

И. А. Случинская

Основы материаловедения и технологии полупроводников

Москва — 2002

Введение

Прогресс современной науки и техники неразрывно связан с успехами как в развитии и совершенствовании технологии получения традиционных полупроводниковых материалов, так и в разработке и исследовании новых перспективных полупроводников.

Поразительные возможности современной полупроводниковой электроники и особенно микроэлектроники реализуются только по мере разработки и освоения выпуска полупроводниковых материалов с разнообразными физическими свойствами. Эти материалы позволили создать на их основе миниатюрные усилители и генераторы электрических сигналов, работающие в широком диапазоне частот; интегральные микросхемы для современных компьютеров; преобразователи одного вида энергии в другой; полупроводниковые светодиоды, лазеры и фотоприемники, работающие в ИК- и видимом диапазонах (полупроводниковые лазеры и фотоприемники — составляющие элементной базы волоконно-оптических линий связи); детекторы излучений и частиц; магнитные, пьезо-, сегнетоэлектрические и многие другие устройства. В то же время открытие новых явлений и потребность создания более совершенных приборов для научных исследований стимулируют поиск, разработку и освоение производства новых материалов с требуемыми свойствами. Между физикой и технологией полупроводников существует тесная взаимосвязь, и часто оказывается, что получение новых физических результатов становится невозможным без постоянного прогресса в технологии.

Работы по эпитаксии прекрасно иллюстрируют эту взаимосвязь. Идея создания тонких эпитаксиальных пленок полупроводников была впервые высказана в 1951 г., а первые эпитаксиальные транзисторы были изготовлены в 1960 г. Быстрый прогресс в области фундаментальных исследований механизма роста тонких пленок, связанный с широким промышленным применением эпитаксиальной технологии, привел к разработке новых методов эпитаксии — молекулярно-лучевой эпитаксии и MOCVD

(химическое осаждение из газовой фазы с использованием металлоорганических соединений). Эти достижения в технологии позволили перейти к получению очень тонких пленок, толщина которых измеряется единицами атомных слоев, и исследованию их физических свойств, а также к созданию и изучению принципиально новых объектов — сверхрешеток.¹

Идея создания искусственных периодических структур, обладающих свойствами, качественно отличающимися от свойств соответствующих объемных материалов, была высказана еще в шестидесятые годы. В начале 70-х годов появились первые сообщения о получении таких структур, после чего начались систематические исследования свойств сверхрешеток, ведущиеся в тесном контакте физиками и технологами. В ходе этой работы были обнаружены новые интересные физические явления, например, квантовый эффект Холла [1, 2] и некоторые другие квантоворазмерные эффекты [3]; были выявлены возможности приборного применения сверхрешеток [4, 5]. Все эти результаты могли быть получены только при детальном исследовании механизмов эпитаксиального роста и процессов, происходящих на гетерограницах в слоистых структурах, что и позволило создать сложные, достаточно совершенные сверхрешетки, состоящие из эпитаксиальных слоев толщиной в 1–5 моноатомных слоя. Этому способствовал выросший уровень исследовательской техники (современных методов исследования структуры и состава поверхности полупроводников) и вакуумной технологии. В настоящее время интерес исследователей в области эпитаксиального роста обращен на обнаруженные недавно особенности островкового роста на вицинальных гранях, напоминающие явление самоорганизации в живых системах.

Раньше, а нередко и сейчас, новые полупроводниковые материалы создавались эмпирически. Однако знание природы химической связи, структуры, физико-химических свойств известных материалов и установление взаимосвязи между свойствами и способами их производства позволяют научно обоснованно прогнозировать наличие полупроводниковых свойств у новых материалов и разрабатывать технологию получения этих новых полупроводников с заданными параметрами.

Материаловедение полупроводников — наука, изучающая закономерности образования полупроводниковых фаз (элементарных веществ,

¹Под сверхрешеткой понимают периодическую структуру, состоящую из тонких (нанометровых) слоев двух или более полупроводников, повторяющуюся в одном направлении. В этой структуре на периодичность самой кристаллической решетки накладывается дополнительная периодичность, обусловленная созданным закономерным изменением электронных или структурных свойств кристалла. Дополнительный период по толщине, обычно изменяющийся от единиц до десятков нанометров, меньше длины волны электронов, но больше постоянной кристаллической решетки.

соединений, твердых растворов) в равновесных и неравновесных условиях, влияние химического состава, атомной структуры и структурных дефектов на свойства материала, а также разрабатывающая принципиальные подходы (способы) управления свойствами получаемого материала.

Технология полупроводников — это совокупность методов, которые позволяют получить полупроводник с заданными параметрами. На основе результатов, полученных в материаловедении, технологи разрабатывают и совершенствуют промышленные способы получения материалов.

Описать подробно все существующие технологические приемы, используемые в производстве материалов для электронной техники, в одном курсе вряд ли возможно и целесообразно, по крайней мере, из-за довольно быстрого совершенствования технологии. Сведения о способах получения того или иного конкретного полупроводника можно получить из статей, справочников, монографий и т. д. Однако существуют некоторые общие закономерности технологических процессов производства разных полупроводниковых материалов. Это позволяет объединить способы получения различных полупроводников в типичные технологические методы, выделить наиболее широко используемые и детально их проанализировать.

В настоящем курсе излагаются основы материаловедения и технологии полупроводников, которые позволяют получить общее представление об основных закономерностях образования полупроводниковых фаз, о механизмах их роста, условиях получения, а также получить представление о наиболее широко используемых в промышленности методах производства объемных полупроводников и эпитаксиальных полупроводниковых пленок с заданными свойствами. Эти знания необходимы студентам, специализирующимся в области физики полупроводников, для понимания специальной литературы. Действительно, почти каждая статья, посвященная исследованию свойств полупроводников или созданию приборов на их основе, начинается с описания метода получения материала, так как его свойства, как будет показано в этом курсе, тесно связаны с методом его получения.

Весь материал курса разбит на главы, тесно связанные друг с другом. Во второй главе на основе теории химических связей излагаются основные закономерности образования полупроводников. В третьей главе рассмотрены структурные дефекты и их влияние на свойства материала. Описание существующих фазовых диаграмм, которые позволяют выбрать оптимальные условия получения полупроводника заданного состава, а также элементы общей теории образования фаз даны в четвертой

главе. Методы получения материалов полупроводниковой степени чистоты рассмотрены в пятой главе. Глава 6 посвящена описанию основных технологических методов получения полупроводниковых монокристаллов из жидкой и газовой фаз. Общие принципы и методы легирования полупроводников изложены в седьмой главе. В восьмой главе рассмотрены основные закономерности диффузии, широко используемой для создания заданного профиля распределения примесей в полупроводниковых структурах. Наконец, в главе 9 описаны основные закономерности роста эпитаксиальных пленок и основные используемые в настоящее время методы эпитаксии — жидкостная и газовая эпитаксия, MOCVD, молекулярно-лучевая эпитаксия и др.

Глава 1

Некоторые свойства полупроводников и основные группы полупроводниковых материалов

1.1. Отличительные свойства полупроводников

Среди твердых тел, как хорошо известно, можно выделить определенные группы веществ, отличающихся природой сил, действующих между атомами в твердом теле, и особенностями их зонной структуры, — металлы, полупроводники, диэлектрики.

В металлах преимущественным типом химической связи между атомами является металлический тип связи, а в полупроводниках и диэлектриках — ковалентный или ковалентно-ионный и ионно-ковалентный типы связи соответственно.

С точки зрения зонной структуры¹ наиболее важной особенностью металлов является то, что в них верхняя энергетическая зона, содержащая электроны (зона проводимости), имеет незанятые уровни. В диэлектриках и полупроводниках при низких температурах (близких к 0 К) верхняя, целиком заполненная электронами энергетическая зона (валентная зона) отделена от следующей — пустой — зоны разрешенных энергий (зона проводимости) запрещенным участком конечной величины E_g

¹Зонная структура образуется всеми возможными электронными состояниями (уровнями энергии) всех атомов, входящих в состав твердого тела. Эти состояния могут быть заполнены электронами или вакантны.

(запрещенной зоной), в котором энергетические уровни электронов отсутствуют. Примеси и дефекты в диэлектриках и полупроводниках могут приводить к появлению уровней в запрещенной зоне, но этих состояний сравнительно мало, так что понятие запрещенной зоны сохраняет смысл.

Полупроводники обычно выделяют в самостоятельную группу веществ. Но если между металлами и диэлектриками имеется качественное различие в зонной структуре, то между диэлектриками и полупроводниками различие только количественное — в величине E_g . Практически многие полупроводники при низких температурах (вблизи 0 К) ведут себя как диэлектрики, а многие диэлектрики при высокой температуре являются полупроводниками. Условно принято считать полупроводниками вещества, у которых при комнатной температуре $E_g < 2\text{--}3$ эВ. Однако следует заметить, что существуют и бесщелевые полупроводники, у которых $E_g = 0$ (например, HgTe и HgSe).

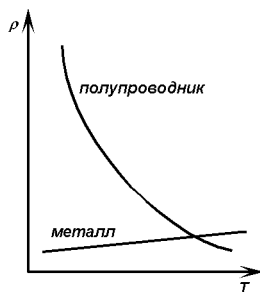
Обычно металлы кристаллизуются в компактные структуры с плотнейшей упаковкой атомов, они обладают одновременно высокой механической прочностью и высокой пластичностью. Полупроводники имеют, как правило, не плотноупакованную структуру, они хрупки вплоть до весьма высоких температур.

Металлы часто легко образуют сплавы с материалами с иной природой химической связи, полупроводники — плохо.

Металлы непрозрачны для электромагнитных волн от самых низких частот вплоть до середины ультрафиолетовой области спектра; для больших частот металлы становятся прозрачными. Кроме того, они в этой широкой области спектра (от низких частот до середины ультрафиолетовой области) хорошо отражают излучение. Диэлектрики и полупроводники в противоположность металлам прозрачны для электромагнитных волн от низких частот до некоторой граничной частоты, характерной для каждого материала и называемой *основной частотой поглощения или краем собственного поглощения*. Как правило, диэлектрики и полупроводники прозрачны в видимой области спектра. Часто у диэлектриков и полупроводников перед краем собственного поглощения наблюдаются пики примесного и экситонного поглощения.

Рассматриваемые группы материалов различаются величиной удельного сопротивления ρ (или электропроводности σ). Различие в величине ρ и дало повод к введению термина полупроводники (в смысле плохой проводник). Разделение материалов по значениям удельного сопротивления (электропроводности) условно. Принято относить к диэлектрикам материалы с $\rho = 10^{10} \div 10^{20}$ Ом · см ($\sigma = 10^{-20} \div 10^{-10}$ Ом⁻¹см⁻¹); к полупроводникам материалы с $\rho = 10^{-4} \div 10^{10}$ Ом · см ($\sigma = 10^{-10} \div$

Рис. 1.1. Характерные температурные зависимости удельного сопротивления в металлах и полупроводниках.



$10^4 \text{ Ом}^{-1}\text{см}^{-1}$); а к металлам материалы с $\rho = 10^{-6} \div 10^{-4} \text{ Ом} \cdot \text{см}$ ($\sigma = 10^4 \div 10^6 \text{ Ом}^{-1}\text{см}^{-1}$).

В металлах основными носителями заряда, как правило, являются электроны проводимости с энергией, близкой к энергии Ферми E_F ; в полупроводниках — электроны в зоне проводимости и дырки в валентной зоне; в диэлектриках — ионы.

В металлах легко возникает электрический ток за счет перехода электронов под влиянием приложенного извне слабого электрического поля на свободные верхние уровни в одной энергетической зоне. Большинство свойств металлов обусловлено именно этой низкой энергией возбуждения электронов. В полупроводниках слабое электрическое поле тока вызвать не может (при низких температурах). Здесь для возникновения тока необходимо перебросить часть электронов в незаполненную зону, преодолев при этом энергетический зазор E_g .

Наличие энергетического зазора приводит к различной зависимости удельного сопротивления от температуры (рис. 1.1). У металлов ρ , как правило, растет с повышением температуры T в первом приближении по степенному закону

$$\rho_m = \rho_0(1 + \alpha T^p), \quad (1.1)$$

где ρ_0 — удельное сопротивление при 0 К, α — температурный коэффициент сопротивления, T — абсолютная температура, $p \sim 1$.

У полупроводников, наоборот, ρ_s падает с повышением температуры по экспоненциальному закону

$$\rho_s = A \exp(E_a/kT), \quad (1.2)$$

где E_a — энергия активации, k — постоянная Больцмана.

Разная температурная зависимость удельного сопротивления в металлах и полупроводниках связана с различной температурной зависи-

мостью концентрации носителей заряда в них. В случае одного типа носителей

$$1/\rho = \sigma = ne\mu, \quad (1.3)$$

где n — концентрация электронов, e — заряд электрона, μ — подвижность электронов. В металлах n практически не зависит от температуры, и температурная зависимость σ_m связана с температурной зависимостью подвижности электронов, которая определяется рассеянием электронов на любых нарушениях периодичности кристаллической решетки (рассеяние на фононах, рассеяние на дефектах). Кроме того, носители могут рассеиваться друг на друге. С повышением температуры подвижность носителей в металлах всегда падает за счет их более интенсивного рассеяния тепловыми колебаниями атомов (рассеяние на фононах). В полупроводниках температурная зависимость подвижности электронов определяется механизмом рассеяния носителей заряда [6]. Наиболее важные механизмы, определяющие подвижность носителей в полупроводниках, — рассеяние на фононах (см. гл. 2) и заряженных примесях [6]. Температурная зависимость, характерная для этих механизмов рассеяния, носит степенной характер, и обычно подвижность падает с повышением температуры (рассеяние на фононах), как и в металлах. Однако n в полупроводниках экспоненциально увеличивается с повышением температуры, и это обуславливает экспоненциальную зависимость ρ от T .

Удельное сопротивление металлов и полупроводников можно изменить с помощью внешних воздействий (механических, облучения, освещения). В металлах эти воздействия, как правило, уменьшают подвижность μ , и удельное сопротивление увеличивается, причем абсолютная величина этого изменения составляет от нескольких сотых процента до десяти процентов. В полупроводниках μ также меняется, но, кроме того, указанные воздействия изменяют концентрацию носителей заряда, причем это изменение может достигать нескольких порядков. Особенно следует выделить высокую чувствительность полупроводников к освещению, что является основой создания полупроводниковых фотоприемников.

Для полупроводников характерна высокая чувствительность физико-химических свойств к содержанию химических примесей и структурных дефектов, для металлов — существенно меньшая. Действительно, примеси и структурные дефекты могут сильно влиять на физические свойства полупроводников, например, на электрические (проводимость). Одни типы примесей и структурных дефектов не дают никакого эффекта или он совершенно незначителен, другие могут привести к увеличению проводимости на несколько порядков. Действие таких примесей и струк-

турных дефектов состоит в основном в том, что они изменяют число носителей заряда. Возможность изменять электрические свойства полупроводника путем точной дозировки добавляемых примесей или путем получения контролируемого числа структурных дефектов имеет огромное технологическое значение и дает основу для массового производства полупроводниковых приборов. Более того, само требование высокой степени чистоты и структурного совершенства явилось одной из основных причин того, что широкое применение полупроводников и развитие полупроводниковой техники стали возможны только с конца 40-х годов этого века, хотя первые упоминания о полупроводниках относятся еще к 1833 году.

Таким образом, различие в типах химической связи приводит к резкому различию энергетических и кристаллических структур материалов и их физических свойств.

1.2. Основные группы полупроводниковых материалов

В настоящее время известно 104 химических элемента, из них 79 металлов и 25 неметаллов. Среди последних 13 элементов проявляют полупроводниковые, а остальные 12 — диэлектрические свойства (см. табл. 1.1).

Но кроме элементарных полупроводников насчитываются сотни и даже тысячи соединений, твердых растворов, обладающих полупроводниковыми свойствами. Поэтому было бы целесообразно классифицировать полупроводниковые материалы.

Все полупроводники могут быть разделены на *кристаллические* и *некристаллические*.

В настоящем курсе не будут рассматриваться особенности материаловедения и технологии некристаллических полупроводников, которые обычно выделяются в особый раздел физики полупроводников и о которых можно прочитать в специальной литературе [7, 8]. Однако для полноты классификации дадим о них самые общие сведения.

К числу некристаллических полупроводников относятся: аморфные полупроводники (*a*-Si, *a*-Ge, Se, Sb, Te); жидкие полупроводники (расплавленный Se, Te, расплавы оксидов, сульфидов, селенидов и теллуридов некоторых металлов); стеклообразные полупроводники.² Наибольшее распространение среди стеклообразных полупроводников получили халь-

²Стекла — сильно переохлажденные жидкости с очень большой вязкостью.

когенидные (сплавы P, As, Sb, Bi с S, Se, Te, например, As_2Se_3 , As_2S_3 , As_2Te_3) и оксидные ($\text{V}_2\text{O}_5 - \text{P}_2\text{O}_5 - \text{MeO}_x$).³

Общей чертой некристаллических полупроводников является отсутствие дальнего порядка в расположении атомов, результатом которого оказываются существенные изменения в энергетическом спектре и низкая подвижность носителей заряда.

Исследования дифракции рентгеновских лучей, электронов и исследование дальней тонкой структуры рентгеновских спектров поглощения (EXAFS-спектроскопия) показывают, что ближайшее окружение и межатомные расстояния в некристаллических полупроводниках практически не отличаются от их кристаллических аналогов. Таким образом, можно сделать вывод, что полупроводниковые свойства материалов обусловлены ближним порядком (типом связи), а не дальним порядком (кристаллической структурой материалов).

Среди кристаллических полупроводников часто выделяют группы со сходными теми или иными свойствами. Так по количеству входящих в состав вещества компонент можно выделить:

1) элементарные (гомополярные) полупроводники: C, Si, Ge, α -Sn, P, As, Sb, Bi, S, Se, Te, I, В. Важнейшими представителями этой группы являются Ge и Si — основные материалы полупроводниковой электроники. Германий и кремний имеют кубическую структуру типа алмаза, где каждый атом окружен 4 ближайшими соседями (см. гл. 2).

2) гетерополярные полупроводники (бинарные, трехкомпонентные и так далее).

В этом курсе будут рассматриваться в основном только элементарные и бинарные полупроводники, что позволит на простых примерах выявить общие закономерности, характерные для полупроводников, и более детально разобраться в физико-химических проблемах, которые возникают при получении этих материалов. Кроме того, следует отметить, что в электронной промышленности из-за сложностей получения чистых многокомпонентных материалов, как правило, применяются элементарные и бинарные полупроводники.

Среди бинарных полупроводников выделим наиболее изученные и широко используемые группы полупроводниковых соединений. В одну группу обычно включают соединения с одинаковой стехиометрической формулой, то есть образованные из элементов, расположенных в одних и тех же группах периодической системы элементов Д. И. Менделеева.

а) Алмазоподобные полупроводники.

³Me — металлы I–VI групп.

Таблица 1.1. Фрагмент таблицы периодической системы Д. И. Менделеева. Для каждого элемента в нижней строке указаны его атомный радиус (в Å) (слева) и величина электроотрицательности по Полингу (справа) [20].

I	II	III	IV	V	VI	VII	VIII			2 He
H 1 0.78 2.1									1.28	
Li 3 1.52 1.0	Be 4 1.13 1.5	B 5 1.31 2.0	C 6 0.77* 2.5	N 7 0.71 3.0	O 8 0.66* 3.5	F 9 0.58* 4.0			10 Ne	
Na 11 1.54 0.9	Mg 12 1.60 1.2	Al 13 1.43 1.5	Si 14 1.17 1.8	P 15 1.15 2.1	S 16 1.04* 2.5	Cl 17 0.99* 3.0			18 Ar 1.74	
K 19 2.27 0.8	Ca 20 1.97 1.0	Sc 21 1.61 1.3	Ti 22 1.45 1.5	V 23 1.32 1.6	Cr 24 1.25 1.6	Mn 25 1.24 1.5	Fe 26 1.24 1.8	Co 27 1.25 1.8	Ni 28 1.25 1.8	
29 Cu 1.28 1.9	30 Zn 1.33 1.6	31 Ga 1.22 1.6	32 Ge 1.22 1.8	33 As 1.25 2.0	34 Se 1.17* 2.4	35 Br 1.14* 2.8			36 Kr 1.89*	
Rb 37 2.47 0.8	Sr 38 2.15 1.0	Y 39 1.81 1.2	Zr 40 1.60 1.4	Nb 41 1.43 1.6	Mo 42 1.36 1.8	Tc 43 1.36 1.9	Ru 44 1.34 2.2	Rh 45 1.34 2.2	Pd 46 1.38 2.2	
47 Ag 1.44 1.9	48 Cd 1.49 1.7	49 In 1.63 1.7	50 Sn 1.40 1.8	51 Sb 1.82 1.9	52 Te 1.43 2.1	53 I 1.33* 2.5			54 Xe 2.18 2.6	
Cs 55 2.65 0.7	Ba 56 2.17 0.9	La 57 1.88 1.1	Hf 72 1.56 1.3	Ta 73 1.43 1.5	W 74 1.37 1.7	Re 75 1.37 1.9	Os 76 1.35 2.2	Ir 77 1.36 2.2	Pt 78 1.38 2.2	
79 Au 1.44 2.4	80 Hg 1.60 1.9	81 Tl 1.70 1.8	82 Pb 1.75 1.8	83 Bi 1.55 1.9	84 Po 1.67 2.0	85 At 2.2			86 Rn	
Fr 87 2.7 0.7	Ra 88 2.23 0.9	Ac 89 1.88 1.1	Ku 104							

Стехиометрическая формула этих полупроводников — $A^N B^{8-N}$, где A и B — компоненты соединения, N — номер группы компонента A в периодической таблице Менделеева. Компоненты этих соединений расположены симметрично относительно IV группы в таблице Менделеева.

Соединения имеют гранецентрированную кубическую решетку и кристаллизуются в структурах типа алмаза или сфалерита (цинковой обманки), либо гексагональную решетку и кристаллизуются в структуре типа вюрцита.

Физические свойства соединений элементов VI группы в значительной мере определяются концентрацией собственных точечных дефектов структуры (см. гл. 3), проявляющих электрическую активность (центры рассеяния и рекомбинации).

Полупроводники $A^N B^{8-N}$ находят широкое применение во многих областях техники и, прежде всего, в оптоэлектронных приборах.

К ним относятся:

$A^{III} B^V$ — GaAs, InSb, InP, GaSb, GaP, AlSb (структура сфалерита).

$A^{II} B^{VI}$ — CdSe, CdTe, HgTe, ZnS, ZnSe (структура сфалерита); CdS, CdSe, ZnS, ZnSe (структура вюрцита).⁴

$A^{I} B^{VII}$ (кристаллизующиеся в кубической структуре типа сфалерита) — CuBr, AgI.

$A^{IV} B^{IV}$ — β -SiC (структура сфалерита).

б) Халькогениды элементов четвертой группы.

Стехиометрическая формула этих полупроводников — $A^{IV} B^{VI}$. Типичные представители — PbS, PbTe, SnTe.

Эти материалы имеют гранецентрированную кубическую решетку типа NaCl или слабо искаженные структуры на ее основе. Они имеют узкую запрещенную зону и используются как источники и приемники ИК-излучения. Собственные точечные дефекты структуры в $A^{IV} B^{VI}$ имеют низкую энергию ионизации и проявляют электрическую активность.

в) Халькогениды элементов пятой группы.

Стехиометрическая формула этих материалов — $A_2^V B_3^{VI}$. К этой группе относятся халькогениды мышьяка (As_2S_3 , As_2Se_3 , As_2Te_3), кристаллизующиеся в структуры с моноклинной решеткой; халькогениды сурьмы (Sb_2Se_3 , Sb_2S_3) — соединения с ромбической решеткой; халькогениды висмута (Bi_2Te_3 , Bi_2Se_3) — соединения с ромбоэдрической решеткой.

Халькогениды сурьмы и висмута используются для создания полупроводниковых термоэлектрических генераторов.

⁴Обратите внимание на то, что некоторые из приведенных полупроводников существуют в двух кристаллических модификациях.

Среди гетерополярных полупроводников принято также выделять следующие группы полупроводников:

а) Магнитные полупроводники [9, 10].

Во всех этих материалах присутствуют атомы с частично заполненными d - и f -оболочками (атомы переходных металлов и редкоземельных элементов). Отличительной особенностью многих из них является возникновение магнитного упорядочения при низких температурах.

Эта группа включает такие соединения как, например, EuS , EuSe , MnTe , простые (NiO , CoO , FeO , ZnO , EuO , Cu_2O) и сложные окислы переходных металлов, например, ферриты типа $\text{MeO} \cdot \text{Fe}_2\text{O}_3$ или MeFe_2O_4 (ZnFe_2O_4 , MnFe_2O_4), кристаллизующиеся в структуре шпинели.

