце изменения поля у катода не так велики, чтобы предотвратить начало второй ударной волны прежде, чем первая достигнет анода, поэтому возникновение волн носит случайный характер. Но, как я уже говорил, я пытаюсь математически обосновать мою идею. Кроме того, есть надежда, что дальнейшее математическое развитие теории позволит объяснить некоторые новые явления, которые наблюдались в германии, легированном золотом, и в полудиэлектрическом арсениде галлия, а также странные фононные эффекты. Все эти явления объясняются тем, что подвижность электронов существенно зависит от их энергии.

Рис. 9.

Прежде чем мы сможем использовать арсенид галлия для микроволнового генератора, нам нужно устранить кое-какие неясности. Например,

Легкое колесо (около 60 *г*) состоит из центрального диска диаметром 5 *см*, обода с внешним и внутренним диаметрами 22 и 20 *см* и 4 резиновых спиц шириной 3—4 *см*. Обод и центральный диск двухслойные, вырезанные из листового дюралюминия толщиной 0,3 *мм*. Спицы из резины для хирургических перчаток закреплены между слоями с помощью клея и легких скрепок, прижимающих слои друг к другу (рис. 1). Колесо может вращаться около горизонтальной оси с очень малым трением. Центр тяжести колеса должен находиться на оси или очень близко к ней. Для облегчения центрировки на обод насажено несколько легких муфт, которые могут перемещаться.

На левую (или правую) половину колеса направляется интенсивный пучок инфракрасного излучения (рис. 2). Под действием излучения спицы этой половины колеса нагреваются и сокращаются, центр тяжести колеса смещается относительно оси, возникающий вращающий момент приводит колесо в движение. При повороте колеса

кристалла. Поэтому нам пришлось сконструировать собственный датчик, в котором заряды наводятся на маленькую пластинку, подвешенную на расстоянии 3 μ от поверхности кристалла и имеющую в направлении движения тока длину 15 μ. На осциллографе с помощью этого прибора можно получить сигнал, пропорциональный производной по времени от потенциала (рис. 7). Отчетливо видно, как происходит распространение возмущения вдоль кристалла. При этом остаются неизменными амплитуда возмущения и его скорость, равная магической величине 10^7 *см/сек*; эту магическую скорость можно получить, умножив частоту колебания на длину кристалла. Из данных, которые были получены

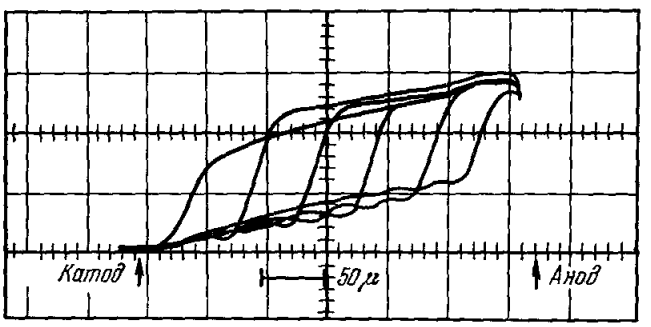


Рис. 8.

о скорости изменения потенциала при помощи этой методики, мы в состоянии сделать вывод о самом характере движения потенциала. В первый момент после того, как приложенное напряжение превысило пороговую величину, при которой ток становится нестабильным, электрическое поле распределено равномерно по всей длине образца, однако вскоре появляется нестабильность — область сильного поля вблизи катода. Разность потенциалов между анодом и катодом постоянна, поэтому любое увеличение поля в одном месте образца должно сопровождаться уменьшением поля в другом месте. В нашем эксперименте мы видели, что с течением времени эта область более сильного поля движется от катода к аноду. Когда эта область достигнет анода, она исчезает, и электрическое поле моментально возвращается к своему первоначальному однородному распределению. Как только это произойдет, вновь создаются условия, при которых образуется область сильного поля у катода, и новая область высокого потенциала движется через кристалл таким же образом. Этот процесс циклически повторяется. Изменения потенциала носят колебательный характер, так как вне района сильного поля арсенид галлия ведет себя как обычный проводник, т. е. с понижением напряжения понижается и ток. Недавно мы смогли получить на осциллографе изменение самого потенциала, напоминающее тот цикл, который мы обсуждали (рис. 8). Следует сказать, что передвижение потенциала очень напоминает аэродинамическую ударную волну, движение фронта которой не зависит от вызвавшего его процесса. Это не означает, что есть какое-то механическое или акустическое возмущение, которое возникает вместе с этой волной, но надо отметить, что поведение потенциала подчиняется нелинейным уравнениям, аналогичным тем, которые описывают распространение ударных волн. В настоящее время я пытаюсь разработать математическую теорию для этого случая. Уже сделано почти $^2/_3$ работы, и то, что пока получается, приводит к довольно простой картине. Дело заключается в следующем. Когда электрическое поле становится достаточно сильным, электроны начинают перемещаться из низшего минимума в более высокий. Как только переместится