Эти материалы находят применение в радиотехнических приборах, оптических устройствах, управляемых магнитным полем, и в волноводных устройствах СВЧ. Оксидные полупроводники (NiO , ZnO , EuO , Cu_2O) используются как разнообразные датчики — температуры, химического состава газа.

б) Полупроводники-сегнетоэлектрики [11, 12].

Отличительными особенностями этой группы кристаллов является наличие электрических моментов в кристалле и возникновение спонтанной поляризации при понижении температуры.

Типичным примером этих материалов являются полупроводники со структурой перовскита и стехиометрической формулой ABO_3 . Их примером могут служить BaTiO_3 и PbTiO_3 . К этой же группе полупроводников-сегнетоэлектриков относятся и некоторые соединения $\text{A}^{\text{IV}}\text{B}^{\text{VI}}$ — GeTe и SnTe .

Используются эти материалы при создании запоминающих и нелинейно-оптических устройств, в качестве пьезодатчиков.

в) Органические полупроводники: антрацен, нафталин, фталоцианины, поливинилкарбазол [13, 14].

В последние годы наметился существенный прогресс в использовании этих материалов в разных областях техники, например, для создания транзисторов и оптоэлектронных приборов.

Такое деление полупроводников можно было бы продолжить и дальше, однако остановимся на приведенных наиболее часто выделяемых группах.

Изучение курса основы материаловедения и технологии полупроводников начнем с обсуждения природы химических связей и анализа особенностей химической связи в полупроводниках.

Глава 2

Основы теории химической связи и закономерности образования полупроводников

2.1. Электронная конфигурация свободных атомов

Химическая связь в твердых телах образуется в результате взаимодействия атомов (ионов). Наиболее существенным результатом этого взаимодействия является расщепление энергетических уровней валентных электронов свободных атомов и образование энергетических зон. С другой стороны, взаимодействие электронов данного атома и соседних атомов не разрушает полностью исходную структуру электронных уровней отдельных атомов. Эти два факта дают основание считать, что электронное строение свободных атомов и, прежде всего, строение их валентных оболочек определяет химическую связь, характер ближнего порядка и, в конечном счете, электронные свойства твердых тел. Периодическая таблица элементов Д. И. Менделеева отражает закономерности электронного строения свободных атомов.

Структура энергетического спектра свободных атомов рассчитывается с помощью методов квантовой механики. Простейшим атомом является атом водорода, обладающий единственным электроном. Уравнение Шредингера для него имеет точное решение и позволяет провести классификацию состояний электрона в атоме. Отметим, что изучение атома водорода дает основу для классификации состояний любого атома, так как электронная структура атома водорода имеет ряд существенных черт, характерных для электронной структурой более сложных атомов.

Уравнение Шредингера для атома водорода можно записать в виде

$$\hat{H}\psi = E\psi, \quad (2.1)$$

где $\hat{H} = \hat{p}^2/2m + U$ — гамильтониан ($\hat{p} = -i\hbar\nabla$ — оператор импульса, $U(\mathbf{r}) = -e^2/r$ — потенциальная энергия электрона в поле, созданном неподвижным ядром, m — масса электрона, e — заряд электрона, $\mathbf{r}$ — радиус-вектор электрона), $\psi(\mathbf{r})$ — волновая функция электрона ($|\psi(\mathbf{r})|^2$ определяет плотность вероятности нахождения электрона в данной точке пространства), а E — энергия состояния электрона.

Для сферически симметричного потенциала $U(r)$ уравнение (2.1) проще всего решать в сферической системе координат. В этом случае уравнение Шредингера для атома водорода допускает разделение переменных. Собственные волновые функции уравнения (2.1) ищутся в виде $\psi = R(r)\Theta(\vartheta)\Phi(\varphi)$, где r , ϑ , φ — сферические координаты. Решение этих уравнений совместно с соответствующими граничными условиями приводит к появлению трех целочисленных квантовых чисел n , l , m_l , которые служат параметрами для собственных волновых функций уравнения (2.1), описывающих состояния электрона в атоме. Учет спина электрона приводит к появлению четвертого квантового числа m_s . Полная волновая функция электрона равна произведению координатной и спиновой волновых функций. Пространственная (координатная) часть волновой функции электрона в атоме называется *атомной орбиталью*.

Таким образом, собственные волновые функции уравнения (2.1) зависят от четырех параметров (четырёх квантовых чисел): n , l , m_l , m_s .

Волновые функции электрона, имеющие одинаковые n , l и m_l , образуют орбиталь. Совокупность орбиталей, имеющих одинаковые n и l , образует оболочку. Совокупность оболочек с одинаковым n образует электронный слой.

Главное квантовое число n может принимать значения 1, 2, 3, ... и определяет номер слоя. Слои с $n = 1, 2, 3, 4, \dots$ принято обозначать буквами K, L, M, N и т. д. Величина n характеризует удаленность электрона от атомного ядра.

Орбитальное квантовое число l определяет величину момента количества движения электрона относительно ядра (или симметрию электронного облака); l может иметь значения от 0 до $(n - 1)$. Оболочки с $l = 0, 1, 2, 3$ называются *s*-, *p*-, *d*-, *f*-орбиталями соответственно.

Орбитальное магнитное квантовое число m_l характеризует проекцию момента количества движения электрона на некоторое выделенное направление в пространстве; m_l может принимать значения от $-l$ до l (включая ноль).

Спиновое магнитное квантовое число m_s характеризует собственный механический момент электрона (ориентацию в пространстве вектора спина электрона); m_s может принимать только два значения: $+1/2$ и $-1/2$.

Можно показать, что уравнение (2.1) имеет решения при любых положительных значениях энергии и при дискретных отрицательных значениях энергии, равных

$$E_n = -\frac{me^4}{2\hbar^2} \frac{1}{n^2} = -\frac{e^2}{2a_0} \frac{1}{n^2} = -\frac{13.6}{n^2} \text{ эВ}, \quad (2.2)$$

где $a_0 = \hbar^2/me^2 = 0.529 \text{ \AA}$ — боровский радиус. Случай $E > 0$ соответствует делокализованным состояниям, характеризваемым волновым вектором и проекцией спина (электрон пролетает вблизи ядра и удаляется вновь на бесконечность). Случай $E < 0$ соответствует локализованным электронным состояниям, характеризваемым набором квантовых чисел n , l , m_l и m_s (электрон находится в пределах атома). Видно, что значения энергии для интересующего нас второго случая зависят от квантового числа n и не зависят от l , m_l и m_s . Именно поэтому число n и называется главным квантовым числом. Таким образом, каждому E_n соответствует несколько волновых функций ψ_{n,l,m_l,m_s} , отличающихся значениями квантовых чисел l , m_l , m_s . Это означает, что атом водорода может иметь одно и то же значение энергии, находясь в нескольких различных состояниях, то есть энергетические уровни вырождены, причем кратность вырождения равна $2n^2$.

Рассмотрим пространственные части (без учета спина электрона) собственных функций, являющихся решениями волнового уравнения (2.1), в порядке возрастания энергии, то есть n . Функции $|\psi_{n,l,m_l}|^2$ описывают пространственное распределение электронной плотности вокруг ядра и имеют независимые друг от друга радиальные $R(r)$ и угловые $\Theta(\vartheta)\Phi(\varphi)$ части.

Радиальная часть $R(r)$ волновой функции определяет распределение электронной плотности по радиусу от ядра. На рис. 2.1,а показаны радиальные части первых трех s -орбиталей. Отметим, что $1s$ -функция никогда не меняет знака, $2s$ -функция меняет его один раз, а $3s$ -функция — дважды. На рис. 2.1,б показано изменение электронной плотности с расстоянием от ядра для этих орбиталей. Электронная плотность концентрируется на все большем расстоянии от ядра по мере того, как значение n растет. Обычно радиальная функция плотности имеет самое большое значение за последней поверхностью, на которой $R(r) = 0$. Для p -, d - и f -орбиталей картины похожи.

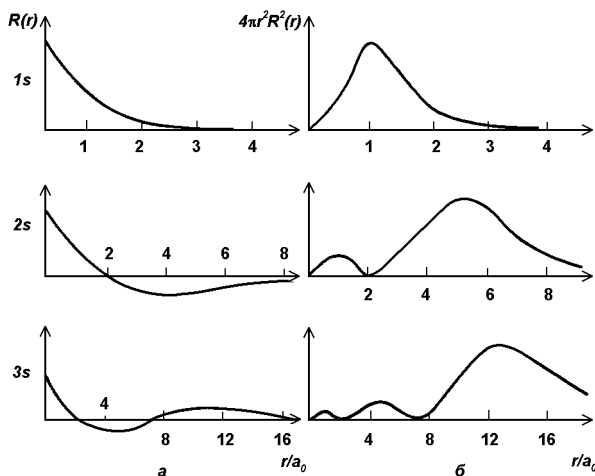


Рис. 2.1. Зависимость радиальной части волновой функции $R(r)$ от r/a_0 для трех первых s -орбиталей (а) и соответствующие радиальные функции распределения электронной плотности $4\pi r^2 R^2(r)$ (б).

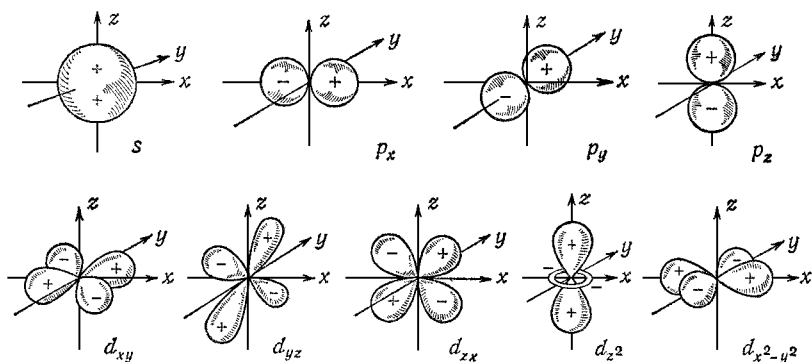
Рассмотрим угловые части $\Theta(\vartheta)\Phi(\varphi)$ волновых функций. Они определяют угловое распределение электронной плотности — вероятность найти электрон в определенном направлении, рассчитанную на единицу телесного угла.

s -орбитали. В каждом электронном слое существует лишь одна s -орбиталь. Все s -орбитали сферически симметричны (рис. 2.2), так как все они описываются волновой функцией, которая не зависит от углов ϑ и φ (таково решение волнового уравнения).

p -орбитали. Начиная с $n = 2$ появляются p -орбитали. Каждый электронный слой, у которого существует $l = 1$, имеет $2l + 1 = 3$ p -орбитали, которые зависят от углов ϑ и φ . Эти три p -орбитали называются p_x , p_y , p_z , так как они полностью симметричны относительно осей x , y , z соответственно, но меняют знак при инверсии (на рис. 2.2 это показано с помощью знаков $+$ и $-$).

Для $n = 3$ и 4 появляются еще d - ($l = 2$) и f -орбитали ($l = 3$) соответственно. При этом каждый электронный слой должен иметь пять d -орбиталей (рис. 2.2) и семь f -орбиталей. Все орбитали с одним и тем же n имеют сходное радиальное распределение электронной плотности и существенно различные угловые распределения.

Такова классификация состояний электрона в атоме водорода.

Рис. 2.2. Орбитали s -, p - и d -типа.

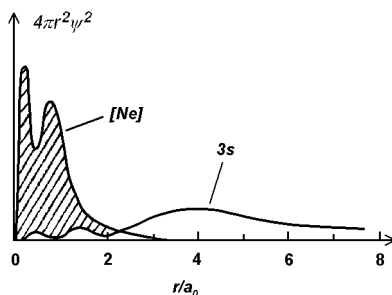
Перейдем к рассмотрению свободных многоэлектронных атомов. Структура энергетического спектра сложных атомов обладает значительным сходством со структурой спектра атома водорода. Физическая причина этого сходства заключается в том, что и в более сложных атомах каждый электрон находится под действием приблизительно центральных сил с центром, расположенным на ядре. Как и в случае атома водорода, потенциальная энергия отдельного электрона является и в рассматриваемом случае функцией только расстояния от ядра. Уравнение Шредингера для многоэлектронного атома, как и для атома водорода, решается в сферических координатах методом разделения переменных, и при его решении находятся аналогичные наборы значений квантовых чисел n , l , m_l , m_s ,¹ а выражения для угловых частей собственных волновых функций этого уравнения оказываются точно такими же, как и в случае атома водорода. Однако энергия отдельного электрона теперь оказывается зависящей не только от n , но и от l ² что приводит к расщеплению атомных уровней энергии.

Обсудим причины, по которым снимается вырождение по l в сложных атомах. При расчетах структуры энергетического спектра многоэлектронного атома предполагается, что выделенный электрон в нем не взаимодействует с другими электронами, а ведет себя так, как если бы он был один в электростатическом поле ядра. Однако в этом случае потенциал, действующий на этот выделенный электрон, принципиально отличается

¹ Однако теперь l , m_l , и m_s относятся к суммарным моментам всех электронов в атоме.

² Энергия не зависит от m_l и m_s по соображениям симметрии в данной постановке задачи, однако реально такая зависимость существует (см. ниже).

Рис. 2.3. Проникновение $3s$ -орбитали в неоновый электронный остов ($[\text{Ne}]$) атома Na .



от кулоновского потенциала в атоме водорода из-за влияния электронов друг на друга. Согласно законам общей физики потенциальная энергия электрона U , находящегося на определенной орбитали в поле сферически-симметричного распределения заряда, пропорциональна Z , где Z — полный заряд, содержащийся внутри сферы, радиус которой равен расстоянию от ядра до электрона. Этот заряд Z состоит из заряда самого ядра минус заряд электронов, находящихся на более близких к ядру орбиталях, чем рассматриваемый электрон. Однако на величину заряда Z , определяющего волновую функцию электрона на рассматриваемой орбитали и его энергию в многоэлектронном атоме, еще оказывает влияние степень проникновения волновой функции этой орбитали в заполненный остов. Поясним этот эффект. В водородоподобном атоме энергия электрона на данной орбитали определяется только главным квантовым числом n и полным зарядом Z : $E_n \sim Ze^2/n^2$, то есть энергии, например, $2s$ - и $2p$ -орбиталей должны быть одинаковы. В многоэлектронном атоме ситуация иная. Так, например, у атома Li уровень $n = 2$ (основное состояние третьего электрона) не является вырожденным, как это было в случае атома водорода. Вместо этого $2s$ -состояния располагаются несколько ниже $2p$ -состояний. Основной причиной этой зависимости энергии от l является то обстоятельство, что волновая функция $2s$ -электрона Li проникает внутрь гелиевого остова больше, чем волновая функция $2p$ -электрона, и при этом заряд ядра экранируется меньше. Аналогичная ситуация наблюдается и в атоме Na . Энергии $3s$ -, $3p$ -, $3d$ -орбиталей значительно различаются, а порядок их расположения в энергетическом пространстве следующий: $3s$, $3p$, $3d$. Это связано с тем, что волновая функция $3s$ -электрона Na значительно проникает внутрь неоновой оболочки, при этом заряд ядра вместо того, чтобы экранироваться полностью электронами неоновой оболочки, экранируется частично

(рис. 2.3). В результате получается, что эффективный (результатирующий) заряд Z^* (ранее полный заряд Z) равен ≈ 2 вместо 1. Волновая функция $3p$ -электрона проникает в неоновый остов меньше, чем $3s$ -, и эффективный заряд становится $1 < Z^* < 2$; волновая функция $3d$ -электрона практически совсем не проникает и, следовательно, $Z^* \approx 1$. Для атома калия ситуация следующая: $4s$ -орбиталь менее экранирована, чем $3d$ -, а поэтому более стабильна и, следовательно, порядок расположения орбиталей в энергетическом пространстве такой: $3s$, $3p$, $4s$, $3d$, $4p$.

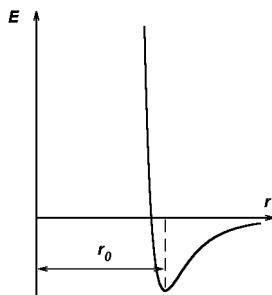
Таким образом, экранирование заряда ядра в многоэлектронном атоме приводит к расщеплению уровней энергии, то есть к снятию вырождения по l . Следует отметить, что экранирование не единственная причина, приводящая к расщеплению энергетических уровней; этот же эффект вызывают малые магнитные силы, действующие внутри атомов, а также и другие специфические квантовомеханические взаимодействия.

Принцип заполнения орбиталей электронами в многоэлектронных атомах состоит в добавлении электронов к ядру атома, то есть в заполнении соответствующего числа орбиталей в порядке возрастания квантового числа n и в согласии с принципом Паули и правилами Хунда и Клечковского. Согласно *принципу Паули* на одной орбитали может находиться не более двух электронов с антипараллельными спинами. Отсюда следует, что максимальное число электронов на s , p , d и f -орбиталях, образующих оболочку, равно 2, 6, 10 и 14 ($2(2l + 1)$) соответственно. Согласно *правилу Хунда* данная электронная конфигурация обладает наименьшей энергией при максимальном числе неспаренных валентных электронов. Согласно *правилу Клечковского* орбитали заполняются в порядке возрастания $n + l$, причем в группе орбиталей с одинаковым значением $n + l$ первыми заполняются орбитали с меньшим значением n : $ns < (n - 1)d \approx (n - 2)f < np$.

2.2. Типы химической связи

При сближении изолированных атомов и образовании кристалла их электронные оболочки перекрываются. Взаимодействие электронов, принадлежащих разным атомам, вызывает расщепление атомных дискретных энергетических уровней изолированных атомов. Первыми начинают расщепляться уровни, соответствующие электронам, наиболее удаленным от ядра, то есть находящимся на внешних оболочках атома. При равновесном межатомном расстоянии r_0 расщеплены уровни всех орбиталей, участвующих в образовании химической связи. Число расщепленных уровней равно числу атомов в кристалле. Каждый тип атомных ор-

Рис. 2.4. Зависимость полной энергии твердого тела E от межатомного расстояния r ; r_0 — равновесное межатомное расстояние.



биталей группируется в определенное число «молекулярных орбиталей», каждая из которых включает большое число атомных расщепленных орбиталей, очень мало отличающихся друг от друга по энергии. Такие «молекулярные орбитали» образуют энергетические зоны. Таким образом, в результате взаимодействия атомов формируется энергетическая структура твердого тела.

Что же удерживает вместе атомы в кристалле? Химическая связь между ними в основном обеспечивается силами электростатического притяжения между отрицательно заряженными электронами и положительно заряженными ядрами. Для того чтобы она образовалась, необходимо выполнение следующих условий.

Положительно заряженные ионные остовы и валентные электроны должны находиться на таком расстоянии друг от друга, чтобы было сведено до минимума электростатическое отталкивание между одноименными зарядами. В то же самое время валентные электроны должны быть настолько близко расположены от положительных ионов, чтобы электростатическое притяжение между разноименными зарядами было максимально. При выполнении этих условий потенциальная энергия системы может немного уменьшиться, а кинетическая может немного возрасти, однако полная энергия должна отвечать минимуму энергии системы (рис. 2.4). В результате ионные остовы будут находиться в твердом теле на равновесном межатомном расстоянии r_0 , которое определяется равенством сил притяжения и отталкивания между ионными остовами и электронами.

Таким образом, существование стабильной химической связи между атомами в кристалле предполагает, что полная энергия кристалла меньше полной энергии такого числа свободных атомов, из которого состоит кристалл, то есть при образовании химической связи всегда вы-

деляется энергия. Количество энергии, выделившейся при образовании химической связи, называется *энергией связи*: $E_{\text{связ}} = E_{\text{своб.ат}} - E_{\text{т.т.}}$. Эта величина является важнейшей характеристикой прочности связи и выражается в килоджоулях на 1 моль образующегося вещества. Энергию связи определяют сравнением с состоянием, предшествующим ее образованию. Определить ее непосредственно из эксперимента довольно трудно, обычно энергию связи вычисляют из данных по теплотам реакций с использованием закона Гесса [15]. Иногда энергию связи можно оценить по теплоте сублимации.

В зависимости от особенностей (определяемых индивидуальной электронной структурой взаимодействующих атомов) сил, действующих между атомами или ионами, в элементарных веществах различают три основных типа химической связи: металлическую, ковалентную и молекулярную. Помимо перечисленных в соединениях и твердых растворах возможен ионный тип связи. В веществах, одним из компонентов которых является водород, возможно появление водородного типа связи. Как правило, в реальных материалах одновременно реализуется несколько различных типов связи. Образующаяся в результате связь называется смешанной.

Рассмотрим особенности основных типов химической связи.

2.2.1. Ковалентная связь

Химическая связь между атомами, осуществляемая обобществленными электронами, называется *ковалентной*. Эта связь обусловлена силами квантовомеханического происхождения — обменным взаимодействием.

Рассмотрим образование гомеоплярной ковалентной связи на примере простейшей молекулы водорода H_2 . В квантовой механике один из методов рассмотрения электронного строения молекул основан на представлении об образовании химической связи в результате движения каждого электрона в поле всех ядер и остальных электронов молекулы. В таком одноэлектронном приближении многоэлектронная волновая функция молекулы представляет собой совокупность одноэлектронных волновых функций (молекулярных орбиталей — МО), каждая из которых описывает один электрон молекулы в определенном состоянии. МО задается определенным набором квантовых чисел и для нее справедлив принцип Паули. При этом сама одноэлектронная МО получается как линейная комбинация одноэлектронных атомных орбиталей (АО). Физическая суть этого метода заключается в следующем (для молекулы водорода). Во время движения электрона вокруг ядерного скелета молекулы H_2 в какой-то

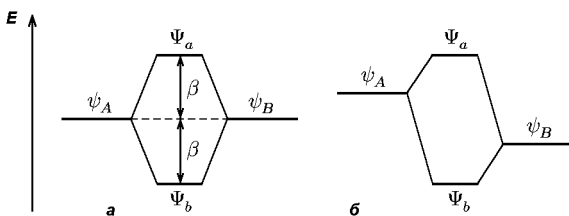


Рис. 2.5. Диаграмма уровней энергии МО для двухатомных молекул, образованных двумя идентичными (а) и различными (б) атомами.

момент электрон может оказаться вблизи одного ядра (назовем его A) и относительно далеко от другого (назовем его B) и наоборот. Когда он оказывается рядом с данным ядром, он может вести себя более или менее так, как будто он находится на атомной орбитали, принадлежащей этому ядру (ψ_A или ψ_B). Если электрон находится в некотором среднем положении относительно скелета из ядер, то его состояние может быть приближенно описано комбинацией обеих функций, то есть $\psi_A \pm \psi_B$.

В этом приближении две нормированные волновые функции (МО) молекулы H_2 для одного электрона имеют вид

$$\Psi_b = N_b(\psi_A + \psi_B), \quad \Psi_a = N_a(\psi_A - \psi_B), \quad (2.3)$$

где ψ_A и ψ_B — волновые функции электрона атома A и B соответственно, $N_{a,b}$ — постоянные нормировки.

Поскольку атомы A и B идентичны, из уравнения Шредингера следует, что энергия орбитали Ψ_b дается выражением $E_b = 2N_b^2(Q + \beta)$, где N_b — постоянная нормировки, $Q = \int \psi_A^* \hat{H} \psi_A dV = \int \psi_B^* \hat{H} \psi_B dV$ — энергия электрона на орбитали ψ_A или ψ_B , то есть энергия атома водорода в основном состоянии; $\beta = \int \psi_A^* \hat{H} \psi_B dV = \int \psi_B^* \hat{H} \psi_A dV$ — обменный интеграл. Параметр β представляет собой энергию взаимодействия между атомами и имеет отрицательное значение. Аналогично для другой орбитали имеем $E_a = 2N_a^2(Q - \beta)$, где N_a — постоянная нормировки. Энергия Ψ_b МО оказывается ниже, чем у исходных АО. Ее заполнение приводит к образованию химической связи между атомами. Это основное состояние молекулы, поэтому Ψ_b называют *связывающей* МО, а находящиеся на ней электроны связывающими электронами. Энергия МО Ψ_a выше, чем у исходных АО. Заполнение этой орбитали электронами ведет к разрыхлению химической связи и распаду молекулы на атомы. Такое состояние молекулы можно рассматривать как возбужденное, поэтому МО Ψ_a называют *разрыхляющей*, а находящиеся на ней электроны разрыхляющи-

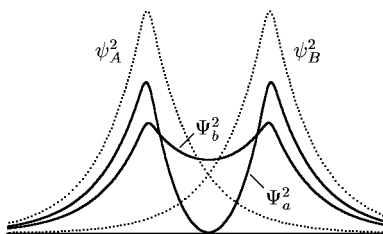


Рис. 2.6. Плотность вероятности нахождения электрона на индивидуальных атомных орбиталях водорода ψ_A^2 и ψ_B^2 (штриховые линии) и на молекулярных орбиталях Ψ_b^2 и Ψ_a^2 (сплошные линии) в молекуле водорода.

ми электронами. Результаты расчетов приведены на диаграмме уровней энергии на рис. 2.5,а.

На рис. 2.6 приведены плотность вероятности нахождения электрона на индивидуальных атомных орбиталях ψ_A и ψ_B и на молекулярных орбиталях Ψ_b и Ψ_a (с учетом спиновой части волновой функции электрона). Функция Ψ_b соответствует электронам с антипараллельно направленными спинами, а Ψ_a — с параллельно направленными спинами. Из рисунка видно, что электрон с волновой функцией Ψ_b имеет более высокую вероятность нахождения между ядрами, чем электрон с волновой функцией Ψ_a . Увеличение электронной плотности между двумя положительно заряженными ядрами сопровождается значительным выигрышем энергии по сравнению с невзаимодействующими атомами водорода, что и является причиной образования химической связи в молекуле. Образование химической связи в молекуле водорода иллюстрирует формирование гомеополлярной ковалентной связи.

В молекулах, образованных из одинаковых атомов, различают два типа МО, формы которых можно представить исходя из форм исходных атомных орбиталей: это молекулярные орбитали σ - и π типов. σ -МО возникает при перекрытии атомных орбиталей вдоль оси, соединяющей ядра взаимодействующих атомов. Например, в молекуле H_2 мы имеем дело с σ -МО (связывающей и разрыхляющей), образующимися при перекрытии s -орбиталей каждого из двух атомов (рис. 2.7,а,б). Если при образовании МО необходимо комбинировать p -орбитали, то при этом могут образовываться σ -МО и π -МО. Например, примем ось, проходящую через два ядра, за ось z . Если при взаимодействии атомов перекрываются p_z -орбитали соседних атомов, то могут образовываться связывающая и разрыхляющая σ -МО (рис. 2.7,в,г). При комбинации p_z -орбитали одного атома и s -орбитали другого также образуется σ -МО. С образованием σ -МО могут комбинировать s -, p - и d -АО, если они ориентированы вдоль оси связи.

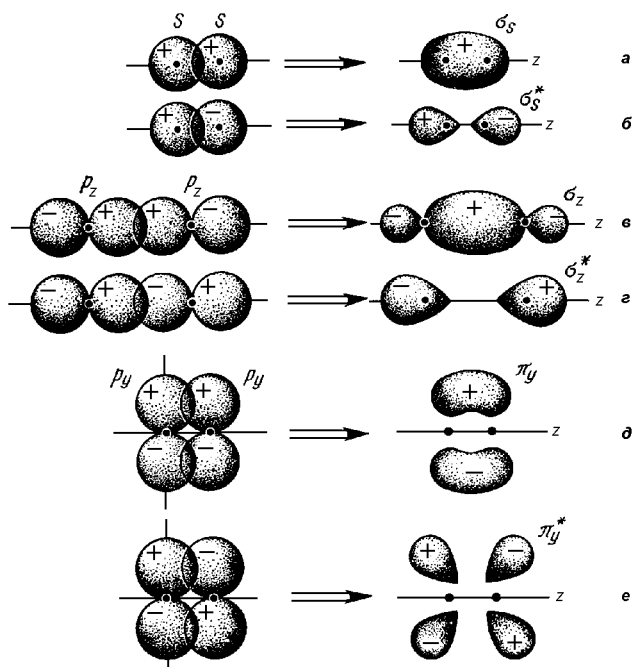


Рис. 2.7. Примеры образования молекулярных σ - и π -орбиталей. а, в, д — связывающие орбитали; б, г, е — разрыхляющие орбитали.

Однако p_x - и p_y -орбитали уже не могут образовывать σ -МО. При комбинировании они образуют π -орбитали. π -орбиталь — это такой тип МО, центроузловая линия которой совпадает с молекулярной осью. При взаимодействии двух p -орбиталей, расположенных перпендикулярно оси, содержащей ядра атомов, возникают две области перекрытия. Соответственно π -МО не обладает цилиндрической симметрией относительно оси, но имеет равное распределение электронной плотности по обе стороны от плоскости, проходящей через эту ось (рис. 2.7, д, е). π -МО также могут образовываться при перекрытии p - и d -орбиталей или двух d -орбиталей.

Нетрудно догадаться, что гомеополярная ковалентная связь тем прочнее, чем сильнее перекрываются орбитали. При этом связывающие электроны больше концентрируются между ядрами, уменьшается взаимное отталкивание ядер и возрастают силы притяжения между электронами и ядрами. Степень перекрытия атомных орбиталей характеризуется интегралом перекрытия $S = \int \psi_A^* \psi_B dV$. Он может иметь положительное, отрицательное и нулевое значение. Если $S < 0$, то ядра отталкиваются, при $S = 0$ взаимодействия нет, а при $S > 0$ между ядрами образуется связь (рис. 2.8).

Вернемся к рис. 2.6. На нем изображена типичная диаграмма уровней энергий, показывающая энергии МО, образованные двумя одинаковыми атомами, каждый из которых для образования связи предоставляет одну орбиталь. При этом электронная плотность в молекуле симметрично распределена относительно обоих ядер.

Если двум исходным атомным орбиталям не соответствует одна и та же энергия (такая ситуация возникает, если два атома различны), то уровни, отвечающие Ψ_a и Ψ_b , должны быть смещены приблизительно на одинаковое расстояние выше и ниже средней энергии этих двух атомных орбиталей (см. рис. 2.6, б). Результирующее электронное облако в молекуле в этом случае оказывается асимметричным, и атомы нельзя уже считать незаряженными. Таким образом, кроме обменного взаимодействия между атомами будет еще и кулоновское взаимодействие; образующуюся при этом связь называют ковалентной полярной связью. Для ее образования необходимо, чтобы из исходных АО образовывались МО. МО образуются при соблюдении следующих условий: а) довольно близкие значения энергии исходных АО; б) значительное перекрытие АО; в) одинаковая симметрия АО относительно оси молекулы. Увеличение степени нарушения этих условий приводит к увеличению асимметрии электронного облака и постепенному изменению типа связи. Например, в молекуле HCl

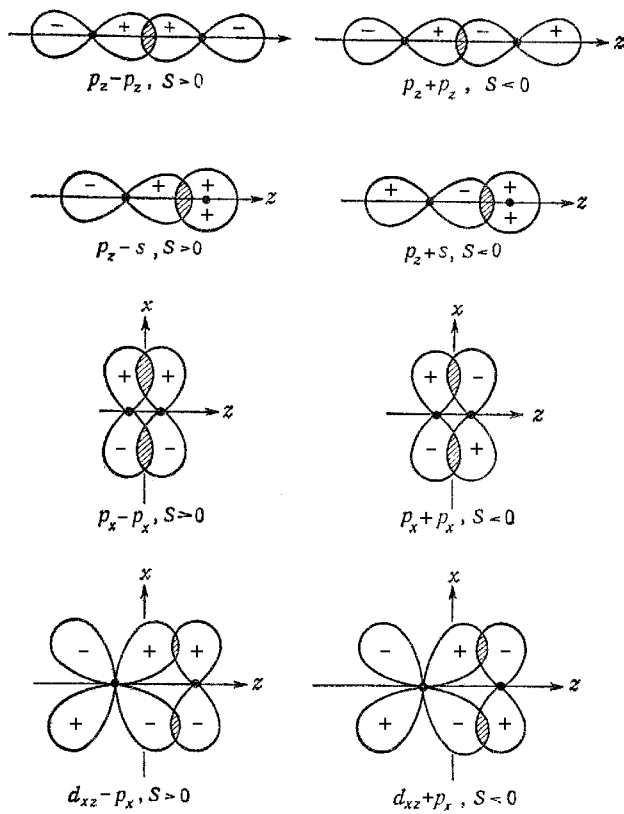


Рис. 2.8. Различные схемы перекрытия атомных орбиталей.

$$\Psi_b = \psi_A + \lambda\psi_B, \quad \Psi_a = \psi_B - \lambda\psi_A. \quad (2.4)$$

Однако для HCl λ настолько мало, что $\Psi_b \approx \psi_A$, а $\Psi_a \approx \psi_B$. Такую полярную ковалентную связь можно рассматривать как переходную между ковалентной и ионной связями. В предельном случае очень сильного различия в энергиях атомных орбиталей, участвующих в образовании связи, эти атомные орбитали уже не будут смешиваться, а образующаяся связь уже не будет ковалентной.

2.2.2. Ионная связь

Ионные кристаллы, то есть те, в которых ионный тип связи является преобладающим, состоят из положительных и отрицательных ионов. Эти ионы образуют кристаллическую решетку в основном за счет электростатического взаимодействия между ионами противоположного знака.

Рассмотрим для простоты простой ионный кристалл.³ Электронные оболочки ионов в таком кристалле похожи на электронные оболочки, характерные для атомов инертных газов, так как атомы, из которых образуется ионный кристалл, достраивают свою внешнюю валентную оболочку до полностью заполненной, присоединяя недостающие электроны (анионы) и отдавая их (катионы). Такой переход электронов оказывается энергетически выгодным (приводит к уменьшению полной энергии системы). Можно ожидать поэтому, что распределение электронной плотности для каждого иона в ионном кристалле будет приближенно обладать сферической симметрией, как у атомов инертных газов, хотя эта симметрия и может несколько нарушаться в области соприкосновения соседних ионов. Этот вывод подтверждается рентгеновскими данными. Так как электрическое поле иона имеет сферическую симметрию, то ионная связь не обладает направленностью. Кроме того, ионная связь не обладает и свойством насыщаемости.

На больших расстояниях взаимодействие между парой противоположно заряженных ионов можно рассматривать как взаимодействие точечных зарядов q с разными знаками. Энергия этого взаимодействия, связанная с силами притяжения, $\sim -q^2/r$, где q — заряд иона, r — расстояние между ионами. Однако в кристаллах надо принимать во внимание взаимодействие между всеми зарядами (как разного знака, так и одного) в структуре. Учет этого, как показывают расчеты [16], приводит к появлению численного множителя α , называемого *постоянной*

³Простой ионный кристалл — это кристалл, имеющий г.ц.к. решетку типа NaCl.

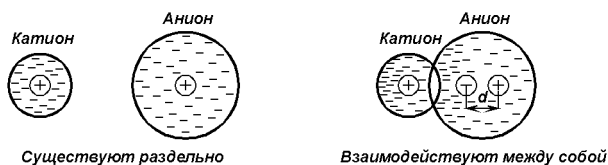


Рис. 2.9. Поляризация ионов (схема).

Маделунга, а энергия взаимодействия в этом случае будет $-\alpha q^2/r$. Значение постоянной Маделунга зависит от типа кристаллической решетки. С другой стороны, при малых r должны проявляться силы отталкивания. Они обусловлены перекрытием электронных облаков ионов, а также взаимным отталкиванием положительно заряженных ядер. Это взаимодействие часто описывают эмпирическим выражением a/r^m , где a и m — постоянные, характеризующие силы отталкивания. Таким образом, полная энергия взаимодействия в ионной кристаллической решетке (энергия решетки) описывается выражением⁴

$$U = a/r^m - \alpha q^2/r. \quad (2.5)$$

Характерная величина энергии связи ионных кристаллов порядка 5–7 эВ на пару ионов.

В ионных кристаллах наблюдается поляризуемость ионов: большие ионы могут деформироваться под действием соседних меньших ионов (рис. 2.9). Например, маленький катион и большой анион. Маленький катион отталкивает ядро большого аниона и при этом индуцирует перераспределение заряда на ближайшей к нему стороне аниона. Вследствие этого плотность собственного электронного облака с этой стороны катиона становится меньше. В результате расстояние между центрами ионов уменьшается по сравнению с расстоянием между «недеформированными» ионами. Сферическая форма аниона нарушается из-за более или менее заметного перекрытия электронных облаков различных ионов в направлениях их ближайших соседей, и центры тяжести отрицательных и положительных зарядов аниона перестают совпадать. Величина этого эффекта характеризуется поляризуемостью χ иона (аниона), которая равна отношению индуцированного дипольного момента μ ($\mu = Zed$, где Ze — перемещаемый заряд, d — длина диполя, Z — номер элемента в таблице Менделеева) к напряженности электрического поля, создавае-

⁴ Данная формула позволяет получать значения энергии близкие к энергии связи для простых ионных кристаллов [16].

мого другим ионом (катионом): $\kappa = \mu/E = Zed/E$. Электрическое поле, вызывающее поляризацию ионов, тем сильнее, чем больше в кристаллах катионов малого радиуса с высоким зарядом.⁵ Поляризация ионов в ионном кристалле сопровождается уменьшением расстояния между ионами и увеличением энергии решетки. Поляризация может вызвать изменение преимущественного типа связи, а соответственно и изменение координационного числа и типа структуры. Так происходит изменение преимущественно типа связи и типа структуры, например, в галоидных соединениях серебра: AgF, AgCl, AgBr имеют ионный тип связи и кристаллизуются в структуру типа NaCl, а AgI — полярную ковалентную связь и кристаллизуется в структуру типа цинковой обманки с одновременным изменением координационного числа с 6 до 4. Такое изменение типа связи и структуры связано с тем, что в ряду F—Cl—Br—I поляризуемость растет от F к I.

Из-за отсутствия у ионной связи направленности ионные кристаллы, как правило, имеют плотноупакованные решетки.⁶ При этом надо иметь в виду, что конкретная структура ионных кристаллов зависит от отношения ионных радиусов аниона и катиона. Для ионных кристаллов существуют пределы устойчивости структур с различными координационными числами. Условие устойчивости такой структуры заключается в том, чтобы каждый ион соприкасался только с ионами противоположного знака. Это обеспечивает компенсацию сил притяжения и отталкивания в ионном кристалле (отталкивание катионов (анионов) должно компенсироваться притяжением катионов (анионов) к анионам (катионам)). Если это условие не выполняется, то структура получается неустойчивой. Значения пределов устойчивости (см. табл. 2.1) определяются из геометрических соотношений радиусов катионов и анионов [17]. Таким образом, образуется одна из тех структур, в которых каждый катион контактирует с геометрически максимально возможным числом анионов с учетом пределов устойчивости данной структуры. Это соответствует условию минимальности полной энергии системы.⁷

Как правило, в ионных кристаллах крупные анионы образуют плотные упаковки, в пустотах которых размещаются катионы (структуры ти-

⁵Обычно поляризуемость анионов заметно больше, чем поляризуемость катионов.

⁶Плотнупакованные решетки — такие решетки, для которых коэффициент упаковки, определяемый как отношение полезного объема, занимаемого атомами, к полному объему элементарной ячейки, оказывается близким к предельному значению 0.74 для трехмерной решетки.

⁷Роль химического фактора оказывается в данном случае незначительной. Таким образом, химически подобные соединения могут иметь совершенно различную кристаллическую структуру, а химически разные соединения — одинаковую.

Таблица 2.1. Пределы устойчивости структур с различными координационными числами [18].

Коорд.число	r_+/r_-	Типичная структура
3	0.15–0.22	BF ₃ (молекула)
4	0.22–0.41	ZnS
6	0.41–0.73	NaCl
8	0.73–1	CsCl

па каменной соли NaCl, флюорита CaF₂ и др.). В итоге, ионные кристаллы обычно имеют плотноупакованные решетки с координационными числами $Z_k = 8, 6, 4$ и 3 , однако координация $Z_k = 12$ исключена, так как такая решетка оказывается неустойчивой.

2.2.3. Металлическая связь

Ранее уже говорилось, что при образовании химической связи происходит расщепление дискретных энергетических уровней изолированных атомов на уровни, которые группируются в зоны, между которыми могут быть энергетические зазоры. Возможны следующие ситуации:

- 1) верхние разрешенные зоны перекрываются;
- 2) между верхними разрешенными зонами возникает энергетический зазор E_g .

Само существование E_g и его величина полностью определяется индивидуальной электронной структурой взаимодействующих атомов.

С точки зрения заполнения энергетических зон электронами при низких температурах возможны три случая:

- 1) энергетические зоны перекрываются, верхняя зона не полностью заполнена;
- 2) перекрытие зон отсутствует, верхняя зона не полностью заполнена;
- 3) верхняя заполненная зона отделена от следующей (пустой) зоны запрещенной зоной E_g .

В первом и втором случаях реализуются металлическая связь, в третьем случае — ковалентная связь.

В металлах принадлежность валентных электронов к определенному атому установить не удастся, так как валентные электроны с небольшой энергией ионизации свободно перемещаются по доступным орбиталям всех соседних атомов, обеспечивая связь между ними, то есть валентные электроны являются коллективизированными и делокализованными. Они равномерно распределены по всему объему металлического кристал-

ла. Распределение электронной плотности заряда вокруг ионов, которые образовались в результате «отдачи» атомами в коллективное пользование своих валентных электронов, является приблизительно сферически-симметричным. Это позволяет представить структуру такого кристалла в виде ионного остова, погруженного в «газ» из коллективизированных валентных электронов, взаимодействие которых с ионами компенсирует силы отталкивания между одноименно заряженными ионами. Таким образом, металлическая связь носит ненаправленный и ненасыщенный характер.

Энергия коллективизированных валентных электронов в металлах квантована, но с очень малой разностью соседних энергетических уровней. Валентные электроны стремятся занять самые низкие энергетические уровни. В металлах существует корреляция между величиной энергии связи и энергией Ферми: $E_{\text{связ}} = A^*F$, где A^* — коэффициент, зависящий от природы металла (обычно $A^* < 1$ [15]).

Для абсолютного большинства металлов следствием отсутствия направленности связи и отсутствия ограничений, обусловленных наличием пределов устойчивости кристаллической структуры, является высокая симметрия и большая компактность кристаллических структур. Металлы кристаллизуются в структуры с плотнейшей упаковкой атомов, что обеспечивает минимальность полной энергии кристалла. Из кристаллографии известно, что наиболее плотную упаковку имеют гексагональная ($Z_k = 12$) и кубическая гранецентрированная ($Z_k = 12$) решетки, а кубическая объемноцентрированная решетка ($Z_k = 8$) имеет чуть меньшую плотность. Большинство металлов и большое число металлических соединений кристаллизуются с образованием одной из этих решеток.

2.2.4. Молекулярная связь (силы Ван-дер-Ваальса)

Этот тип связи реализуется в чистом виде в кристаллах инертных газов. Известно, что внешняя оболочка инертных газов полностью заполнена и поэтому весьма устойчива. Устойчивость внешней оболочки из 8 электронов проявляется в том, что взаимодействие атомов инертных газов с одноименными или другими атомами чрезвычайно слабо (слабая химическая активность). Однако тот факт, что их все-таки можно превратить в жидкость или твердое тело, доказывает наличие некоторых сил притяжения между атомами; в то же время исключительно низкие температуры, необходимые для их конденсации, показывают, что эти силы чрезвычайно малы. Силы, проявляющиеся у атомов инертных газов при низких температурах, называются *силами Ван-дер-Ваальса*. Энер-

гия притяжения, обусловленная этими силами, мала. Так, например, для криптона энергия связи ~ 11.2 кДж/моль (0.116 эВ/атом) или в температурных единицах ~ 100 К, что по порядку величины равно температуре плавления кристаллов инертных газов [16].

Рассмотрим причину возникновения этих сил. Атомы с заполненными валентными оболочками имеют сферическое распределение электронного заряда и не обладают постоянным электрическим моментом. Происхождение сил Ван-дер-Ваальса обусловлено наличием у таких атомов мгновенных индуцированных дипольных моментов. Если бы среднее положение ядра атома всегда совпадало с центром сферического электронного облака, окружающего ядро, то ван-дер-ваальсово взаимодействие между атомами равнялось бы нулю, а твердое тело не могло бы образоваться. Однако электроны в атоме постоянно движутся относительно ядер, даже находясь в наини́зшем энергетическом состоянии. В результате этого движения мгновенное положение центра электронного облака может не совпадать в точности с положением ядра атома. В эти моменты у атома появляется отличный от нуля электрический дипольный момент. Этот мгновенный дипольный момент создает в центре второго атома электрическое поле, которое в свою очередь наводит мгновенный дипольный момент у этого второго атома. Эти два дипольных момента взаимодействуют друг с другом, приводя к ван-дер-ваальсову взаимодействию (диполи ориентируются друг к другу противоположно заряженными концами, в результате чего происходит их электростатическое взаимодействие). Энергия этого взаимодействия выражается формулой $U = -c/r^6$, где c — некоторая эмпирическая константа, характеризующая силы взаимного притяжения.

На малых расстояниях между атомами с заполненными электронными оболочками проявляется действие сил отталкивания, обусловленных перекрытием их электронных облаков и электростатическим отталкиванием их ядер. Силы отталкивания на расстояниях в интервале от $0.5 \text{ \AA}$ до $5 \text{ \AA}$ становятся больше сил взаимного притяжения. Потенциал сил отталкивания описывается эмпирическим выражением $U = a/r^{12}$, где a — некоторая эмпирическая константа, характеризующая силы отталкивания.⁸ Энергия отталкивания в этом случае обусловлена главным образом действием принципа запрета Паули. Действительно, перекрытие запол-

⁸Эти силы отталкивания аналогичны силам отталкивания, обсуждаемым в разделе, посвященном ионной связи. Для описания изменения сил отталкивания с расстоянием широко используется, помимо приведенной формулы, и формула $U = \lambda \exp(-r/\rho)$, где ρ — размер области взаимодействия, λ — эмпирическая константа, характеризующая силу взаимодействия.

ненных электронных оболочек — это добавление электронов в уже заполненные электронные состояния, что противоречит принципу Паули. Реально перекрытие возможно только при переходе части электронов в более высокие незаполненные состояния. Это означает увеличение энергии системы, а это эквивалентно взаимному отталкиванию.

Таким образом, полную энергию межмолекулярного взаимодействия в кристаллах инертных газов можно представить в виде

$$U = a/r^{12} - c/r^6. \quad (2.6)$$

Следует заметить, что действие этих сил проявляется всегда, безотносительно и независимо от других сил. Для ван-дер-ваальсовых сил свойственно отсутствие направленности (в силу сферической симметрии распределения электронной плотности) и насыщаемости. Следствием этого является то, что инертные газы кристаллизуются подобно металлам в структуры с плотнейшей упаковкой (обычно кубическая гранецентрированная решетка с $Z_K = 12$, иногда гексагональная $Z_K = 12$).

2.2.5. Водородная связь

Поскольку нейтральный атом водорода имеет только один электрон, то, естественно, считать, что он может обладать только одной химической связью. Однако известно, что при некоторых условиях атом водорода может быть связан значительными силами притяжения одновременно с двумя атомами, образуя тем самым так называемую водородную связь между ними. Энергия такой связи порядка 0.1 эВ. Водородная связь имеет электростатическую природу. Она возникает при участии атома водорода лишь между наиболее электроотрицательными атомами, в частности между атомами F, O и N. Рассмотрим образование этой связи. Атом водорода может образовать полярную ковалентную связь с одним из соседних атомов, обладающих большой электроотрицательностью (см. ниже). При этом атом водорода приобретает положительный заряд и способен образовать электростатическую связь с другим соседним атомом, обладающим неподеленной электронной парой. В предельном случае при формировании этой связи атом водорода теряет свой единственный электрон, отдавая его одному из двух соседних атомов, и превращается в протон, который и осуществляет связь между атомами. Малые размеры протона не позволяют ему иметь ближайшими соседями более чем два атома. Таким образом, водородная связь осуществляется только между двумя атомами.

Водородная связь является основной формой взаимодействия (наряду с электростатическим притяжением дипольных моментов) между моле-

кулами H_2O при образовании кристаллов льда; она определяет размеры и геометрическую структуру белковых молекул; играет важную роль в молекулярной генетике (спаривание двух спиралей молекулы ДНК); существенна для объяснения свойств некоторых сегнетоэлектрических кристаллов (например, KDP).

2.3. Основные свойства ковалентной связи

Ковалентная связь осуществляется обобществленными валентными электронами, находящимися на общей для двух соседних атомов *связывающей орбитали* (см. разд. 2.2.1). При этом в случае элементарных веществ каждый из атомов «отдает» на связь одинаковое число валентных электронов и достраивает свою валентную оболочку до полностью заполненной за счет связывающих электронов ближайших соседних атомов в решетке. С этим свойством *насыщаемости* ковалентной связи мы познакомились на примере молекулы водорода. Его следствием является правило, установленное Юм-Розери. Согласно ему для кристаллов элементарных веществ, в которых реализуется преимущественно ковалентная связь, выполняется следующее соотношение между координационным числом Z_k и номером группы N , где расположен данный элемент: $Z_k = 8 - N$. Структура элементарных полупроводников (преимущественная связь — ковалентная) задается этим простым эмпирическим правилом и направлением ковалентной связи.

Направление ковалентной связи определяется распределением электронной плотности в кристалле, которое можно установить с помощью, например, рентгеновских данных. Они свидетельствуют, что в кристаллах с ковалентной связью электронная плотность валентных электронов существенно неравномерно распределена в пространстве. В направлениях, являющихся кратчайшими для двух соседних атомов, электронная плотность выше, чем в других направлениях. Это означает, что валентные электроны как бы локализованы в пространстве и образуют «электронные мостики», то есть ковалентная связь имеет резко выраженный *направленный характер*. Характер распределения электронной плотности при образовании ковалентной связи зависит от конкретной электронной структуры взаимодействующих атомов.