хотя бы небольшое количество электронов, проводимость понижается из-за уменьшения подвижности электронов в более высоком минимуме. Это вызывает дальнейшее возрастание электрического поля в этой области, что в свою очередь вызывает перемещение еще большего количества электронов. Таким образом, скорость переноса электронов непрерывно возрастает, и происходит лавинообразное нарастание поля. Мы сделали киносъемку распределения потенциала в арсениде галлия и нашли, что такое описание соответствует действительности. Как только электроны начинают переходить из одного минимума в другой, все происходит очень быстро. Если появилась ударная волна, то поле и ток пропорционально уменьшаются, а волна начинает перемещаться. Можно предполагать, что ядро этой волны состоит из очень энергичных электронов, переходящих в состояние с более высокой энергией. Когда волна проходит через кристалл, она оставляет за собой след в виде возбужденных электронов. Можно показать, что эти замедляющиеся электроны отстают от ударной волны, теряют энергию, пока не вернутся в свое первоначальное состояние, после чего их скорость возрастает настолько, что они вновь догонят ударную волну. Вы можете мысленно представить, что электрон совершает некий цикл (рис. 9), то догоняя ударную волну, то отставая от нее. На основе этой модели можно понять, почему длина кристалла является критической характеристикой для когерентных колебаний. В большом образце изменения поля у катода не так велики, чтобы предотвратить начало второй ударной волны прежде, чем первая достигнет анода, поэтому возникновение волн носит случайный характер. Но, как я уже говорил, я пытаюсь математически обосновать мою идею. Кроме того, есть надежда, что дальнейшее математическое развитие теории позволит объяснить некоторые новые явления, которые наблюдались в германии, легированном золотом, и в полудиэлектрическом арсениде галлия, а также странные фононные эффекты. Все эти явления объясняются тем, что подвижность электронов существенно зависит от их энергии.

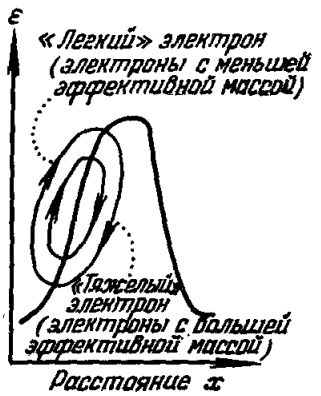


Рис. 9.

Прежде чем мы сможем использовать арсенид галлия для микроволнового генератора, нам нужно устранить кое-какие неясности. Например, приходится иметь дело с огромной мощностью в единице объема, порядка 10^7 *вт/см*^{3}, для чего необходимы более совершенные устройства для отвода тепла. Есть еще некоторые неразрешенные вопросы, касающиеся взаимосвязи между прибором и внешней цепью. Неясно, почему мощность именно такая, а не в два раза меньше или в два раза больше. Вероятно, эти проблемы будут разрешены, так как многие лаборатории занимаются ими. Мои коллеги из ИБМ и я бóльшую часть своего времени посвятим, возможно, выяснению основного механизма — механизма ударной волны.

ЛИТЕРАТУРА

1. B. K. Ridley, T. B. Watkins, Proc. Phys. Soc. 78, 293 (1961).
2. B. K. Ridley, Proc. Phys. Soc. 82, 954 (1963).
3. J. B. Gunn, IBM J. Res. and Dev. 8, 141 (1964).
4. A. B. Hutson et al., Phys. Rev. Letts, 14, 638 (1965).
5. C. Hilsum, Proc. IRE 50, 185 (1962).
6. Ч. Киттель, Введение в физику твердого тела, М., Физматгиз, 1963.