Так как ковалентная химическая связь в отличие от других типов химической связи имеет локализованный характер, то она определяется не только энергией связи, но имеет и геометрические характеристики. Геометрическими характеристиками ковалентной связи являются ее длина и углы между связями в молекуле или кристалле. *Длиной ковалентной*

химической связи называется расстояние между ядрами атомов, объединенных ковалентной связью, в кристалле. Она зависит от размеров взаимодействующих атомов и степени перекрытия их электронных облаков. Длину связи и углы между связями определяют экспериментально с помощью методов молекулярной спектроскопии, дифракции рентгеновских лучей и другими методами.

Свойство направленности ковалентной связи рассмотрим на примере образования химической связи в алмазе.⁹ При этом будем руководствоваться правилами, составляющими содержание теории направленных валентностей.

1. Ковалентная единичная связь образуется при взаимодействии двух электронов с противоположными спинами, принадлежащих разным атомам.

2. Направление ковалентной связи должно отвечать направлению, в котором орбитали данного валентного электрона в максимальной степени перекрывается орбиталями другого спаренного с ним валентного электрона, принадлежащего соседнему атому.

Следует иметь в виду, что в ряде случаев форма орбиталей валентных электронов, участвующих в образовании связи, изменяется незначительно, а в ряде случаев происходит ее резкое изменение. В последнем случае возникают смешанные, так называемые *гибридные орбитали*.

В алмазе, состоящем из атомов углерода, как хорошо известно, реализуется ковалентная химическая связь. Электронная конфигурация валентной оболочки атома углерода в невозбужденном состоянии — $2s^2 2p^2$. На $1s$ -орбитали и на $2s$ -орбитали находятся по два спаренных электрона, спины которых антипараллельны. Валентными оказываются две $2p$ -орбитали, на которых располагается по одному электрону, способному принимать участие в образовании химической связи. Эти $2p$ -орбитали образуют между собой угол в 90° . Таким образом, углерод в своих соединениях должен быть двухвалентным и формировать связи, между которыми прямой угол.

Однако, как показывают экспериментальные данные, в большинстве своих соединений углерод четырехвалентен и все четыре связи углерода одинаково прочны и имеют одинаковую ориентацию относительно друг друга: угол между связями равен $109^\circ 28'$. Это обстоятельство объясняется гибридизацией орбиталей, происходящей в два этапа. Сначала атом углерода переходит из основного состояния в возбужденное, при котором один из электронов с заполненной $2s^2$ -орбитали переходит на

⁹Ковалентная связь в молекуле водорода не обладает свойством направленности из-за сферически симметричного распределения электронной плотности в s -состояниях.

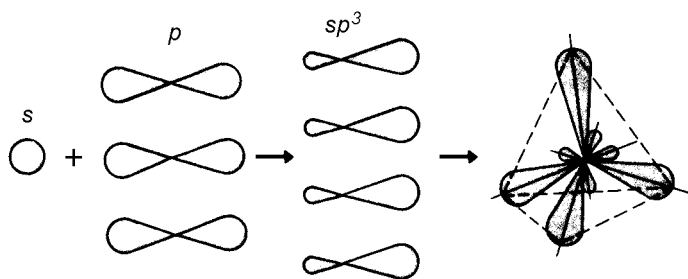


Рис. 2.10. Схема гибридизации s - и p -электронных облаков и пространственная ориентация sp^3 -гибридов.

пустую $2p$ -орбиталь. Далее происходит «перемешивание» четырех волновых функций и образование четырех новых одинаковых волновых функций, которые не являются ни s -, ни p -функциями. Это — гибридные sp^3 -функции. Они эквивалентны и имеют ориентацию, показанную на рис. 2.10. Таким образом, в результирующем состоянии атом углерода в алмазе имеет четыре неспаренных электрона. Направления максимальной плотности обменных облаков, как видно из рис. 2.12, сосредоточены вдоль пространственных диагоналей куба $\langle 111 \rangle$. Атом с четырьмя соседями образует правильный тетраэдр, а из совокупности тетраэдров строится «бесконечная ковалентная молекула».

Прочность химической связи зависит от степени перекрытия орбиталей валентных электронов: чем больше перекрытие, тем прочнее связь. Расчеты показывают, что перекрытие электронных оболочек соединяющихся атомов в случае гибридных sp^3 -орбиталей оказывается существенно больше, чем в случае негибридных s - и p -орбиталей. Хотя гибридным состояниям соответствует более высокая энергия электронов в атоме, чем негибридизованным, тем не менее полная энергия кристалла оказывается ниже в случае образования связи из sp^3 -гибридов, поэтому гибридизация и оказывается энергетически выгодной.

В образовании ковалентной связи могут принимать участие s -, p -, d и f -орбитали. Если происходит гибридизация при образовании связи, то в зависимости от того, сколько и каких орбит гибридизуется, различают sp -, sp^2 -, dsp^2 -, sp^3 - и d^2sp^3 -гибридные орбитали (рис. 2.11).

Таким образом, структура ковалентных кристаллов определяется правилом Юм-Розери, дающим число ближайших соседей, необходимых для полного насыщения химической связи, и направлением химической связи, которое может быть установлено из анализа волновых функций ва-

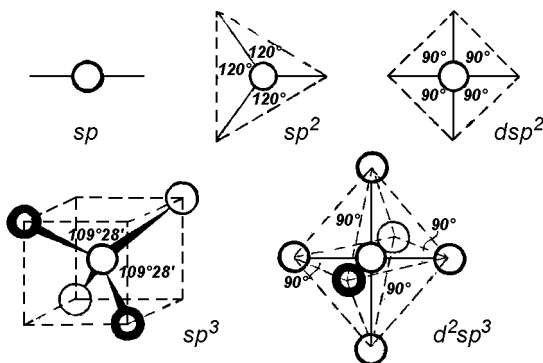


Рис. 2.11. Пространственная ориентация sp -, sp^2 -, dsp^2 -, sp^3 - и d^2sp^3 -гибридных орбиталей.

лентных электронов или экспериментально. Характерная величина энергии ковалентной связи составляет величину порядка 5–7 эВ. В полупроводниковых материалах прослеживаются следующие общие закономерности в изменении свойств при изменении энергии связи. С увеличением энергии связи между атомами уменьшается период кристаллической решетки, возрастают температура плавления и ширина запрещенной зоны.

Наиболее характерные свойства простых кристаллов, в которых реализуется преимущественно один из четырех типов химической связи, приведены в табл. 2.2.

Таблица 2.2. Свойства четырех групп твердых тел, отличающихся типами межатомных связей.

Ионная	Тип межатомной связи		
	Ковалентная	Металлическая	Ван-дер-Ваальсова
<i>Характеристика и энергия связи</i>			
Ненаправленная и ненасыщенная, сильная связь; 5–7 эВ на пару ионов.	Направленная и насыщенная, сильная связь; 5–7 эВ на атом.	Ненаправленная и ненасыщенная связь; около 3.5 эВ на атом.	Ненаправленная и ненасыщенная, слабая, короткодействующая связь; ≈ 0.1 эВ на атом.

Ионная	Ковалентная	Металлическая	Ван-дер-Ваальсова
<i>Структурные свойства</i>			
Крупные анионы образуют структуры с плотной упаковкой, в пустотах которых размещены катионы ($Z_k = 8, 6, 4$ и 3).	Структуры с неплотной упаковкой решет-ки (например, $Z_k = 4$) и низкой плотностью.	Компактные кристаллические структуры с плотнейшей упаковкой ($Z_k = 12$ и 8) и высокой плотностью.	Компактные кристаллические структуры с плотнейшей упаковкой ($Z_k = 12$) и высокой плотностью.
<i>Тепловые свойства</i>			
Довольно высокие температуры плавления. Низкий коэффициент расширения.	Высокие температуры плавления. Низкий коэффициент расширения.	Различные температуры плавления.	Низкие температуры плавления. Высокий коэффициент расширения.
<i>Электрические свойства</i>			
Изоляторы. Проводимость носит в основном ионный характер и увеличивается с ростом температуры.	Электронный тип проводимости (два типа носителей). Активационная зависимость проводимости от температуры.	Проводники. Основной тип проводимости — электронный. Проводимость с ростом температуры уменьшается.	Изоляторы.
<i>Оптические свойства</i>			
Прозрачны для электромагнитного излучения от низких частот до края поглощения. Обычно прозрачны в видимой области спектра.	Прозрачны для электромагнитного излучения от низких частот до края поглощения.	Непрозрачны для электромагнитных волн от самых низких частот вплоть до середины ультрафиолетовой области; хорошо отражают свет.	Прозрачны для электромагнитного излучения от низких частот до дальней ультрафиолетовой области.

Ионная	Ковалентная	Металлическая	Ван-дер-Ваальсова
<i>Зонная структура</i>			
Полностью заполненная верхняя валентная зона отделена от следующей пустой зоны проводимости широкой запрещенной зоной ($E_g > 2-3$ эВ).	Полностью заполненная верхняя валентная зона отделена от следующей пустой зоны проводимости запрещенной зоной. $E_g < 2-3$ эВ.	Зона проводимости частично заполнена.	Полностью заполненная верхняя валентная зона отделена от следующей пустой зоны проводимости очень широкой запрещенной зоной.

2.4. Элементарные полупроводники

Элементы, обладающие полупроводниковыми свойствами, то есть имеющие преимущественно ковалентный тип связи, образуют в таблице Д. И. Менделеева компактную группу. Снизу и слева с нею граничат металлы, сверху и справа — вещества, в твердом состоянии являющиеся изоляторами (табл. 1.1).

Рассмотрим элементарные полупроводники с точки зрения особенностей реализующихся в них межатомных связей. Учтем, что у всех этих полупроводников часть электронов валентной оболочки находится в p -орбиталях, причем их число меняется от 1 (бор) до 5 (йод). Покажем, что в зависимости от числа p -электронов, принимающих участие в образовании ковалентной связи, оказывается различной кристаллическая структура вещества и его полупроводниковые свойства.

2.4.1. Подгруппа IVA: углерод, кремний, германий, α -олово

Валентные оболочки свободных атомов этих элементов описываются формулой ns^2np^2 -электронов, где главное квантовое число $n = 2$ для углерода, 3 для кремния, 4 для германия и 5 для олова.

Химическая связь в элементарных полупроводниках из IVA подгруппы образуется с помощью sp^3 -гибридных σ -связей. Поэтому расположение атомов в пространстве таково, что каждый атом окружен четырьмя равноотстоящими ближайшими атомами, которые образуют тетраэдр. Такой тетраэдр может быть вписан в куб (рис. 2.12). Повторяющимся элементом, который образует бесконечную периодическую структуру в этих полупроводниках, является большой куб (рис. 2.13), содержащий 8 малых кубов, типа изображенных на рис. 2.12, из которых 4 центрированы,

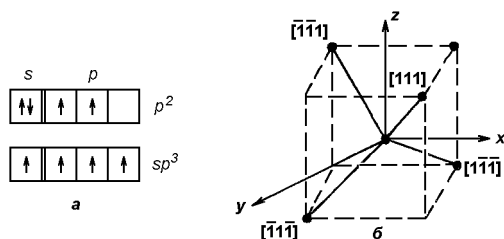


Рис. 2.12. *a* — Схема распределения электронов по квантовым ячейкам валентной оболочки C, Si, Ge и α -Sn. *б* — Пространственная ориентация sp^3 -гибридов в структуре типа алмаза.

а 4 не центрированы. В большом кубе каждый малый куб, содержащий центрированный тетраэдр, отделен от соседнего малого центрированного куба нецентрированным малым кубом. Полученную таким образом структуру называют структурой типа алмаза (координационное число в ней $Z_k = 4$).¹⁰ На одну элементарную ячейку решетки алмаза приходится 8 атомов. Таким образом, элементарные полупроводники IVA подгруппы имеют структуру типа алмаза, которую можно представить как образованную двумя гранецентрированными решетками, вставленными одна в другую и взаимно смещенными вдоль пространственной диагонали на одну четверть ее длины. В такой решетке каждый атом из одной подрешетки окружен четырьмя равноотстоящими ближайшими атомами другой подрешетки.

Кратчайшее расстояние между ядрами атомов дает длину связи d , а постоянная решетки $a = 4d\sqrt{3}/3$ характеризует расстояние между атомами, расположенными в вершинах кубов.

Коэффициент упаковки в структуре типа алмаза (плотность упаковки) мал — 0.34. Такая неплотная упаковка решетки, обусловленная направленностью связей, существенно сказывается на особенностях образования точечных дефектов, растворимости и диффузии примесей в алмазоподобных полупроводниках.

Углерод существует в нескольких модификациях, например, алмаз и графит. У алмаза (метастабильная при нормальных условиях модификация углерода) под внешней оболочкой находится всего два электрона, поэтому связь валентных электронов с ядром наименее экраниро-

¹⁰ Структуру алмаза нельзя получить путем набора только малых центрированных кубов, так как в этом случае каждый атом имел бы 8 ближайших соседей и решетка была бы объемноцентрированной.

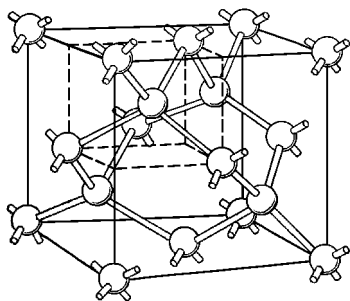


Рис. 2.13. Структура алмаза.

вана по сравнению с другими элементами этой подгруппы. Прочность связей валентных электронов с ядром проявляется как в механических, так и в электрических свойствах алмаза. Его ширина запрещенной зоны ($E_g = 5.48$ эВ) максимальна среди элементарных полупроводников. Несмотря на большую величину E_g и на высокое удельное сопротивление, алмаз включают в группу полупроводников, так как при наличии примесей или при освещении в нем обнаруживается проводимость с типичными для полупроводников высокими подвижностями электронов (см. табл. 2.3). Область прозрачности алмаза охватывает весь спектр видимого света, область собственного поглощения $h\nu > E_g$ находится в ультрафиолетовой области, а прочность — максимальна среди всех элементов периодической системы Д. И. Менделеева.

Существование второй модификаций углерода — *графита* (стабильной при нормальных условиях) — связано с иным, чем в алмазе, характером гибридизации волновых функций, участвующих в образовании химической связи. Графит существенно отличается от алмаза структурой и своими физическими свойствами. Он имеет слоистое строение и гексагональную структуру, показанную на рис. 2.14. Атомы углерода лежат в параллельных плоскостях, в каждой из которых они расположены в вершинах правильных шестиугольников. Расстояние между атомами в слое ($d_1 = 1.42$ Å) близко к длине связи в алмазе, что указывает на ковалентный характер связи в слое, а большое расстояние между слоями ($d_2 = 3.35$ Å) указывает на слабую связь между ними, по-видимому, обусловленную действием ван-дер-ваальсовых сил. Естественно предположить, что в графите валентные связи строятся не на sp^3 - (как в алмазе), а на sp^2 -гибридных и p -орбиталях. Действительно, эксперимент показывает, что три эквивалентные связи образуют плоскую структуру (слой) и расположены под углом близким к 120° (sp^2 -гибридные орби-

Таблица 2.3. Некоторые физические свойства элементарных полупроводников [19]. (В таблице везде, где не отмечено специально, приводимые значения относятся к $T = 290$ К.)

Элемент	Кристаллич. структура	E_g , эВ	Характерные значения подвижности, $\text{см}^2/\text{В} \cdot \text{с}$	
			μ_n	μ_p
С (алмаз)	Куб., O_h^b $a=3.566 \text{ \AA}$	5.48	1800	1400
Si	Куб., O_h^7 $a=5.43 \text{ \AA}$	1.11	3000	500
Ge	Куб., O_h^7 $a=5.66 \text{ \AA}$	0.664	3800	1820
α -Sn	Куб., O_h^7 $a=6.49 \text{ \AA}$	0.09	2500	2400
Р (черный)	Орторомб., D_{2h}^{18} $a=3.31$, $b=4.37$, $c=10.47 \text{ \AA}$	0.33 (0 К)	220	350
As (серый)	Ромбоэдр., D_{3d}^5 $a=3.76$, $c=10.76 \text{ \AA}$	0.175 (< 20 К)	40–550 (анизотр.)	50–1210 (анизотр.)
Sb	Ромбоэдр., D_{3d}^5 $a=4.308$, $c=11.274 \text{ \AA}$	0.1 (1.15 К)	110–274 (анизотр.)	180–363 (анизотр.)
α -S	Орторомб., D_{2h}^{24} $a=10.46$, $b=12.88$, $c=24.48 \text{ \AA}$	3.8	7.5	10
Se	Триг., D_3^4 $a=4.35$, $c=4.94 \text{ \AA}$	1.9	—	40
Te	Триг., D_3^4 $a=4.94$, $c=5.95 \text{ \AA}$	0.335	2380 ($\parallel c$) 1150 ($\perp c$)	1260 ($\parallel c$) 650 ($\perp c$)
α -B	Ромбоэдр., D_{3d}^5 $a=5.057 \text{ \AA}$, $\alpha=58^\circ$	1.97 ($E \parallel c$) 1.90 ($E \perp c$)	—	120
β -B	Ромбоэдр., D_{3d}^5 $a=10.14 \text{ \AA}$, $\alpha=65^\circ$	1.63 ($E \parallel c$) 1.61 ($E \perp c$)	1	300

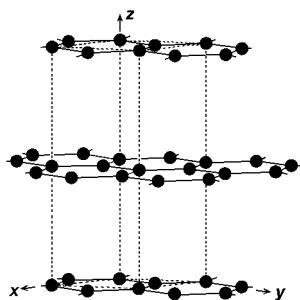


Рис. 2.14. Структура графита. Элементарная ячейка указана штриховыми линиями.

тали), а четвертая связь направлена перпендикулярно плоскости слоя и построена на p -орбиталях. Отсюда следует, что три sp^2 σ -связи являются ковалентными, то есть строго локализованными по соответствующим направлениям, а перекрытие между p -орбиталями соседних атомов приводит к образованию π -связывающих орбиталей, которые расположены сверху и снизу от плоскости. Высокая электропроводность графита в направлениях, параллельных плоскостям, и металлический блеск обусловлены лишь частичным заполнением p -орбиталей, из которых образуются π -связывающие молекулярные орбитали. Таким образом, особенности образования ковалентной химической связи в графите приводят к тому, что этот материал имеет слоистую структуру с резкой анизотропией механических и электронных свойств.

Кремний Si имеет 10 экранирующих электронов, которые ослабляют связь валентных электронов с ядром. О величине этого эффекта можно судить по уменьшению первого потенциала ионизации. Физические параметры Si приведены в табл. 2.3.

Германий Ge имеет 28 экранирующих электронов. Физические параметры приведены в табл. 2.3.

У олова между валентной оболочкой и ядром находится 46 электронов и их экранирующее действие столь велико, что ковалентная связь оказывается неустойчивой. Алмазной структурой с $E_g = 0.09$ эВ и типичными для этой группы подвижностями носителей заряда обладает лишь одна из модификаций — α -Sn (серое олово), устойчивая при температурах ниже -13°C ; обычное β -Sn, устойчивое при более высоких температурах, — металл.

Таким образом, видно, что по мере возрастания атомного номера элемента ковалентная составляющая связи закономерно ослабляется, уменьшается E_g и возрастает доля металлической составляющей связи. У ато-

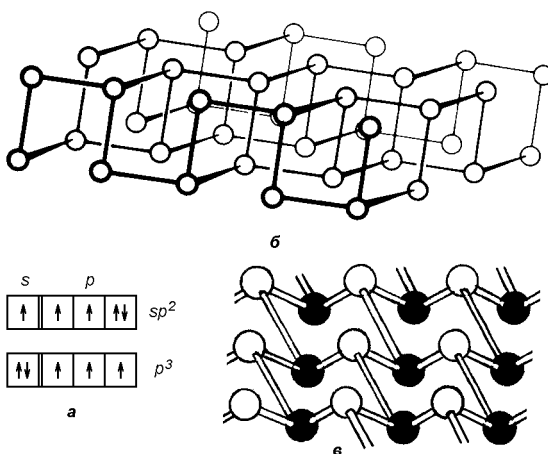


Рис. 2.15. *а* — Схема возможного распределения электронов по квантовым ячейкам валентной оболочки элементов подгруппы VA. *б* — Один слой кристаллической структуры черного фосфора. *в* — Кристаллической структура мышьяка, сурьмы и висмута.

ма Pb, следующего за Sn, химическая связь теряет направленность и становится металлической (для атома свинца характерно сильное влияние внутренних электронов на потенциал взаимодействия внешних электронов с ядром, что приводит к значительному расщеплению s - и p -состояний и, как следствие, s -электроны не участвуют в образовании химической связи — нет sp^3 -гибридизации).

2.4.2. Подгруппа VA: фосфор, мышьяк, сурьма и висмут

Структура валентной оболочки — ns^2np^3 с $n = 3$ для фосфора, 4 для мышьяка, 5 для сурьмы, 6 для висмута. Распределение электронов по ячейкам валентной оболочки показано на рис. 2.15, *а*: основное состояние характеризуется тремя электронами с неспаренными спинами (p^3), причем число их не будет меняться даже при переходе одного из s -электронов на p -орбиту. Таким образом, во всех указанных выше элементах этой подгруппы три связи любого атома направлены к трем другим соседним атомам, при этом образуются структуры с координационным числом $Z_k = 3$. Физические параметры элементов этой подгруппы приведены в табл. 2.3.

Фосфор. Фосфор имеет по крайней мере пять полиморфных модификаций (белый (две модификации), красный, коричневый и черный). Лишь одна из них — черный фосфор, который образуется при повышенной температуре ($\approx 200^\circ\text{C}$) и повышенном давлении, — обладает полупроводниковыми свойствами. Черный фосфор — наиболее устойчивая кристаллическая форма фосфора. Он кристаллизуется в орторомбической системе; его элементарная ячейка содержит восемь атомов. Кристалл имеет слоистую структуру, в которой каждый атом связан с тремя другими. Один слой этой структуры показан на рис. 2.15,б. Длина связи Р–Р равна $2.23 \text{ \AA}$. Два угла между направлениями валентных связей составляют 102° , а один — $96,5^\circ$. По-видимому, ковалентные связи образуются в результате частичного смешивания и перекрытия sp^2 -гибридных и p -орбиталей.

Мышьяк. Мышьяк также существует в трех разных модификациях (белый, серый и аморфный), из которых серая кристаллическая и аморфная модификации обладают полупроводниковыми свойствами. Термодинамически устойчивым при нормальных условиях является серый мышьяк, кристаллизующийся в ромбоэдрической решетке. Кристалл серого мышьяка имеет гофрированно-слоистую структуру (рис. 2.15,в). Ее можно представить состоящей из двухслойных пакетов, суперпозиция которых дает ромбоэдрическую структуру кристалла. Таким образом, каждый атом имеет три ближайших соседа в соседнем слое того же пакета на расстоянии $d_1 = 2.52 \text{ \AA}$, с которыми он связан ковалентными связями, и три более удаленных соседа из соседнего пакета на расстоянии $d_2 = 3.12 \text{ \AA}$. Углы между направлениями межатомных связей в одном пакете (96.65°) указывают на то, что ковалентные связи образуются в результате перекрытия простых p^3 -орбиталей. Соседние пакеты связаны между собой в основном слабыми силами Ван-дер-Ваальса, однако в химической связи между соседними пакетами присутствует значительная доля металлической составляющей. Смешанный характер связей накладывает отпечаток на степень совершенства структуры и на электрические свойства: они явно анизотропны, а подвижность носителей заряда оказывается существенно ниже, чем в элементах с трехмерной ковалентной структурой.

Сурьма, висмут. Сурьма и висмут имеют ту же кристаллическую структуру, что и мышьяк. Однако с увеличением атомного номера увеличивается доля металлической составляющей связи (в обычном виде сурьма и висмут — полуметаллы). Этот факт находит свое отражение и в кристаллической структуре: для сурьмы $d_1 = 2.91 \text{ \AA}$ и $d_2 = 3.36 \text{ \AA}$ и разность между ними ($d_2 - d_1$) оказывается существенно меньшей, чем

в случае мышьяка. Для Bi $d_1 = 3.07 \text{ \AA}$ и $d_2 = 3.53 \text{ \AA}$. Однако следует заметить, что несмотря на очевидно металлический характер Sb и Bi существует различие между их структурой с относительно низкой координацией и структурами с высокой координацией, типичными для настоящих металлов.

Таким образом, закономерности изменения E_g с увеличением порядкового номера элемента в этой группе полупроводников аналогичны описанным выше для IVA подгруппы. Кроме того, видно, что химическая связь, образующаяся в подгруппе VA, оказывается заметно слабее связи, построенной на sp^3 -гибридных орбиталях.

2.4.3. Подгруппа VIA: сера, селен, теллур

Структура валентной оболочки — ns^2np^4 ($n = 3$ для серы, 4 для селена, 5 для теллура). Элементы этой группы содержат два неспаренных p -электрона на внешней оболочке, поэтому две ковалентные связи от каждого атома могут быть направлены либо к одному атому (при этом будут образовываться двухатомные молекулы как в O_2), либо к двум различным атомам (образуются многоатомные молекулы в виде замкнутых колец или бесконечных протяженных цепочек). Угол между соседними ковалентными связями, направленными к двум различным атомам, не должен сильно отличаться от 90° . Отличие может быть связано либо с частичной sp -гибридизацией волновых функций электронов, либо со взаимодействием между вторыми соседями. В кристалле кольца или изогнутые цепочки связаны между собой силами Ван-дер-Ваальса. Можно ожидать, что связи здесь будут еще слабее, чем в кристаллах VA подгруппы. Физические параметры элементов подгруппы приведены в табл. 2.3.

Сера. Существует несколько различных полиморфных модификаций серы, однако полупроводниковыми свойствами обладает только наиболее стабильная при нормальных условиях орторомбическая модификация. Структура этой фазы строится из молекул S_8 , которые имеют кольцевую форму, показанную на рис. 2.16,б. Расстояние S–S составляет $d = 2.06 \text{ \AA}$, что соответствует ковалентному радиусу атома серы. Углы между ковалентными связями равны 108° . Отдельные компактные молекулы S_8 относительно слабо связаны между собой силами Ван-дер-Ваальса.

Селен. Для наиболее часто встречающейся полупроводниковой модификации селена (тригональной) характерно образование длинных спиральных цепочек, оси которых образуют оси c гексагональной структуры. Угол между связями в такой цепочке составляет 103° , поэтому можно

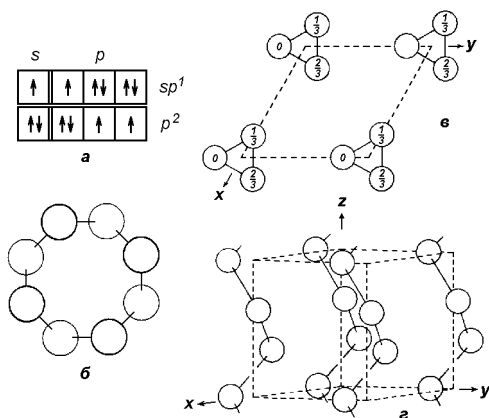
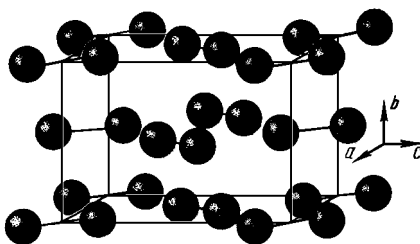


Рис. 2.16. *а* — Схема возможного распределения электронов по квантовым ячейкам валентной оболочки элементов подгруппы VIA. *б* — Молекула S₈ в кристаллической структуре орторомбической серы. *в*, *г* — Схема элементарной ячейки гексагональной структуры селена, спроектированная на плоскость, перпендикулярную оси *z*; клинографическая проекция той же самой ячейки.

говорить, что эти связи осуществляются не чистыми *p*-, а частично гибридованными *sp*-волновыми функциями. Расстояние Se–Se в цепочке равно $d_1 = 2.37 \text{ \AA}$, а между атомами в соседних цепочках — $d_2 = 3.49 \text{ \AA}$, что соответствует слабым ван-дер-ваальсовым силам, удерживающим цепочки вместе. Заметим, что расстояние d_2 значительно меньше, чем соответствующий радиус Ван-дер-Ваальса для атома селена, поэтому, вероятно, что в связи между цепочками присутствует металлическая составляющая, и следует ожидать усиления металлических свойств элемента. Каждый атом в структуре селена оказывается симметрично окруженным шестью соседями в базисной плоскости, а цепочки образуют структуру подобную прутикам в связке. Такая сложная природа связей и такая кристаллическая структура селена приводят к тому, что для селена характерна сильная анизотропия электропроводности и подвижности (вдоль и перпендикулярно оси *c*). Общее же несовершенство структуры создает в кристаллах большие концентрации активных дефектов, поэтому Se всегда имеет проводимость *p*-типа, которую не удастся изменить даже введением больших концентраций примесей.

Теллур. Теллур имеет такую же структуру, что и селен (расстояние Te–Te в ней оставляет $2.835 \text{ \AA}$, а угол между связями в цепочке — 103.2°). Однако она более совершенная, чем у селена, вследствие чего

Рис. 2.17. Структура йода.



подвижности носителей в теллуре значительно выше ($\mu_n = 1700 \text{ см}^2/\text{В}\cdot\text{с}$, $\mu_p = 1200 \text{ см}^2/\text{В}\cdot\text{с}$). Свойства Те анизотропны, хотя и не в такой степени, как у Se. Тип проводимости Те также всегда дырочный.

Из табл. 2.3 видно, что корреляция между величиной E_g и порядковым номером элементов VIA подгруппы такая же, как и у элементов IVA и VA подгрупп.

2.4.4. Подгруппа VIIA: йод

Из элементов VIIA подгруппы к полупроводникам относится только твердый йод. Структура йода (рис. 2.17) состоит из молекул I_2 , в которых пары атомов прочно связаны между собой ковалентными связями. Силы Ван-дер-Ваальса, удерживающие молекулы в кристалле, очень слабы. При слабом нагревании кристалл разрушается, причем молекулы сохраняются и в жидкой и газообразной фазе. Лишь при нагревании до температуры 1200°C при давлении в 0.1 МПа около половины молекул диссоциирует на отдельные атомы. Расстояние между атомами I—I в молекуле составляет $d = 2.7 \text{ \AA}$, а ближайшее расстояние между различными молекулами — $d = 3.54 \text{ \AA}$.

2.4.5. Подгруппа IIIA: бор

Бор — единственный элементарный полупроводник, расположенный левее IVA подгруппы в периодической системе. Структура его валентной оболочки — $2s^2p^1$. В большинстве соединений он трехвалентен, поэтому распределение электронов по валентным орбиталям в нем можно представить в виде $2s^12p^2$. Однако при образовании простых sp^2 -гибридных связей в боре валентная оболочка атомов бора не заполняется полностью, тем не менее бор, как полупроводник, должен иметь полностью заполненную валентную зону. По-видимому, заполнение происходит при участии в гибридизации, помимо $2s$ -электронов, еще и $1s$ -электронов.

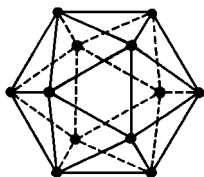


Рис. 2.18. Икосаэдрическая структура бора.

В этом случае полное число электронов, участвующих в формировании химических связей у элементарного бора, будет равно 5. Этому предположению соответствует значительно более сложная по сравнению с другими элементарными полупроводниками кристаллическая структура бора, основной структурной единицей которой являются икосаэдрические группы B_{12} (правильные двадцатигранники, в которых каждый атом имеет пять соседей, см. рис. 2.18). Существует несколько полиморфных модификаций бора, среди которых полупроводниковыми свойствами обладают только ромбоэдрические модификации бора (метастабильная α -форма и стабильная β -форма). Элементарная ячейка α -В состоит из 12 атомов, а ячейка β -В — из 105 атомов. Бор химически инертен и обладает твердостью, близкой к твердости алмаза. Это обусловлено образованием прочных ковалентных связей (расстояние В–В равно 1.71 Å) и трудностью их разрыва, что определяет и высокую температуру кипения данного вещества.

2.5. Атомные и ионные радиусы

Одной из важнейших характеристик химических элементов, участвующих в образовании химической связи, является размер атома (иона): с его увеличением прочность межатомных связей снижается. Размер атома (иона) принято определять величиной его радиуса или диаметра. Так как атом (ион) не имеет четких границ, то понятие «атомный (ионный) радиус» подразумевает, что 90–98 % электронной плотности атома (иона) заключено в сфере этого радиуса. Знание величин атомных (ионных) радиусов позволяет оценивать межъядерные расстояния в кристаллах (то есть структуру этих кристаллов), так как для многих задач кратчайшие расстояния между ядрами атомов (ионов) можно считать суммой их атомных (ионных) радиусов, хотя такая аддитивность приближенна и выполняется не во всех случаях.

Под *атомным радиусом* химического элемента (об ионном радиусе см. ниже), участвующего в образовании химической связи, в общем случае договорились понимать половину равновесного межъядерного расстояния между ближайшими атомами в кристаллической решетке элемента. Это понятие, весьма простое, если рассматривать атомы (ионы) в виде жестких шаров, фактически оказывается сложным и часто не однозначным. Атомный (ионный) радиус химического элемента не является неизменной величиной, а изменяется в зависимости от ряда факторов, важнейшими из которых являются тип химической связи

и координационное число.

Если один и тот же атом (ион) в различных кристаллах образует разные типы химической связи, то у него будет несколько радиусов — ковалентный в кристалле с ковалентной связью; ионный в кристалле с ионной связью; металлический в металле; ван-дер-ваальсов в молекулярном кристалле. Влияние типа химической связи можно проследить на следующем примере. В алмазе все четыре химические связи являются ковалентными и образованы sp^3 -гибридами, поэтому все четыре соседа данного атома находятся на одном и том же расстоянии от него ($d = 1.54 \text{ \AA}$) и ковалентный радиус углерода в алмазе будет равен $0.77 \text{ \AA}$. В кристалле мышьяка расстояние между атомами, связанными ковалентными связями ($d_1 = 2.52 \text{ \AA}$), значительно меньше, чем между атомами, связанными силами Ван-дер-Ваальса ($d_2 = 3.12 \text{ \AA}$), поэтому у As будет ковалентный радиус, равный $1.26 \text{ \AA}$, и ван-дер-ваальсов, равный $1.56 \text{ \AA}$.

Очень резко изменяется атомный (ионный) радиус и при изменении координационного числа (это можно наблюдать при полиморфных превращениях элементов). Чем меньше координационное число, тем меньше степень заполнения пространства атомами (ионами) и меньше межъядерные расстояния. Увеличение же координационного числа всегда сопровождается увеличением межъядерных расстояний.

Из сказанного следует, что атомные (ионные) радиусы разных элементов, участвующих в образовании химической связи, можно сравнивать только тогда, когда они образуют кристаллы, в которых реализуется один и тот же тип химической связи, и у этих элементов в образуемых кристаллах одинаковые координационные числа.

Рассмотрим основные особенности атомных и ионных радиусов более подробно.

Под *ковалентными радиусами элементов* принято понимать половину равновесного межъядерного расстояния между ближайшими атомами, соединенными ковалентной связью.

Особенностью ковалентных радиусов является их постоянство в разных «ковалентных структурах» с одинаковым координационным числом Z_k . Кроме того, ковалентные радиусы, как правило, аддитивны: связаны друг с другом, то есть расстояние А–В равно полусумме расстояний А–А и В–В при наличии ковалентных связей и одинаковых координационных чисел во всех трех структурах.

Различают нормальный, тетраэдрический, октаэдрический, квадратичный и линейный ковалентные радиусы.

Нормальный ковалентный радиус атома отвечает случаю, когда атом образует столько ковалентных связей, сколько соответствует его месту в периодической таблице: для углерода — 2, для азота — 3 и т. д. При этом получаются разные значения нормальных радиусов в зависимости от кратности (порядка) связи (единичная связь, двойная, тройная). Если связь образуется при перекрытии гибридных электронных облаков, то говорят о тетраэдрических ($Z_k = 4$, sp^3 -гибридные орбитали), октаэдрических ($Z_k = 6$, d^2sp^3 -гибридные орбитали), квадратичных ($Z_k = 4$, dsp^2 -гибридные орбитали), линейных ($Z_k = 2$, sp -гибридные орбитали) ковалентных радиусах.

О ковалентных радиусах полезно знать следующее (значения величин ковалентных радиусов для ряда элементов приведены в [15, 20]).

1. Ковалентные радиусы, в отличие от ионных, нельзя интерпретировать как радиусы атомов, имеющих сферическую форму. Ковалентные радиусы применяются только для вычисления межъядерных расстояний между атомами, объединенными ковалентными связями, и ничего не говорят о расстояниях между атомами того же типа не связанными ковалентно.

2. Величина ковалентного радиуса определяется кратностью ковалентной связи. Тройная связь короче двойной, которая в свою очередь короче единичной, поэтому ковалентный радиус тройной связи меньше, чем ковалентный радиус двойной связи, который меньше

единичного. Следует иметь в виду, что порядок кратности связи не обязательно должен быть целым числом. Он может быть и дробным, если связь носит резонансный характер (молекула бензола, соединение Mg_2Sn , см. ниже). В этом случае ковалентный радиус имеет промежуточное значение между значениями, соответствующими целым порядкам кратности связи.

3. Если связь носит смешанный ковалентно-ионный характер, но с высокой степенью ковалентной составляющей связи, то можно вводить понятие ковалентного радиуса, но нельзя пренебрегать влиянием ионной составляющей связи на его величину. В некоторых случаях это влияние может приводить к значительному уменьшению ковалентного радиуса, иногда до $0.1 \text{ \AA}$. К сожалению, попытки предсказать величину этого эффекта в различных случаях пока не увенчались успехом.

4. Величина ковалентного радиуса зависит от типа гибридных орбиталей, которые принимают участие в образовании ковалентной связи.

Ионные радиусы, естественно, не могут быть определены как полусумма расстояний между ядрами ближайших ионов, так как, как правило, размеры катионов и анионов резко различаются. Кроме того, симметрия ионов может несколько отличаться от сферической. Тем не менее для реальных ионных кристаллов под *ионным радиусом* принято понимать радиус шара, которым аппроксимируется ион.

Ионные радиусы используются для приближенных оценок межъядерных расстояний в ионных кристаллах. При этом считают, что расстояния между ближайшими катионом и анионом равно сумме их ионных радиусов. Типичная погрешность определения межъядерных расстояний через ионные радиусы в таких кристаллах составляет величину $\approx 0.01 \text{ \AA}$.

Существует несколько систем ионных радиусов, отличающихся значениями ионных радиусов индивидуальных ионов, но приводящих к примерно одинаковым межъядерным расстояниям. Первая работа по определению ионных радиусов была проведена В. М. Гольдшмитом в 20-х годах XX века. В ней автор использовал, с одной стороны, межъядерные расстояния в ионных кристаллах, измеренные методами рентгеновского структурного анализа, а, с другой стороны, — значения ионных радиусов F^- и O^{2-} , определенные методом рефрактометрии. Большинство других систем также опирается на определенные дифракционными методами межъядерные расстояния в кристаллах и на некоторые «реперные» значения ионного радиуса конкретного иона. В наиболее широко известной системе Полинга этим реперным значением является ионный радиус пероксид-иона O_2^{2-} , равный $1.40 \text{ \AA}$. Эта величина для O_2^{2-} хорошо согласуется с теоретическими расчетами. В системе Г. Б. Бокия и Н. В. Белова, считающейся одной из наиболее надежных, ионный радиус O^{2-} принимается равным $1.36 \text{ \AA}$.

В 70–80-х годах были сделаны попытки прямого определения радиусов ионов путем измерения электронной плотности методами рентгеновского структурного анализа при условии, что минимум электронной плотности на линии, соединяющей ядра, принимается за границу ионов. Оказалось, что этот прямой метод приводит к завышенным значениям ионных радиусов катионов и к заниженным значениям ионных радиусов анионов. Кроме того, оказалось, что значения ионных радиусов, определенные прямым способом, нельзя переносить от одного соединения к другому, а отклонения от аддитивности слишком велики. Поэтому такие ионные радиусы, не используются для предсказания межъядерных расстояний.

О ионных радиусов полезно знать следующее (в таблицах, приведенных ниже, даны величины ионных радиусов по Бокию и Белову).

1. Ионный радиус для ионов одного и того же элемента меняется в зависимости от его заряда, а для одного и того же иона зависит от координационного числа. В зависимости от координационного числа различают тетраэдрический и октаэдрический ионные радиусы.

2. Внутри одного вертикального ряда, точнее внутри одной группы, периодической

системы радиусы ионов с одинаковым зарядом возрастают с увеличением атомного номера элемента, поскольку растет число занятых электронами оболочек, а значит и размер иона.

Ион	Li ⁺	Na ⁺	K ⁺	Rb ⁺	Cs ⁺
<i>Z</i>	3	11	19	37	55
Радиус, Å	0.70	1.00	1.33	1.52	1.70

3. Для положительно заряженных ионов атомов из одного периода ионные радиусы быстро уменьшаются с увеличением заряда. Быстрое уменьшение объясняется действием в одну сторону двух основных факторов: сильное притяжение «своих» электронов катионом, заряд которого увеличивается с увеличением атомного номера; увеличение силы взаимодействия между катионом и окружающими его анионами с увеличением заряда катиона.

Ион	Au ⁺	Hg ²⁺	Tl ³⁺	Pb ⁴⁺
Радиус, Å	1.37	1.12	1.05	0.76

4. Для отрицательно заряженных ионов атомов из одного периода ионные радиусы увеличиваются с увеличением отрицательного заряда. Два фактора, о которых шла речь в предыдущем пункте, в этом случае действуют в противоположные стороны, причем преобладает первый фактор (увеличение отрицательного заряда аниона сопровождается возрастанием его ионного радиуса), поэтому увеличение ионных радиусов с ростом отрицательного заряда происходит существенно медленнее, чем уменьшение в предыдущем случае.

Ион	F ⁻	O ²⁻	Cl ⁻	S ²⁻
Радиус, Å	1.33	1.36	1.81	1.82

5. Для одного и того же элемента, то есть при одинаковой исходной электронной конфигурации, радиус катиона меньше, чем аниона. Это обусловлено уменьшением притяжения внешних «добавочных» электронов к ядру аниона и увеличением эффекта экранирования за счет внутренних электронов (у катиона недостаток электронов, а у аниона избыток).

Ион	Pb ⁴⁻	Pb	Pb ²⁺	Pb ⁴⁺
Радиус, Å	2.15	1.74	1.18	0.84

6. Размеры ионов с одинаковым зарядом следуют периодичности таблицы Менделеева. Однако величина ионного радиуса не пропорциональна заряду ядра *Z*, что обусловлено сильным притяжением электронов ядром. Кроме того, исключение из периодической зависимости представляют лантаноиды и актиноиды, в рядах которых радиусы атомов и ионов с одинаковым зарядом не растут, а уменьшаются с ростом атомного номера (так называемые лантаноидное сжатие и актиноидное сжатие).¹¹

¹¹Лантаноидное сжатие и актиноидное сжатие обусловлены тем, что у лантаноидов и актиноидов добавляющиеся при увеличении атомного номера электроны заполняют *внутренние d- и f-оболочки* с главным квантовым числом меньшим, чем главное квантовое число данного периода. При этом согласно квантовомеханическим расчетам в *d-* и особенно в *f-* состояниях электрон находится гораздо ближе к ядру, чем в *s-* и *p-* состояниях данного периода с большим квантовым числом, поэтому *d-* и *f-* электроны размещаются во внутренних областях атома, хотя заполнение этих состояний электронами (речь идет об электронных уровнях в энергетическом пространстве) происходит по другому.

Металлические радиусы считаются равными половине кратчайшего расстояния между ядрами атомов в кристаллизующейся структуре элемента-металла. Они зависят от координационного числа. Если принять металлический радиус какого-либо элемента при $Z_k = 12$ за единицу, то при $Z_k = 8, 6$ и 4 металлические радиусы того же элемента будут соответственно равны $0.98; 0.96; 0.88$. Металлические радиусы обладают свойством аддитивности. Знание их величин позволяет приблизительно предсказывать параметры кристаллических решеток интерметаллических соединений.

Для атомных радиусов металлов характерны следующие особенности (данные о величинах атомных радиусов металлов можно найти в [15, 20]).

1. Металлические атомные радиусы переходных металлов, как правило, меньше, чем металлические атомные радиусы непереходных металлов, что отражает большую прочность связи в металлах переходных элементов. Эта особенность обусловлена тем, что металлы переходных групп и ближайшие к ним в периодической системе металлы имеют электронные d -оболочки, а электроны в d -состояниях могут принимать участие в образовании химической связи. Упрочнение связи может быть связано отчасти с появлением ковалентной составляющей связи и отчасти с ван-дер-ваальсовым взаимодействием ионных остовов. В кристаллах железа и вольфрама, например, электроны в d -состояниях вносят существенный вклад в энергию связи.

2. В пределах одной вертикальной группы по мере продвижения сверху вниз атомные радиусы металлов возрастают, что обусловлено последовательным увеличением числа электронов (растет число занятых электронами оболочек).

Элемент	Li	Na	K	Rb	Cs
Радиус, Å	1.52	1.85	2.31	2.46	2.62

3. В пределах одного периода, точнее начиная с щелочного металла до середины группы переходных металлов, в направлении слева направо атомные металлические радиусы уменьшаются. В той же последовательности возрастает электрический заряд атомного ядра и происходит увеличение числа электронов, находящихся на валентной оболочке. При возрастании числа связывающих электронов, приходящихся на один атом, металлическая связь упрочняется, и вместе с тем из-за увеличения заряда ядра усиливается притяжение остовных (внутренних) электронов ядром, поэтому величина металлического атомного радиуса уменьшается.

Элемент	K	Ca	Sc	Ti	V	Cr
Радиус, Å	2.31	1.96	1.60	1.46	1.31	1.25

4. Переходные металлы VII и VIII групп из одного периода в первом приближении имеют почти одинаковые металлические радиусы. По-видимому, когда речь идет об элементах, имеющих 5 и большее число d -электронов, увеличение заряда ядра и связанные с этим эффекты притяжения остовных электронов, ведущие к уменьшению атомного металлического радиуса, компенсируются эффектами, обусловленными увеличивающимся в атоме (ионе) числом электронов, не участвующих в образовании металлической связи, и ведущими к увеличению металлического радиуса (растет число занятых электронами состояний).

Элемент	Fe	Co	Ni
Радиус, Å	1.23	1.25	1.24

5. Увеличение радиусов (см. п. 2) у переходных элементов, которое имеет место при переходе от четвертого к пятому периоду, не наблюдается у переходных элементов при

переходе от пятого к шестому периоду; металлические атомные радиусы соответствующих (сравнение идет по вертикали) элементов в этих двух последних периодах почти одинаковы. По-видимому, это связано с тем, что у элементов, расположенных между ними, дстраивается сравнительно глубоко расположенная f -оболочка, поэтому увеличение заряда ядра и связанные с этим эффекты притяжения оказываются более существенными, чем эффекты, связанные с увеличивающимся числом электронов (лантаноидное сжатие).

Элемент из 4 периода	Cr	Fe	Ni	Cu
Радиус, Å	1.28	1.27	1.24	1.28
Элемент из 5 периода	Mo	Ru	Pd	Ag
Радиус, Å	1.40	1.33	1.37	1.44
Элемент из 6 периода	W	Os	Pt	Au
Радиус, Å	1.41	1.35	1.38	1.44

6. Обычно металлические радиусы много больше, чем ионные радиусы, однако не столь значительно отличаются от ковалентных радиусов тех же элементов, хотя и все без исключения больше ковалентных. Большая разница в величинах металлических атомных и ионных радиусов одних и тех же элементов объясняется тем, что связь, обязанная своим происхождением почти свободным электронам проводимости, не является сильной (отсюда наблюдаемые относительно большие межатомные расстояния в решетке металлов). Существенно меньшую разницу в величинах металлических и ковалентных радиусов одних и тех же элементов можно объяснить, если рассматривать металлическую связь как некоторую особенную «резонансную» ковалентную связь [15,20].

Под *ван-дер-ваальсовым радиусом* принято понимать половину равновесного межъядерного расстояния между ближайшими атомами, соединенными ван-дер-ваальсовой связью. Ван-дер-ваальсовы радиусы определяют эффективные размеры атомов благородных газов. Кроме того, как следует из определения, ван-дер-ваальсовым атомным радиусом можно считать половину межъядерного расстояния между ближайшими одноименными атомами, связанными ван-дер-ваальсовой связью и принадлежащими разным молекулам (например, в молекулярных кристаллах). При сближении атомов на расстояние, меньшее суммы их ван-дер-ваальсовых радиусов, возникает сильное межатомное отталкивание. Поэтому ван-дер-ваальсовы атомные радиусы характеризуют минимально допустимые контакты атомов, принадлежащих разным молекулам. Данные о величинах ван-дер-ваальсовых атомных радиусов для некоторых атомов можно найти в [15,20]).

Знание ван-дер-ваальсовых атомных радиусов позволяет определять форму молекул, их упаковку в молекулярных кристаллах. Ван-дер-ваальсовы радиусы много больше всех перечисленных выше радиусов тех же самых элементов, что объясняется слабостью ван-дер-ваальсовых сил.

2.6. Полупроводниковые соединения

2.6.1. Электроотрицательность элементов

Одной из принципиальных особенностей химических соединений по сравнению с элементами является их гетерополярность. Компоненты соединения обладают разной способностью отдавать и присоединять валентные электроны. Эта способность характеризуется *электроотрицатель-*

ностью атома X . Атом B при взаимодействии с атомом A может захватить у него электрон, а при взаимодействии с атомом C отдать ему свой электрон. По величине электроотрицательности атомы при этом располагаются в ряд: $A < B < C$. Электроотрицательностью атома принято называть полусумму потенциала ионизации I и энергии сродства атома к электрону E (шкала электроотрицательностей по Малликену):

$$X = (I + E)/2. \quad (2.7)$$

Первый потенциал ионизации I атома равен энергии, необходимой для отрыва его первого валентного электрона от атома. Сродством атома к электрону называется способность атомов присоединять добавочный электрон и превращаться в отрицательный ион.¹² Мерой сродства атома к электрону является энергия электронного сродства, равная разности энергии атома в основном состоянии и энергии основного состояния соответствующего отрицательного иона. Энергия сродства атома к электрону для устойчивого отрицательного иона положительна.

Потенциалы ионизации измеряются методом фотоионизации и спектроскопическими методами. Обычно потенциал ионизации выражают в электрон-вольтах (эВ), отнесенных к одному атому. Для экспериментального определения величины сродства атома к электрону существуют такие прямые методы, как метод фотоэлектронной эмиссии, метод захвата электронов и др; кроме того, ее можно вычислить из известных энергий решеток ионных кристаллов. Выражается энергия сродства атома к электрону, как правило, также в эВ.

Если расположить элементы в ряд по электроотрицательности, то этот ряд начнется с наименее электроотрицательных элементов, то есть наиболее склонных к образованию положительных ионов (щелочные металлы, для которых I и E малы). На другом конце этого ряда будут наиболее электроотрицательные элементы, то есть наиболее легко образующие отрицательно заряженные ионы (галогены, для которых I и E велики).

Следует отметить, что в то время как потенциалы ионизации могут быть определены с большой степенью точности, энергия сродства атома к электрону не всегда может быть определена с высокой надежностью. Для построения практической шкалы электроотрицательностей элемен-

¹²Статического электрического поля нейтрального атома недостаточно для того, чтобы атом смог принять дополнительный электрон, однако этот электрон наводит в атоме электрический дипольный момент, в результате чего возникает потенциал притяжения, действующий на больших расстояниях. Во многих случаях этот поляризационный потенциал достаточно велик для того, чтобы свободный атом смог присоединить добавочный электрон.

тов поэтому применяются условные единицы, значения которых выводят следующим образом (шкала электроотрицательностей по Полингу).

Рассмотрим два сорта атомов А и В. Пусть когда одинаковые атомы А или В соединяются в молекулы, то между ними образуется ковалентная связь. Обозначим энергию связи в таких молекулах D_{AA} и D_{BB} . Если ковалентная связь образуется между разнородными атомами А и В, то энергию связи D_{AB} для молекулы АВ можно, вообще говоря, представить как среднее энергий связей D_{AA} и D_{BB} — $D_{AB} = (D_{AA} + D_{BB})/2$. Однако опыт показывает, что D_{AB} фактически больше этой величины, то есть $D_{AB} - (D_{AA} + D_{BB})/2 = \Delta > 0$.¹³ Причина этого несоответствия заключается в поляризации связи, то есть образовании положительно и отрицательно заряженных ионов A^{q*+} и B^{q*-} , между которыми возникает кулоновское взаимодействие. Появляется ионная составляющая в энергии связи — Δ . Л. Полинг установил, что по мере увеличения разности электроотрицательностей $\Delta X = X_A - X_B$ между атомами А и В увеличивается и Δ . Кроме того, он выразил параметр Δ через разность электроотрицательностей взаимодействующих атомов следующим образом: Δ (эВ) $\sim (X_A - X_B)^2$. Согласно этой формуле, вклад в Δ дает только разность электроотрицательностей атомов А и В, а абсолютные значения X_A и X_B не являются определяющими. Они могут выражаться в относительных единицах. В периодической таблице элементов значение $X = 1$ приписано атому лития, а для остальных элементов значения X найдены из известных разностей ΔX .

На рис. 2.19 приведены значения электроотрицательностей элементов по шкале Полинга. Видно, что электроотрицательность элементов в периодической таблице слева направо и снизу вверх (в основном) растет. При движении слева направо в рамках одного периода и снизу вверх в пределах группы совершенно отчетливо проявляется тенденция к накоплению металлоидных свойств, то есть по мере роста электроотрицательности X сродство электрона к атому увеличивается и облегчается образование аниона.

Следует иметь в виду, что шкала Полинга, также как и шкала Малликена, не является универсальной. Существуют и другие шкалы электроотрицательностей. Естественно, что абсолютные значения электроотрицательностей и даже порядок их расположения на разных шкалах может быть различным. Тем не менее понятие электроотрицательности элементов используется, главным образом для оценки степени ионности

¹³Истинные энергии связи в двухатомных молекулах равны энергиям диссоциации молекулы на атомы и могут быть определены термохимическими или спектроскопическими методами.

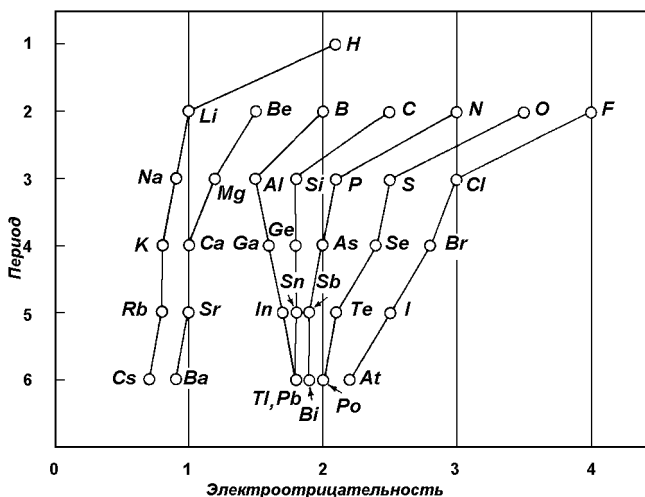


Рис. 2.19. Электроотрицательность некоторых элементов по шкале Полинга [20].

рассматриваемой связи, когда имеет значение только разность электроотрицательностей. На практике наиболее широко используется шкала, предложенная Полингом.

2.6.2. Степень ионности связи

Очевидно, что химическая связь в полупроводниковых соединениях является полярной ковалентной связью (ковалентно-ионной). Рассмотрим теперь вопрос о распределении электронной плотности между компонентами соединения.

При «чистой ковалентной» связи электронная плотность распределена совершенно симметрично между одинаковыми атомами, и ее центр симметрии располагается в середине межатомного расстояния. В «чисто ионных» соединениях распределение электронной плотности таково, что она в основном сосредоточена вокруг каждого иона и обладает (приблизленно) сферической симметрией. В случае же промежуточной, ковалентно-ионной связи, логично предположить, что центр симметрии электронного облака смещен от середины межатомного расстояния к одному из соседних атомов. Величина этого смещения определяет *степень ионности связи* δ . Очевидно, что степень ионности связи должна быть пропорциональна разности электроотрицательностей.

Таблица 2.4. Разность электроотрицательностей, степень ионности связи и структура бинарных соединений $A^N B^{8-N}$ [20].

Соединение	ΔX	δ	Стр. тип ^a	Соединение	ΔX	δ	Стр. тип
SiC	0.7	0.21	сф.	InSb	0.3	0.09	сф.
InAs	0.3	0.09	сф.	InP	0.4	0.12	сф.
GaSb	0.2	0.06	сф.	GaAs	0.4	0.12	сф.
GaP	0.5	0.15	сф.	AlAs	0.5	0.15	сф.
AlP	0.6	0.18	сф.	HgTe	0.2	0.06	сф.
HgSe	0.5	0.15	сф.	HgS	0.6	0.18	сф.
CdTe	0.4	0.12	сф.	CdSe	0.7	0.21	сф., в.
CdS	0.8	0.24	сф., в.	ZnTe	0.5	0.15	сф.
ZnSe	0.8	0.24	в., сф.	ZnS	0.9	0.27	сф., в.
AgI	0.6	0.18	сф., в.	AgBr	0.9	0.27	NaCl
CuCl	1.1	0.33	сф., в.	CuBr	0.9	0.27	сф., в.
ZnO	1.9	0.57	в.	BeO	2.0	0.60	в.
MgO	2.3	0.69	NaCl	NaCl	2.1	0.63	NaCl
LiF	3.0	0.90	NaCl	CsF	3.3	1.0	NaCl

^a Для обозначения структурных типов использованы следующие обозначения: сф. — сфалерит, в. — вюртцит, NaCl — структура каменной соли.

Найдем степень ионности связи бинарных соединений $A^N B^{8-N}$, где N — номер группы элемента А. Если предположить, что степень ионности связи δ линейно связана с разностью электроотрицательностей компонент ΔX_{AB} , то для соединений $A^N B^{8-N}$ ее можно рассчитать по формуле $\delta = (\Delta X_{AB}/\Delta X_{CsF})\delta_{CsF}$, где ΔX_{CsF} — разность электроотрицательностей CsF, степень ионности связи которого принимается за 1. Результаты расчетов представлены в табл. 2.4. Из нее следует, что соединения $A^{IV}B^{IV}$, $A^{III}B^V$ и $A^{II}B^{VI}$, а также часть соединений $A^I B^{VII}$ имеют степень ионности связи менее 0.3 при разности электроотрицательностей менее 1. Вся эта многочисленная группа отделена значительным промежутком от небольшой группы соединений $A^I B^{VII}$, у которых δ больше 0.5, а $\Delta X > 1$. К этой группе относятся соединения $A^I B^{VII}$ — CsF, LiF, NaCl, KCl и т. д.

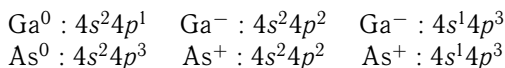
Выделенная многочисленная группа соединений проявляет полупроводниковые свойства. Межатомная связь в этих соединениях представляет собой комбинацию ковалентного и ионного типов связи с преимуществом ковалентного. Соединения этой группы кристаллизуются в структуры типа сфалерита или вюртцита, в которых $Z_k = 4$. Ко второй, небольшой, группе относятся соединения с преимущественно ионным типом

химической связи и кристаллической структурой типа NaCl ($Z_k = 6$). Это — диэлектрики.

2.6.3. Химическая связь в полупроводниковых соединениях $A^N B^{8-N}$

Проследим, каким образом реализуется ковалентная связь в полупроводниковых бинарных соединениях $A^N B^{8-N}$, у которых $\Delta X < 1$. Было установлено, что неполярная ковалентная связь, приводящая к формированию алмазоподобной структуры ($Z_k = 4$), наблюдается в гомеополярных кристаллах при образовании в них sp^3 -гибридных связей. Такие гибридные связи могут образовываться и в соединениях $A^N B^{8-N}$, поскольку во всех них среднее число валентных электронов на атом компонента соединения равно 4, как и у элементов IVA подгруппы.

Рассмотрим, например, соединение GaAs. Известно, что оно имеет структуру типа сфалерита с координационным числом $Z_k = 4$ и обладает полупроводниковыми свойствами. Наличие sp^3 -гибридных ковалентных связей обеспечивается в нем обобществлением электронов по схеме:



Таким образом, видно, что для образования ковалентной связи в соединениях $A^N B^{8-N}$ необходим «переход» валентных электронов от одного компонента соединения к другому и, как следствие, возникновение у них зарядов, с последующим обобществлением валентных электронов. Отсюда ясно, что ковалентная связь в «чистом» виде в соединениях не является нейтральной, как в элементах IVA подгруппы; причем образующиеся заряды имеют *антиионные* знаки, которые противоположны знакам, ожидаемым на основании положения атомов в ряду электроотрицательностей. Особенно большие антиионные заряды получаются у соединений $A^I B^{VII}$.

Найдем для бинарных соединений такие распределения валентных электронов между атомами компонентов соединения A и B, которые в одном случае позволяют представить это соединение как соединение с «чисто» ковалентной связью, а в другом как с «чисто» ионной. Снова рассмотрим GaAs. Для ковалентной связи соответствующая ковалентная формула, как было показано выше, — $\text{Ga}^{I-} \text{As}^{I+}$. Соединение с «чисто» ионной связью можно получить, превращая атомы компонентов соединения в противоположно заряженные ионы с таким зарядом, который бы позволил каждому атому компонента соединения приобрести элек-

тронную конфигурацию благородного газа. Для GaAs ионная формула — $\text{Ga}^{3+}\text{As}^{3-}$.

Запишем промежуточную ковалентно-ионную связь как комбинацию «чисто ионного» и «чисто ковалентного» типов связи. Это означает, что в соединениях волновая функция, описывающая состояния валентных электронов, может быть формально представлена в виде

$$\Psi = C_1\Psi_{\text{к}} + C_2\Psi_{\text{и}}, \quad C_1^2 + C_2^2 = 1, \quad (2.8)$$

где $\Psi_{\text{к}}$ описывает электронные состояния для «чистой ковалентной», а $\Psi_{\text{и}}$ — для «чистой ионной» связи. Коэффициенты C_1 и C_2 характеризуют относительный вклад обеих составляющих в реальную систему связей. Их величины зависят от свойств атомов компонентов соединения А и В, включая их разность электроотрицательностей. Реальная связь должна соответствовать таким значениям коэффициентов C_1 и C_2 , при которых энергия состояния, описываемого волновой функции Ψ , была бы минимальна.

Если атомы компонентов соединения идентичны (элементарные вещества), то $C_2 = 0$ и связь — «чистая ковалентная». В этом случае обменное электронное облако полностью симметрично относительно центра между атомами, а взаимодействующие атомы нейтральны. Если атомы компонентов соединения различны, но $C_2 \ll C_1$, то связь будет преимущественно ковалентной, а атомы компонентов соединения заряжены антионно и электронная плотность смещена в сторону компонента А. С ростом ионной составляющей связи C_2 антионный заряд будет уменьшаться благодаря смещению электронного облака в сторону более электроотрицательного атома компонента В. Таким образом, при смешанной ковалентно-ионной связи электронное облако является *подвижным и асимметричным*. Это свойство полупроводниковых соединений составляет их важнейшую особенность. При сохранении кристаллической структуры с $Z_{\text{к}} = 4$ подвижное электронное облако с увеличением разности электроотрицательностей компонентов соединения может смещаться к более электроотрицательному элементу, в результате чего могут не только полностью исчезать антионные заряды, но и создаваться эффективные заряды, соответствующие ионным, и тем не менее соединение будет оставаться полупроводником.

Степень ионности полупроводникового соединения принято характеризовать такой величиной как ионность решетки. В качестве параметра ионности решетки примем величину вероятности ионной конфигурации $\lambda = C_2^2$. Введем также понятие *равновесной* ионности, определив ее как $\lambda_0 = 1 - n_{\text{A}_e}/N_{\text{к}}$, где n_{A_e} — число электронов, отдаваемых компонентом

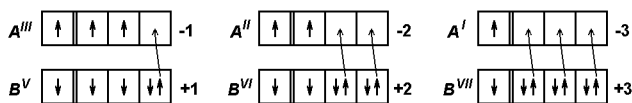


Рис. 2.20. Схемы образования sp^3 -гибридных связей у соединений $A^N B^{8-N}$.

соединения А при образовании «чисто» ионной связи, а N_k — число ковалентных связей, задаваемых данной кристаллической структурой. Величина λ_0 соответствует такой ионности решетки λ , при которой атомы компонентов соединения не заряжены из-за равенства зарядов, образующихся в результате «чисто ковалентного» и «чисто ионного» взаимодействий. Для рассматриваемых нами соединений $A^N B^{8-N}$ $N_k = 4$, а n_{A_e} равно номеру группы А компонента. Таким образом, равновесная ионность разных групп будет

соединение	$A^{IV}B^{IV}$	$A^{III}B^V$	$A^{II}B^{VI}$	A^IB^{VII}
равновесная ионность	0	0.25	0.5	0.75

Заметим, что равновесная ионность в ряду $A^{IV}B^{IV}$, $A^{III}B^V$, $A^{II}B^{VI}$, A^IB^{VII} увеличивается. Это и понятно, поскольку в $A^{IV}B^{IV}$ гибридизация происходит без передачи электронов, в $A^{III}B^V$ — при передаче одного электрона; в $A^{II}B^{VI}$ — двух электронов и в A^IB^{VII} — трех электронов (рис. 2.20).

Соотношение между λ и λ_0 имеет вид:

$$\lambda = \lambda_0 + q^*/eN_k, \quad (2.9)$$

где λ — ионность решетки, λ_0 — равновесная ионность, q^* — эффективный заряд, N_k — число ковалентных связей. Величину q^*/eN_k можно интерпретировать как относительный перемещенный заряд связи или полярность связи. Таким образом, понятие степень ионности соединения можно характеризовать двумя параметрами: собственно ионностью решетки, выраженной через $\lambda = C_2^2$, то есть полный перемещенный заряд; и полярностью связи, выраженной через q^*/eN_k , то есть относительный перемещенный заряд.

Значения эффективных зарядов часто используются для того, чтобы охарактеризовать полярность связи и степень ионности полупроводникового соединения. Величины q^* могут быть определены экспериментально из спектров ИК-поглощения и рамановского рассеяния, неупругого рассеяния нейтронов; могут быть теоретически рассчитаны. Однако понятию «эффективный заряд» атомов в соединениях с преимущественно

ковалентной связью присуща некоторая неопределенность. Действительно, под зарядом атома обычно понимают заряд, находящийся внутри объема атома. Для ионного кристалла это определение вполне приемлемо, так как электронная плотность сосредоточена в основном вокруг иона, а его радиус можно определить достаточно точно. В кристаллах с преимущественно ковалентной связью определить объем, ограничивающий один атом, довольно трудно. На основании этого понятию «эффективный заряд» не следует придавать буквального значения, а нужно рассматривать его как некоторый экспериментально определяемый параметр соединений, позволяющий их классифицировать и характеризовать. В табл. 2.5 приведены величины эффективных зарядов, определенные различными методами, и ионности решетки, рассчитанные с помощью соотношения (2.9), для соединений $A^N B^{8-N}$.

2.6.4. Металлическая составляющая связи, ионная составляющая связи и ширина запрещенной зоны в полупроводниках

Деление веществ по преимущественному характеру межатомной связи (ионная, ковалентная, металлическая) соответствует их качественному делению на диэлектрики, полупроводники и металлы, поскольку характер химической связи определяет физические свойства материалов. Естественно, поэтому предпринимались и предпринимаются многочисленные попытки установить корреляцию между важнейшими параметрами полупроводниковых материалов (шириной запрещенной зоны E_g , подвижностью носителей заряда μ , теплопроводностью κ) и их кристаллохимиче-

Таблица 2.5. Значения эффективных зарядов и ионности решетки в соединениях $A^N B^{8-N}$ [21–24].

Соединение	AlSb	GaSb	InSb	GaAs	InAs	GaP	InP
q^*/e	0.48	0.33	0.42	0.46	0.56	0.58	0.68
λ	0.37	0.33	0.36	0.40	0.39	0.40	0.42
Соединение	ZnO	ZnS	ZnSe	CdS	CdSe	CdTe	HgS
q^*/e	1.06	0.81	0.7	0.73	0.55	0.71	0.77
λ	0.77	0.70	0.68	0.68	0.64	0.68	0.69
Соединение	HgSe	HgTe	CuCl	CuBr	CuI	AgI	
q^*/e	0.71	0.52	0.59	0.38	0.18	0.23	
λ	0.68	0.63	0.9	0.85	0.79	0.81	

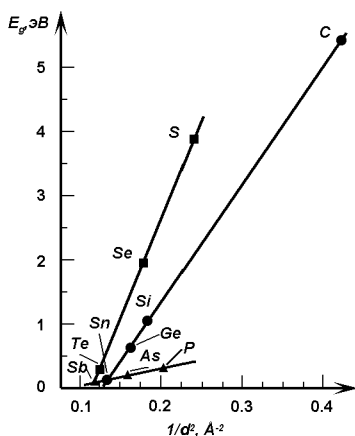


Рис. 2.21. Зависимость ширины запрещенной зоны E_g от длины связи d в полупроводниках IVА–VІА подгрупп (см. табл. 2.3).

скими характеристиками (длиной связи d , эффективными зарядами q^* , разностью электроотрицательностей ΔX и т. д.).

Для полупроводниковых элементов различных подгрупп из таблицы Д. И. Менделеева была построена зависимость ширины запрещенной зоны E_g этих элементов от длины связи d (рис. 2.21). Оказалось, что для элементов разных подгрупп ширина запрещенной зоны пропорциональна обратному квадрату длины связи. Такая зависимость обусловлена, как уже обсуждалось ранее, уменьшением ковалентной составляющей связи и увеличением металлической составляющей связи (эффект экранирования ядра: связь ядро — валентные электроны ослабевает из-за увеличения числа заполненных оболочек и уменьшается прочность ковалентной связи).

О влиянии ионной составляющей связи на ширину запрещенной зоны в полупроводниковых соединениях можно достаточно однозначно судить по изменениям E_g в *изоэлектронных рядах* этих соединений. Такой ряд образуют элементы и соединения, компонентами которых являются элементы из одного периода периодической таблицы, сумма атомных номеров которых постоянна и среднее число электронов на атом одинаково.

Из табл. 2.6 следует, что в каждом изоэлектронном ряду α -Sn, Ge и Si длины связей d с высокой точностью совпадают, в то время как ширина запрещенной зоны E_g монотонно возрастает при переходе от элемента-родоначальника ряда к соединениям с все более увеличивающейся степенью ионности (рост λ связан с увеличением разности электроотрицательностей атомов компонентов соединения; см. табл. 2.4 и табл. 2.5).

Таблица 2.6. Значения d , λ , E_g , μ_n и μ_p в изоэлектронных рядах Sn, Ge, Si приведены для 300 К [24].

Ряд	d , Å	λ	E_g , эВ	μ_n , см ² /В·с	μ_p , см ² /В·с
α - Sn	2.82	0	0.09	2500	2400
InSb	2.82	0.36	0.18	78000	750
CdTe	2.81	0.68	1.5	1200	80
AgI	2.81	0.81	2.80	50	—
Ge	2.44	0	0.664	3800	1820
GaAs	2.44	0.4	1.43	8500	420
ZnSe	2.45	0.68	2.80	260	23
CuBr	2.46	0.85	2.90	30	—
Si	2.33	0	1.11	3000	500
AlP	2.33	0.40	2.45	50	150

Изоэлектронные ряды Si и C представлены в таблице в укороченном виде или отсутствуют, соответственно, потому, что последующие за приведенными представители этих рядов не обладают алмазоподобной структурой и не являются полупроводниками.

Постоянство d в рядах дает основание полагать, что ковалентная составляющая связи в них не изменяется, а увеличение E_g с одновременным ростом степени ионности соединения с алмазоподобной структурой обусловлено ростом ионной составляющей связи. Кроме того, из табл. 2.6 видно, что, во-первых, происходит замедление роста E_g при переходе от одного ряда к другому, и, во-вторых, происходит замедление нарастания E_g при переходе к ионным соединениям в каждом ряду. Замедление роста E_g при переходе от одного ряда к другому свидетельствует о том, что ширина запрещенной зоны является сложной функцией λ . Замедление же нарастания E_g в каждом ряду обусловлено эффектом деполяризации ковалентно-ионной связи. Действительно, с ростом степени ионности соединения возрастает степень компенсации между «антиионным» и «ионным» зарядами компонентов соединения. В результате в определенном интервале достаточно больших λ ионная составляющая связи стабилизируется. Подобный эффект является спецификой ковалентно-ионной связи, имеющей подвижный заряд.

В полупроводниковых соединениях, не принадлежащих к изоэлектронным рядам, зависимости $E_g(\lambda)$ определяется взаимным положением элементов А и В в периодической таблице. Так, например, в ряду соединений AlSb, GaSb и InSb уменьшение E_g ($E_g = 1.63$ эВ; 0.7 эВ; 0.18 эВ соответственно) происходит в основном за счет уменьшения ковалентной составляющей связи практически без изменения ионной (см. табл. 2.5).

Прочность ковалентной составляющей связи в этом ряду падает, а доля металлической составляющей связи увеличивается с возрастанием атомного номера A компонента соединения из-за роста эффекта экранирования. В ряду AlSb , AlAs , AlP рост E_g ($E_g = 1.63$ эВ; 2.14 эВ; 2.45 эВ соответственно) обусловлен как небольшим ростом степени ионности соединения ($\lambda = 0.37$; 0.39 ; 0.68 соответственно), так и увеличением ковалентной составляющей связи, связанной с уменьшением эффекта экранирования.

2.6.5. Ионная составляющая связи и подвижность носителей заряда

Установление взаимосвязи между степенью ионности полупроводниковых соединений λ и подвижностями носителей заряда μ в них затруднительно, прежде всего, из-за сильной чувствительности μ к дефектам кристаллов. В то время как измерения E_g в образцах с широким диапазоном значений концентрации примесей и дефектов дают одно и то же значение, для измерения «решеточной» подвижности μ_L (см. ниже) необходимо иметь чистые и достаточно совершенные кристаллы. Поэтому далеко не во всех полупроводниках по измеренной при некоторой температуре подвижности можно говорить о надежном определении μ_L . С другой стороны, сама теория химических связей не дает возможности рассчитывать численные величины подвижности носителей заряда из значений кристаллохимических параметров, характеризующих тип химической связи, можно говорить лишь о тенденциях, знание которых полезно для прогнозирования свойств получаемых полупроводниковых материалов.

Прежде чем рассматривать влияние ионной составляющей связи на подвижность носителей заряда μ , обсудим некоторые вопросы, связанные с самим понятием μ . Подвижность носителей заряда определяется их рассеянием в кристалле. В идеальной кристаллической решетке рассеяние отсутствует. В реальных же кристаллах рассеяние носителей происходит на нарушениях периодичности кристаллической структуры, то есть любых отклонениях силового поля решетки от идеально периодического. Эти нарушения могут быть вызваны разными причинами: 1) тепловыми колебаниями атомов или ионов, составляющих кристаллическую решетку; 2) примесями, которые могут находиться в кристалле либо в ионизованном, либо в нейтральном состоянии; 3) всевозможными дефектами решетки (вакансиями, дислокациями, границами кристаллических зерен, трещинами и т. д.; см. гл. 3). Кроме того, носители могут

рассеиваться друг на друге.

Ограничимся рассмотрением рассеяния носителей заряда на тепловых колебаниях решетки, то есть рассмотрением «решеточной» подвижности (при этом необходимо пользоваться данными для чистых и структурно совершенных кристаллов). При достаточно высоких температурах атомы решетки совершают малые тепловые колебания около своих равновесных положений. Среди возможных типов колебаний выделяют акустические колебания¹⁴ и оптические колебания.¹⁵ Акустические колебания отвечают смещениям элементарной ячейки как целого, а оптические — внутренним деформациям в ней при почти неподвижном центре тяжести ячейки. Эти малые колебания распространяются по всему кристаллу в виде волн. Введя специальные, так называемые нормальные, координаты, полную энергию колеблющегося кристалла можно представить как сумму энергий невзаимодействующих квазичастиц с энергией $\hbar\omega(\mathbf{q})$ и квазимпульсом $\hbar\mathbf{q}$, где $\omega(\mathbf{q})$ — частота колебаний атомов кристалла, а $\mathbf{q}$ — волновой вектор волны. Эти квазичастицы носят название *фононов*. Согласно существующим представлениям, рассеяние носителей заряда на тепловых колебаниях решетки можно рассматривать как их взаимодействие с фононами или, что тоже самое, с колеблющейся решеткой. Это взаимодействие сводится к поглощению или испусканию фонона, при этом увеличивается или уменьшается, соответственно, энергия электрона.

Различают неполярное и полярное рассеяние. Неполярное рассеяние на акустических фононах возникает вследствие того, что при распространении волн в кристалле на периодический потенциал идеальной кристаллической решетки накладывается дополнительный периодический потенциал, вызванный смещением атомов из положений равновесия (деформацией решетки). Это приводит к изменению потенциальной энергии носителя заряда, выражение для которой оказывается того же типа, что и аналогичное выражение при однородной статической деформации, возникающей при сжатии или растяжении кристалла. Поэтому метод описания рассеяния (изменения энергии) носителей заряда на акустических фононах называли методом потенциала деформации [6]. Аналогичные рассуждения могут быть проведены и для неполярного рассеяния на оптических фононах. Однако конкретное выражение для потенциала деформации при рассеянии на акустических и оптических фононах получается разное [6].

¹⁴При длинноволновых акустических колебаниях все атомы элементарной ячейки смещаются практически синфазно.

¹⁵При оптических колебаниях атомы движутся в противофазе, а центр тяжести элементарной ячейки остается на месте.

Практически неполярное рассеяние наиболее ярко проявляется в гомеопольярных (атомных) полупроводниках с преимущественно ковалентными связями (Si, Ge и др.).

Механизм полярного рассеяния связан с существованием эффективных зарядов у атомов решетки. Оптические колебания решетки вызывают изменение дипольного момента элементарной ячейки, то есть колебания вектора поляризации среды.¹⁶ Это приводит к изменению энергии электрона, не связанному с потенциалом деформации. Следует заметить, что в гетерополярных кристаллах одновременно и независимо с полярным рассеянием может происходить и неполярное рассеяние, однако оно менее эффективно. Это связано с тем, что силы, отвечающие за возникновение полярного рассеяния, — дальнodelствующие, в результате чего на электрон одновременно оказывают влияние изменения дипольного момента во многих элементарных ячейках. Поэтому обычно полярное рассеяние оказывается преобладающим в гетерополярных полупроводниках с ковалентно-ионными связями и не слишком малыми эффективными зарядами (InAs, GaAs и др.). Как правило, подвижность носителей при достаточно высоких температурах (когда преобладает рассеяние на колебаниях решетки)¹⁷ в кристаллах с большой долей ионной составляющей связи значительно ниже, чем в кристаллах с меньшей долей.

В кристаллах, не имеющих центра инверсии и состоящих из разноименно заряженных ионов, механическая деформация сопровождается их электрической поляризацией и возникновением электрического поля. Этот пьезоэлектрический эффект обусловлен тем, что при механической деформации такого кристалла отдельные подрешетки, каждая из которых состоит из ионов одного и того же знака, могут смещаться относительно друг друга и при этом возникает электрический момент. При колебаниях решетки в таких кристаллах на носители заряда действуют силы со стороны электрического поля, что приводит к дополнительному механизму рассеяния. Это так называемое пьезоэлектрическое рассеяние на акустических фоновых. Оно наблюдается в кристаллах соединений $A^{III}B^V$ (GaAs, InSb), $A^{II}B^{VI}$ (CdS, CdSe) и других.

Экспериментально установлено, что в чистых элементарных полупроводниках (с преобладающей долей ковалентной связи) при комнатной температуре основную роль играет неполярное рассеяние носителей заря-

¹⁶По этой причине колебания рассматриваемого типа иногда называют поляризационными.

¹⁷При низких температурах, когда колебания решетки не интенсивны, то есть число фононов мало, преобладает рассеяние носителей на нейтральных атомах, точечных дефектах или ионах примеси.

да на акустических фононах. К нему добавляется неполярное рассеяние на оптических фононах. При переходе к соединениям (чистые образцы), имеющим ионную составляющую связи, начинает преобладать рассеяние носителей заряда на поляризационных фононах. Поэтому можно ожидать уменьшения μ_L при переходе от элементарных полупроводников к соединениям изоэлектронного ряда. Однако в наиболее обширных изоэлектронных рядах α -Sn и Ge подвижность электронов μ_n сначала растет, а лишь затем начинает резко падать (табл. 2.6). Для того чтобы объяснить этот факт, необходимо иметь в виду, и это очень существенно, что при сравнении подвижностей носителей в разных кристаллах следует учитывать, что на их величину, кроме степени ионности соединения, также влияет и масса носителей заряда, и сама зонная структура кристалла [6]. Какой из факторов окажется преобладающим, часто трудно предсказать.

В соединениях $A^{III}B^V$ максимум подвижности $\mu \simeq 78000 \text{ см}^2/\text{В} \cdot \text{с}$ наблюдается у InSb, в $A^{II}B^{VI}$ максимум $\mu \simeq 2300 \text{ см}^2/\text{В} \cdot \text{с}$ наблюдается у HgTe, в $A^I B^{VII}$ максимум $\mu \simeq 50 \text{ см}^2/\text{В} \cdot \text{с}$ наблюдается у AgI. Во всех этих трех соединениях q^* оказывается минимальным или близким к минимальному значению эффективных зарядов, характерных для соответствующей группы соединений, поэтому полярное рассеяние в них минимально. Таким образом, в каждой из групп соединений $A^N B^{8-N}$ максимум подвижности наблюдается вблизи минимального значения q^* , а при переходе от $A^{III}B^V$ к $A^{II}B^{VI}$ и далее к $A^I B^{VII}$ величина максимальной подвижности падает с увеличением равновесной ионности соединения, так что и относительная полярность связи $\sim q^*$, и равновесная ионность определяют величины подвижности электронов в полупроводниковых соединениях с ковалентно-ионной межатомной связью.

Подвижности дырок во всех соединениях $A^N B^{8-N}$, как правило, не превышают их величин в полупроводниках IVA подгруппы. В изоэлектронных рядах, приведенных в табл. 2.6, подвижность дырок резко уменьшается при переходе от элементарных полупроводников к соединениям и далее продолжает уменьшаться с ростом степени ионности соединения.

2.6.6. Химические связи в полупроводниках, производных от $A^N B^{8-N}$

Для развития прикладной полупроводниковой электроники требуется получение полупроводниковых материалов с заданными свойствами. В рамках обычного эмпирического подхода с этой целью синтезируют большое число кристаллов разного состава, изучают их свойства и находят тот

материал, который обладал бы заданными параметрами. В такой ситуации, естественно, весьма желательно сформулировать некоторые критерии, которые давали бы возможность, не затрачивая больших средств и времени на изготовление огромного числа разных кристаллов, заранее исключить из рассмотрения те материалы, которые не являются полупроводниками или не обладают требуемыми свойствами. Проще всего решения этой проблемы можно достичь используя теорию химической связи. Несмотря на то, что значения кристаллохимических параметров этой теории, как правило, не являются строго определенными, знание основных закономерностей изменения физико-химических свойств известных полупроводников с изменением этих параметров позволяет в ряде случаев синтезировать аналоги известных материалов с заранее прогнозируемыми свойствами.

Анализируя химическую связь в полупроводниках, Музер и Пирсон [25] ввели понятие о некоторой специфической «полупроводниковой связи»,¹⁸ основной составной частью которой является ковалентная связь, обеспечивающая полупроводниковые свойства кристаллов. Они пришли к выводу, что для образования этой связи в элементарных полупроводниках необходимо наличие полностью заполненных *s*- и *p*-орбиталей в валентных оболочках всех атомов. В полупроводниковых бинарных соединениях достаточно, чтобы хотя бы один из двух связанных атомов обладал заполненными *s*- и *p*-валентными орбиталями. Присутствие пустых «металлических» орбиталей у другого атома, входящего в состав соединения, не уничтожает полупроводниковых свойств, если эти атомы не связаны друг с другом. Связи в полупроводниках должны образовывать одно-, двух- или трехмерную решетку, простирающуюся на весь кристалл.

Принимая во внимание выводы Музера и Пирсона и учитывая накопленный при исследовании полупроводников опыт, установили, что большинство бинарных полупроводников подчиняются следующим эмпирическим закономерностям.

1. Полупроводник должен содержать в своем составе атом элемента, принадлежащего к IVA–VIIA подгруппам периодической системы. Поясним это правило.

Это правило эквивалентно требованию присутствия в полупроводниковом соединении атомов В, у которых валентная оболочка будет заполняться полностью и которые, как правило, имеют большую электроотрицательность, чем электроотрицательность атомов А. При этом предполагается, что у атомов А возможно наличие свободных «металлических»

¹⁸На самом деле химические связи в полупроводниках, конечно же, смешанные.

орбиталей. Принцип полного заполнения валентной зоны соединения не нарушается, если атомы А непосредственно не связаны друг с другом. Однако наличие свободных «металлических» орбиталей может вызвать образование резонансной связи (см. ниже). В этом случае число ближайших соседей, с которыми образуется химическая связь, у атомов В будет превышать их валентность.

2. В полупроводнике для валентных электронов выполняется соотношение, называемое правилом Музера-Пирсона:

$$n_e/N_B + N_{BB} = 8, \quad (2.10)$$

где n_e — суммарное число валентных электронов на атом или молекулу (число валентных электронов, участвующих в образовании связи), N_B — число атомов В (анионов из IVA–VIIA группы) в молекуле, N_{BB} — число связей, образуемых В атомами между собой. Определение N_{BB} требует знания кристаллической структуры соединения.

3. Разность электроотрицательностей, входящих в состав полупроводниковых соединений элементов, должна быть менее 0.8–1. Это требование связано с тем, что степень ионности связи пропорциональна разности электроотрицательностей. Увеличение степени ионности связи приводит, как мы выяснили ранее, к увеличению E_g и к уменьшению $\mu_{n,p}$, то есть переходу к диэлектрикам.

4. Большинство полупроводников подчиняются еще и правилу *нормальной валентности*: для бинарных соединений $A_a^K B_b^L$

$$Ka = (8 - L)b, \quad (2.11)$$

где K и L — валентности элементов А и В, соответственно, в соединении $A_a B_b$, а a и b — числа, определяющие стехиометрический состав соединения.

Легко проверить, что эти правила выполняются во всех рассмотренных выше полупроводниковых соединениях $A^N B^{8-N}$. Например,

$$\begin{aligned} \text{GaAs (A}^{\text{III}}\text{B}^{\text{V}}\text{): } n_e &= 8; N_B = 1; N_{BB} = 0; n_e/N_B + N_{BB} = 8 \\ \text{CdTe (A}^{\text{II}}\text{B}^{\text{VI}}\text{): } n_e &= 8; N_B = 1; N_{BB} = 0; n_e/N_B + N_{BB} = 8 \\ \text{CuBr (A}^{\text{I}}\text{B}^{\text{VII}}\text{): } n_e &= 8; N_B = 1; N_{BB} = 0; n_e/N_B + N_{BB} = 8 \end{aligned}$$

Приведенные простые эмпирические закономерности оказались полезными с точки зрения прогнозирования полупроводниковых производных от известных полупроводниковых соединений, в частности большой группы полупроводников сложного состава (твердые растворы, бинарные, тройные и др. соединения), производной от $A^N B^{8-N}$.

Поскольку далее мы будем рассматривать не только соединения, производные от $A^N B^{8-N}$, но и твердые растворы, то необходимо дать определения понятиям «соединение» и «твердый раствор».

Все однофазные (см. гл. 4) твердые кристаллические тела представляют собой либо соединения, либо твердые растворы.

1. Соединения.

Соединение определяется стехиометрической формулой, с помощью которой задается строгое соотношение между числами атомов элементов, входящих в состав соединения. Также отличительным признаком соединения является наличие у него определенной кристаллической решетки, отличной от решеток элементов, из которых оно образовано. Например, соединение InSb имеет кубическую решетку, а элементы, из которых оно образовано имеют: In — тетрагональную решетку, а Sb — ромбоэдрическую. Для соединения характерно строгое расположение атомов, из которых оно состоит, в элементарной ячейке. Например, кристаллическая решетка InSb (структурный тип сфалерит) состоит из двух г.ц.к.-подрешеток, сдвинутых относительно друг друга на $1/4$ объемной диагонали, каждая из которых заполняется своим сортом атомов: одна — In, другая — Sb.

2. Твердые растворы.

Понятие «твердый раствор» было введено по аналогии с жидкими растворами, так как в твердых растворах для атомов растворяемого вещества характерно случайное расположение атомов в решетке растворителя, как в жидкостях, а не строго определенное, как в соединениях. Взаимодействующие вещества (растворитель и растворяемое вещество) могут иметь как неограниченную, так и ограниченную растворимость друг в друге с разной степенью этого ограничения и разной температурной зависимостью предела растворимости (см. гл. 4 и гл. 8). Таким образом, твердые растворы существуют не при определенном численном соотношении своих компонентов, а в интервале концентраций, величина которого определяется пределом растворимости. Кроме того, при образовании твердого раствора тип кристаллической решетки растворителя сохраняется, хотя постоянная решетки изменяется.

Таким образом, *твердые растворы* — это твердотельные двух- или многокомпонентные однородные системы переменного состава (например, типа $A_x B_{1-x}$), в которых атомы компонентов, смешиваясь в различных соотношениях ($0 \leq x \leq 1$), образуют общую кристаллическую решетку, характерную для одного из компонентов.

На основе растворителя — химического элемента — возможны два типа твердых растворов:

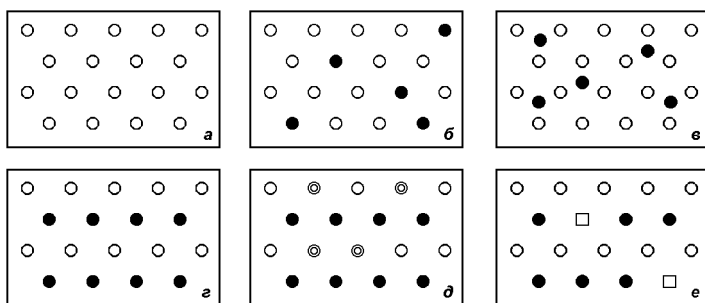


Рис. 2.22. Типы твердых растворов: *а* — растворитель; *б* — раствор замещения; *в* — раствор внедрения; *г* — растворитель; *д* — раствор замещения; *е* — раствор вычитания. *а*, *б*, *в* — растворитель–химический элемент, *г*, *д*, *е* — растворитель–соединение АВ. $\circ$ — атомы А, $\bullet$ — атомы В, $\odot$ — атомы С, $\square$ — незанятые узлы.

а) Растворы замещения: атомы растворяемого элемента занимают позиции атомов растворителя в узлах решетки, то есть замещают их (рис. 2.22,б). Примером может служить твердый раствор Ge–Si.

б) Растворы внедрения: атомы растворяемого элемента занимают междоузельные позиции в решетке растворителя, внедряясь между атомами последнего (рис. 2.22,в). Примеры: Ge(Li), Si(Li).

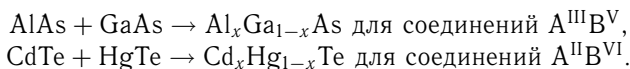
Если твердые растворы образуются на основе соединений, то кроме растворов замещения (рис. 2.22,д) и внедрения образуются еще и

в) Растворы вычитания: один из элементов, образующих соединение, присутствует в количестве, превышающем формульное, но при этом занимает в решетке соединения присущие ему позиции, а соответствующая часть позиций другого элемента остается вакантной (рис. 2.22,е). Например, $\text{Ni}_{49}\text{Sb}_{51}$, In_2Te_3 .

Обратимся вновь к вопросу о полупроводниковых производных от $\text{A}^{\text{N}}\text{B}^{8-\text{N}}$ с алмазоподобной структурой. Заметим, что помимо четырех общих закономерностей для бинарных полупроводников, для алмазоподобных выполняется следующее правило: в них на каждый атом в среднем приходится четыре валентных электрона и химические связи строятся на sp^3 -гибридных орбиталях. Поскольку правило Музера-Пирсона определяется суммарным числом валентных электронов и кристаллической структурой вещества, то это дает возможность получать новые вещества, обладающие полупроводниковыми свойствами, из полупроводниковых соединений АВ, замещая часть атомов А или (и) В элементами

из той же подгруппы с учетом остальных правил для сохранения полупроводниковых свойств. Практически это замещение осуществляется следующим образом. Возьмем два полупроводниковых соединения, например, с разными металлическими компонентами АС и ВС, в которых элементы А и В находятся в одной и той же подгруппе периодической системы. «Смешивание» этих двух соединений приведет к образованию твердого раствора $A_{1-x}B_xC$ в области его существования. Область существования — это часть фазовой диаграммы (см. гл. 4), в которой образуется непрерывный ряд твердых растворов и которая имеет тем большую протяженность, чем меньше различия между периодами решеток смешиваемых соединений и чем меньше различия между степенями ионности химических связей, то есть чем меньше различие между разностями электроотрицательностей в смешиваемых соединениях.

Примерами твердых растворов с замещением компонента А могут быть:



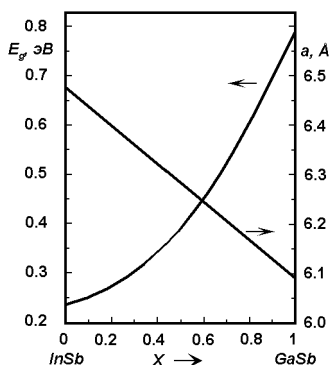
Выполнение закономерностей, характерных для полупроводников, в том числе и правила Музера-Пирсона, и сохранение алмазоподобной структуры для результирующих твердых растворов очевидно, так как в них, как и в исходных соединениях, реализуются смешанные ковалентно-ионные sp^3 -гибридные связи.

Как показывает опыт, при «смешении» полупроводниковых соединений с одинаковыми атомами А и разными атомами В, расположенными в одной и той же подгруппе периодической системы, также образуются полупроводниковые твердые растворы. Например, $GaSb_{1-y}As_y$ ($GaSb+GaAs$) и $CdSe_{1-y}Te_y$ ($CdSe+CdTe$). При таком замещении также очевидно выполнение всех закономерностей, приведенных выше, и сохранение алмазоподобной структуры.

Возможно одновременное замещение сразу двух компонент соединения атомами из соответствующих групп периодической системы с образованием четверного твердого раствора. Например, $(GaAs)_{1-x}(InP)_x$ ($GaAs+InP$).

Практическая важность получения подобных материалов для полупроводниковой электроники заключается в возможности, как правило, монотонного изменения за счет изменения состава ширины запрещенной зоны E_g от максимальной до минимальной величин, определяемых шириной запрещенных зон смешиваемых соединений (рис. 2.23). Таким образом, можно получить ряд кристаллов с непрерывно меняющейся E_g

Рис. 2.23. Зависимость ширины запрещенной зоны E_g и параметра решетки от состава твердого раствора $\text{In}_{1-x}\text{Ga}_x\text{Sb}$.



и даже кристаллы, в которых E_g меняется от точки к точке. Период кристаллической решетки обычно линейно зависит от состава (концентрации растворенного компонента). Такая зависимость называется *правилом Vegarda*. Концентрационная зависимости подвижности носителей, времени их жизни, интенсивности излучательной рекомбинации и оптического поглощения в твердых растворах прямозонных полупроводниковых материалов описываются плавными кривыми между значениями, характерными для составляющих их компонентов. В твердых растворах, образованных прямозонным и непрямозонным полупроводниковыми материалами, в области составов, где происходит изменение зонной структуры, наблюдаются резкие изменения свойств.

Далее следует обратить внимание на то, что в правиле Музера-Пирсона оговаривается только число В атомов элемента, входящего в состав полупроводника, а роль А атомов сводится к добавлению электронов в суммарное число валентных электронов n_v . Это позволяет предполагать, что можно посредством замены компонента А в исходном соединении получать не только полупроводниковые твердые растворы, производные от соединений $\text{A}^{\text{N}}\text{B}^{8-\text{N}}$, но и другие бинарные, тройные и более сложные полупроводниковые соединения, но уже, возможно, со структурой, производной от алмазоподобной. В этом случае замещающие элементы выбираются из групп периодической таблицы, отличных от той, в которой расположен замещаемый атом А, однако при этом должны удовлетворяться общие закономерности образования полупроводников (см. выше).

Ниже приведены ряды соединений, получающихся из $\text{A}^{\text{N}}\text{B}^{8-\text{N}}$ при горизонтальном замещении, с примерами полупроводниковых соединений

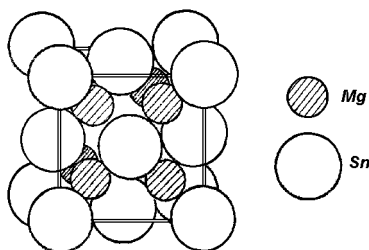


Рис. 2.24. Структура Mg_2Sn (антифлюорита).

соответствующих типов.

1.	$\text{A}^{\text{IV}}\text{B}^{\text{IV}} \rightarrow \text{A}_2^{\text{II}}\text{B}^{\text{IV}}$	$\text{Mg}_2(\text{Si}, \text{Ge}, \text{Sn}, \text{Pb});$
2.	$\text{A}^{\text{III}}\text{B}^{\text{V}} \rightarrow \text{A}_3^{\text{I}}\text{B}^{\text{V}}$	$(\text{Li}, \text{Rb}, \text{Cs})_3(\text{Sb}, \text{Bi});$
3.	$\text{A}^{\text{II}}\text{B}^{\text{VI}} \rightarrow \text{A}_5^{\text{I}}\text{B}^{\text{VI}}$	$(\text{Cu}, \text{Ag})_2(\text{O}, \text{S}, \text{Se}, \text{Te});$
4.	$3\text{A}^{\text{II}}\text{B}^{\text{VI}} \rightarrow \text{A}_2^{\text{III}}\text{B}_3^{\text{VI}}$	$(\text{Al}, \text{Ga}, \text{In})_2(\text{O}, \text{S}, \text{Se}, \text{Te})_3$

Рассмотрим, как в приведенных примерах реализуется ковалентная связь и какое влияние на кристаллическую структуру оказывает горизонтальное замещение компонента А в соединениях $\text{A}^{\text{N}}\text{B}^{8-\text{N}}$.

Соединения с резонансной связью

В соединениях первого типа наиболее электроотрицательный элемент IVA подгруппы играет роль компонента В. Рассмотрим этот тип замещения на примере соединения Mg_2Sn . Mg_2Sn кристаллизуется в структуре антифлюорита, которая показана на рис. 2.24. Видно, что эта структура отличается от структуры алмаза, а в молекуле два атома А и число валентных электронов в расчете на один атом меньше 4 ($8/3 < 4$). Тем не менее, Mg_2Sn является полупроводником, а правила нормальных валентностей ($K = 2$; $a = 2$; $L = 4$; $b = 1$), Музера-Пирсона ($n_e = 8$; $N_B = 1$; $N_{BB} = 0$) и остальные общие полупроводниковые закономерности выполняются.

В структуре антифлюорита каждый атом олова окружен восемью атомами магния, а каждый атом Mg — четырьмя атомами Sn, расположенными в вершинах тетраэдра. Атомы олова образуют плотноупакованную (гранецентрированную кубическую) решетку, все тетраэдрические пустоты которой (вместо половины в алмазоподобных структурах) заняты атомами магния. Формально можно рассматривать связи у Sn как sp^3 -гибридные с удвоенным числом атомов Mg.

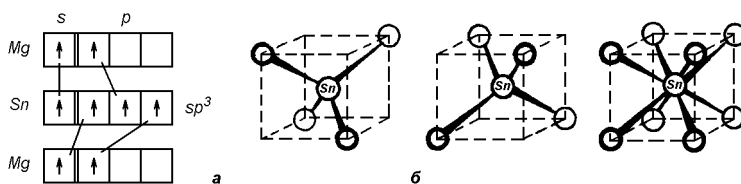


Рис. 2.25. *а* — Схема образования резонансных (кратность связи 1/3) ковалентных связей в Mg_2Sn ; *б* — локальное окружение атомов Sn в структуре антифлюорита и схемы виртуальных тетраэдрических связей.

Для описания таких структур было введено понятие о *резонансной или кратной* ковалентной связи. Используя все свои валентные электроны, каждый атом Sn образует четыре ковалентные sp^3 -гибридные связи, которые испытывают поворотный резонанс между восемью положениями. Таким образом, каждый атом Sn образует 8 «полусвязей» с соседними атомами Mg, заполняя тем самым его s- и p-оболочки, причем свою до конца (до 8 электронов), а оболочки атомов Mg — только наполовину (рис. 2.25,а). Поскольку каждый атом Mg использует свои два электрона для насыщения «полусвязей» с четырьмя соседними атомами Sn, то для образования связи Mg–Mg уже не остается электронов и, в результате, выполняются условия образования «полупроводниковой связи».

Таким образом, реальное кубическое окружение атома Sn можно представить как результат наложения двух виртуальных тетраэдрических, и при этом считается, что 4 электрона олова попеременно взаимодействуют с двумя группами атомов Mg (в каждой по четыре, см. рис. 2.25,б). Кратность связи показывает, что общее число валентных электронов на атом в молекуле меньше четырех, то есть число полных sp^3 -гибридных связей атома В меньше числа соседей А. Численно кратность связей определяется отношением числа полных ковалентных связей к числу ближайших соседей, то есть в случае Mg_2Sn равна 1/2 (4/8).

Аналогичные выводы о появлении резонансных связей можно сделать относительно соединений третьего типа, в которых число валентных электронов на атом также меньше четырех.¹⁹

Такой же подход к образованию химических связей применим и для

¹⁹В соединениях $\text{A}_2\text{B}^{\text{VI}}$ схема образования резонансных связей однотипна и аналогична схеме, характерной для соединений $\text{A}_2\text{B}^{\text{IV}}$, однако кристаллизуются они не только в структуре антифлюорита, что связано с индивидуальными особенностями образующих соединения элементов.

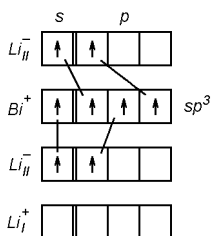


Рис. 2.26. Схема образования резонансных (кратность связи 1/2) ковалентных связей в Li_3Bi .

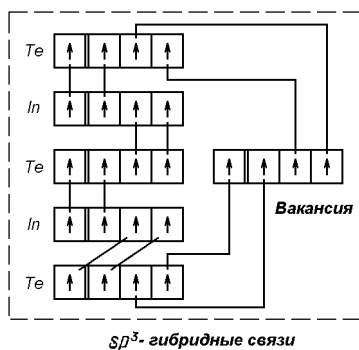
соединений второго типа, но с небольшими осложнениями. Типичный представитель соединений второго типа — Li_3Bi . В структуре этого соединения атомы висмута образуют кубическую плотную упаковку, а атомы Li размещаются в двух различных положениях в пустотах этой решетки. Две трети атомов лития (положение Li_{II}) занимают все тетраэдрические пустоты, образуя вместе с атомами Bi структуру типа антифлюорита, а оставшаяся треть атомов Li (Li_{I}) занимает октаэдрические пустоты в подрешетке Bi.

Схема образования ковалентных связей в соединении Li_3Bi показана на рис. 2.26. Вследствие перехода по одному электрону от атомов Bi и Li_{I} к атомам Li_{II} , каждый из последних имеет по два электрона и в антифлюоритной решетке $(\text{Li}_{\text{II}})_2\text{Bi}$ может наблюдаться резонанс того же типа, что и в Mg_2Sn . Этот пример показывает, как некоторые атомы, не принимая активного участия в образовании полупроводниковой связи, поставляют необходимые для ее реализации электроны. Таким образом, в Li_3Bi наблюдаются одновременно как электронные переходы, так и поворотный резонанс.

Соединения с дефектной структурой

Кроме полупроводников, структура которых является производной от структуры алмаза и которые образуются путем заполнения дополнительных мест в этой решетке, существуют полупроводники с *дефектными структурами* на основе структур сфалерита и вюртцита. Это соединения четвертого типа. Они образуются при замене атомов A из II подгруппы на атомы из III подгруппы и изменении их числа с трех на два для сохранения общего числа электронов в молекуле. С точки зрения правила нормальной валентности и правила Музера-Пирсона такое изменение роли не играет — в обоих случаях $n_e = 24$, $N_B = 3$, $N_{BB} = 0$. Однако число валентных электронов на атом становится больше 4 ($24/5 = 4.8$). Такая ситуация приводит к образованию *дефектной* тетраэдрической структу-

Рис. 2.27. Схема образования ковалентных связей в In_2Te_3 .



ры, в которой $1/3$ узлов компонента А оказывается вакантной.

Например, в In_2Te_3 (дефектная структура сфалерита) $1/3$ узлов в подрешетке In свободна; на каждый атом Te приходится 2.66 окружающих его атомов In ($4 \cdot 2/3 = 2.66$), тогда как каждый атом In окружен 4 атомами Te. Формальная схема образования sp^3 -гибридных связей в этом соединении приведена на рис. 2.27. Предполагается, что вакансия участвует в образовании ковалентной связи: между вакансией и атомами аниона (Te) образуются sp^3 -гибриды. Очевидно, что связи Te–In не идентичны связям Te–вакансия, что подтверждается экспериментальными данными. Если вакансии рассматривать как компонент соединения, то соединения этого типа будут подчиняться правилу 4 электрона на атом ($24/(5 + 1) = 4$).

Из-за высокой концентрации собственных дефектов ($5 \cdot 10^{21} \text{ см}^{-3}$) соединения этого типа имеют устойчивый тип проводимости, не меняющийся при введении в кристалл атомов компонентов в концентрациях выше стехиометрических; физические свойства этого типа соединений слабо зависят от примесей, а подвижности носителей заряда низки.

2.6.7. Полупроводниковые соединения других групп

Соединения $\text{A}^{\text{IV}}\text{B}^{\text{VI}}$

Из рассмотренных выше примеров может создаться неверное впечатление, что полупроводниковыми свойствами обладают только соединения с тетраэдрической или производной от нее координацией атомов. Соединения $\text{A}^{\text{IV}}\text{B}^{\text{VI}}$, к которым относятся халькогениды свинца, олова и германия, также обладают полупроводниковыми свойствами, хотя их структу-

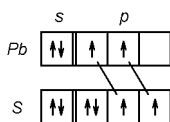


Рис. 2.28. Схема образования кратных (кратность связи 1/3) связей в PbS.

ра аналогична структуре каменной соли (NaCl) с обычным для нее координационным числом 6. Свойства этих соединений позволяют считать межатомные связи в них преимущественно ковалентными, но с заметной долей ионной составляющей связи.

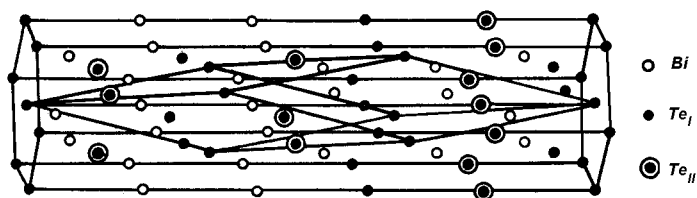
Соединения этого типа образуются элементами, симметрично расположенными относительно элементов V подгруппы, в них среднее число валентных электронов на атом равно 5. У обоих атомов *s*-электроны практически не принимают участия в образовании связей. Согласно представлениям Музера-Пирсона в этих соединениях реализуются две резонансные с кратностью 1/3 (2/6) чистые *p*-связи, а не sp^3 -гибридные связи. Об этом свидетельствует и кристаллическая структура рассматриваемых соединений. Схема образования *p*-связей приведена на рис. 2.28. Пустая орбиталь каждого атома Pb допускает поворотный резонанс двух *p*-связей между шестью положениями, так что образуется шесть 1/3-связей. Для заполнения *s*- и *p*-орбиталей валентной оболочки у каждого атома S достаточно двух *p*-электронов свинца, и условия для образования «полупроводниковой связи» выполняются.

Таким образом, первая и третья из эмпирических закономерностей, характерных для полупроводников, для соединений этого типа выполняются. Кроме того, к соединениям $A^{IV}B^{VI}$ формально можно применить правило Музера-Пирсона, если считать, что оба атома соединения являются компонентами В, тогда $n_e = 4$, $N_B = 2$, $N_{BB} = 6$. Однако правило нормальных валентностей для этого типа соединений не выполняется. Иное описание ковалентного характера связей в PbS было дано Шоттки и Кребсом.

Соединения $A_2^V B_3^{VI}$

К этой группе полупроводниковых соединений относятся халькогениды мышьяка, сурьмы и висмута, кристаллизующиеся в структуры с моноклинной, ромбической и ромбоэдрической решеткой (типа тетрадимита Bi_2Te_3), соответственно.

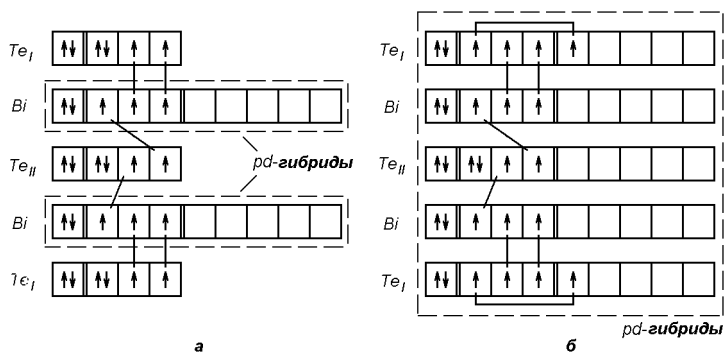
Рассмотрим вопрос об образовании химических связей в этой группе соединений на примере соединения Bi_2Te_3 с решеткой типа тетра-

Рис. 2.29. Структура соединения Bi_2Te_3 .

димита (такие полупроводники находят широкое применение в качестве термоэлементов). Структура Bi_2Te_3 образуется путем последовательного чередования пятислойных пакетов, расположенных перпендикулярно оси c [0001]. Каждый слой пакета уложен атомами одного сорта (рис. 2.29). Слои в структуре Bi_2Te_3 чередуются следующим образом: $-\text{Te}_I-\text{Bi}-\text{Te}_{II}-\text{Bi}-\text{Te}_I-\text{Te}_I-\text{Bi}-\text{Te}_{II}-\text{Bi}-\text{Te}_I-$. Слои смещены относительно друг друга так, что каждый атом внутри пакета имеет несколько искаженное октаэдрическое окружение. Очевидно, что связи с ближайшим окружением слоев Te_I и Te_{II} различаются. Координационное число для атомов, расположенных внутри пакета (Bi и Te_{II}), равно 6; атомы, расположенные в периферийных слоях пакета (Te_I), также имеют октаэдрическое окружение: три атома висмута из своего пакета и на несколько больших расстояниях три атома Te_I соседнего пакета.

Одна из схем образования связей предполагает, что внутри пятислойного пакета существуют pd -гибридные связи, которые являются резонансными. Они испытывают поворотный резонанс между шестью положениями в случае $-\text{Te}_{II}-\text{Bi}-$ (кратность связи равна $1/3$) и между тремя положениями в случае $-\text{Te}_I-\text{Bi}-$ (кратность связи равна $2/3$); для этого используются $6d$ -орбитали атомов висмута (рис. 2.30). Далее было обнаружено, что расстояние между атомами Te_I соседних пакетов ($-\text{Te}_I-\text{Te}_I-$) оказывается меньше, чем это вытекает из предположения о связи только с помощью ван-дер-ваальсовых сил. Схема, приведенная на рис. 2.30 показывает возможность возникновения ковалентной связи между пакетами: происходит гибридизация $5p$ - и $5d$ -орбиталей атомов Te_I , один из p -электронов переходит на d -орбиталь и связь между пакетами осуществляется за счет резонирующих гибридных pd -связей; кратность связи $\approx 2/3$. Таким образом, химические связи в Bi_2Te_3 образуются в результате суммирования двух предложенных схем образования ковалентных связей.

Рассмотренные схемы предполагают, что связь между атомами одного

Рис. 2.30. Схемы образования химических связей в соединении Bi_2Te_3 .

слоя отсутствует, однако существуют и другие представления. В частности, считают, что между атомами одного слоя возникают связи, причем в слоях Bi значительна доля металлической составляющей связи.

Следует отметить, что для двух рассмотренных выше схем образования ковалентной связи формально выполняется правило Музера-Пирсона, атом аниона эффективно достраивает свою внешнюю оболочку до восьми электронов. Действительно, по первой из рассмотренных схем $n_e = 12$, $N_B = 3$, $N_{BB} = 0$, но у каждого атома Te имеется по две неподеленные пары электронов, поэтому эффективно на каждый анион приходится восемь электронов. По второй схеме $n_e = 16$; поведение Te_I и Te_{II} при образовании связей различно: у Te_I одна неподеленная пара электронов, 4 валентных электрона участвуют в образовании связей и $N_{BB} = 2$; у Te_{II} две неподеленные пары валентных электронов и два валентных электрона участвуют в образовании связей ($N_{BB} = 0$). Следовательно, число валентных электронов, эффективно приходящихся на один атом, также равно 8, как и по первой схеме.

В заключении следует отметить, что помимо рассмотренных бинарных полупроводников существуют и другие, для которых могут не выполняться и правило Музера-Пирсона, и правило нормальных валентностей, например, полупроводниковые соединения $\text{A}^{\text{III}}\text{B}^{\text{VI}}$ (GaSe), $\text{A}^{\text{II}}\text{B}^{\text{V}}$ (Mg_3Sb_2).

2.6.8. Тройные полупроводниковые соединения

Тройные соединения и твердые растворы на их основе являются поистине неисчерпаемым источником полупроводниковых материалов с широ-

ким спектром физико-химических свойств. Химические связи в тройных полупроводниковых фазах, также как и в двойных, носят смешанный характер — ковалентно-ионный-металлический. Специфика связей в этих материалах обусловлена наличием атомов трех сортов.

Закономерности образования тройных полупроводниковых соединений

Тройные полупроводники могут быть разделены на одноанионные (двухкатионные) и двуханионные (однокатионные). Большинство тройных полупроводниковых одноанионных соединений подчиняется правилу Музера-Пирсона: анион эффективно достраивает валентную оболочку до восьми электронов. Однако для применения этого правила на практике необходимо знать стехиометрическую формулу и кристаллическую структуру анализируемой фазы для точного определения величин n_e , N_B и N_{BV} . Оценить, является ли химическая связь «полупроводниковой» для двуханионных тройных соединений, более сложно, чем для одноанионных.

Эмпирически установлены следующие закономерности образования тройных полупроводников — аналогов известных двойных.

Первое правило. *В тройной фазе суммарное число валентных электронов должно быть таким же, как и в бинарной полупроводниковой фазе-аналоге.*

Второе правило. *Тип химической связи в тройной фазе должен быть таким же, как и в бинарной полупроводниковой фазе-аналоге.*²⁰

Третье правило. *Разности электроотрицательностей компонент соединения не должны быть велики.*

Четвертое правило. Большинство тройных полупроводниковых соединений подчиняется правилу нормальной валентности. Для одноанионных соединений с общей формулой $A_x^K B_y^L C_z^M$ оно записывается так:

$$Kx + Ly = (8 - M)z, \quad (2.12)$$

где K , L и M — число валентных электронов, которые отдает компонент А, В, С, соответственно, соединения на образование химических связей в нем.

Для двуханионных тройных полупроводников:

$$Kx = (8 - L)y + (8 - M)z. \quad (2.13)$$

²⁰ Первые два правила, по-существу, близки к правилу Музера-Пирсона.

Эти правила выполняются для большинства полупроводников и основаны на том, что элементы одной группы периодической системы обладают похожими физико-химическими свойствами, что обусловлено одинаковым числом валентных электронов у них.

Тройные полупроводники с алмазоподобной структурой

Тройные алмазоподобные полупроводники образуются при возникновении sp^3 -гибридных связей и характеризуются тетраэдрическим расположением атомов в пространстве. Для них кроме перечисленных выше закономерностей должно выполняться дополнительное правило: *четыре электрона на атом*

$$\frac{Kx + Ly + Mz}{x + y + z} = 4. \quad (2.14)$$

Пользуясь всеми этими правилами, можно получить следующие тройные соединения-аналоги, например, двойных соединений типа $A^{III}B^V$ (поперечное замещение): $2A^{III}B^V \rightarrow A^{II}A^{IV}B_2^V$, или для $A^{II}B^{VI}$: $2A^{II}B^{VI} \rightarrow A^IA^{III}B_2^{VI}$.

В рассматриваемом ниже примере замещение производится в катионной подрешетке.

Ga		As	$A^{III}B^V$
Zn	Ge	As ₂	$A^{II}B^{IV}C_2^V$
Cu	Ga	Ge ₂ As ₄	$A^IB^{III}C_2^{IV}D_4^V$

В следующей таблице приведены формулы возможных тройных алмазоподобных соединений производных от A^NB^{8-N} .

Одноанионные		Двуханионные	
тройные	двойные	тройные	двойные
$A^{II}B^{IV}C_2^V$	$A^{III}B^V$	$A_2^{II}B^VC^{VII}$	A^IB^{VII}
$A^IB^{III}C_2^{VI}$	$A^{II}B^{VI}$	$A_3^{III}B_2^{IV}C^{VII}$	A^IB^{VII}
$A^IB_2^{IV}C_3^V$	$A^{III}B^V$	$A_3^{II}B^{IV}C_2^{VII}$	A^IB^{VII}
$A_2^IB^{IV}C_3^{VI}$	$A^{II}B^{VI}$	$A_4^{III}B^{III}C_3^{VII}$	A^IB^{VII}
$A_3^IB^VC^{VI}$	$A^{II}B^{VI}$	$A_2^{III}B^{IV}C^{VI}$	$A^{II}B^{VI}$

Таким образом, возможно существование десяти классов недефектных тройных алмазоподобных соединений (по сравнению с тремя классами недефектных двойных алмазоподобных соединений), причем число соединений в одном классе тройных в несколько раз превышает число соединений в одном классе двойных. Большая группа сложных алмазоподобных полупроводниковых соединений описана в [26].

Глава 3

Дефекты в полупроводниковых материалах

Все природные и синтезированные монокристаллы и в еще большей степени кристаллиты поликристаллов отличаются от идеальных тем, что содержат различные нарушения структуры кристалла. Нарушения идеальной трансляционной симметрии кристалла называются *структурными дефектами*. Дефекты оказывают существенное влияние на многие параметры твердых тел. К таким параметрам относятся электропроводность, фотопроводимость, теплопроводность, скорость диффузии, магнетизм, твердость, прочность и пластичность, плотность и т. д. Зависимость этих параметров твердого тела от дефектов может оказаться настолько велика, что в итоге они будут определяться не столько исходной структурой материала, сколько типом и числом дефектов в нем. Параметров, не чувствительных к структурным дефектам, строго говоря, нет, но практически такие параметры, как температура плавления, диэлектрическая проницаемость, парамагнитные и диамагнитные характеристики, упругие модули, можно отнести к параметрам, менее чувствительным к дефектам.

Дефекты могут быть разделены на две основные группы: *собственные дефекты* и *примеси* (рис. 3.1). Если за основу классификации принять размеры и протяженность областей решетки кристалла, на которые распространяется действие дефектов, то все дефекты могут быть разделены на:

1. *Точечные*: вакансии, междоузельные атомы основного вещества, примесные атомы в узлах и междоузлиях, антиструктурные дефекты, комплексы из простых точечных дефектов.

2. *Линейные дефекты*: дислокации.

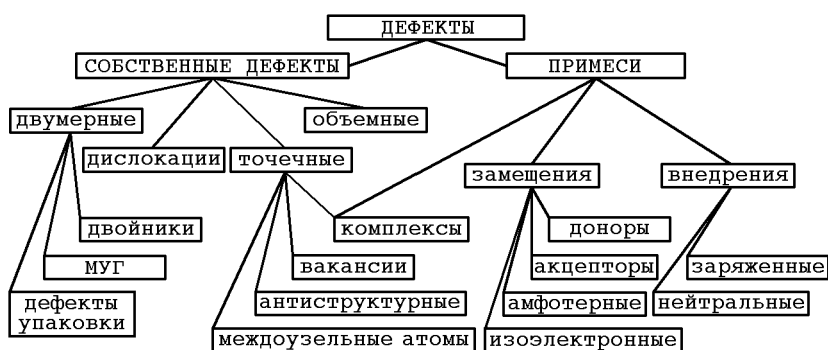


Рис. 3.1. Дефекты в кристаллических полупроводниках.

3. *Двумерные дефекты*: малоугловые границы, границы зерен и блоков, двойниковые границы, дефекты упаковки.

4. *Объемные нарушения*: зональные напряжения, поры, включения второй фазы.

Основное отличие точечных дефектов от линейных, двумерных и объемных дефектов состоит в том, что они могут существовать в кристалле как в термодинамически равновесном, так и в метастабильном состоянии при конечной температуре. Линейные, двумерные и объемные дефекты являются метастабильными образованиями, возникающими при росте, механической деформации или при термической обработке кристалла. Таким образом, теоретически можно получить кристалл, содержащий только точечные дефекты.

Дефекты могут перемещаться по кристаллу, взаимодействовать друг с другом, с образованием новых, более сложных или, наоборот, более простых, чем исходные, дефектов под влиянием внешних воздействий. Эти явления приводят к изменению с течением времени свойств материала, из которого изготавливается прибор, что, естественно, ведет к изменению выходных параметров прибора в процессе его работы. Изменение параметров прибора со временем называется *деградацией* или «старением» прибора. При получении материалов для изготовления полупроводниковых приборов учет этого эффекта и сведение его к минимально возможному чрезвычайно важно.

Каждый тип дефектов по-своему влияет на структурно-чувствительные параметры материала. При наличии в кристалле большого числа различных дефектов невозможно получить материал с контролируемыми параметрами. В связи с этим перед материаловедением и технологией

стоит задача получения структурно совершенных и чистых материалов или материалов с определенным типом и концентрацией дефектов, то есть таких материалов, параметрами которых можно было бы управлять. Поэтому выяснение причин возникновения различных структурных дефектов, изучение их поведения в кристалле, их влияния на те или иные параметры материала, разработка методов управления концентрацией дефектов являются важнейшими задачами материаловедения и технологии.

Рассмотрим сначала собственные дефекты, а затем примеси.

3.1. Собственные дефекты

3.1.1. Точечные дефекты

Точечный дефект — это нарушение кристаллической структуры, размеры которого во всех трех измерениях сравнимы с одним или несколькими межатомными расстояниями. Точечный дефект может иметь простую или сложную структуру.

Вакансия — это простейший структурный дефект, представляющий собой свободный узел решетки, который должен быть занят атомом или ионом в совершенном кристалле. Вакансии обозначаются V_A , V_B , ... , где индексы A, B ... указывают тип отсутствующего атома. Например, в GaAs V_{Ga} и V_{As} — вакансии Ga и As, соответственно. Образование вакансий сопровождается упругой деформацией решетки. Соседние с вакансией атомы сближаются, смещаясь в направлении центра пустого узла.

Междоузельным называют атом (ион), расположенный в междоузлии (простой дефект). A_i означает атом A в междоузлии. Вероятность размещения атома в междоузлии определяется прежде всего относительным объемом междоузлия (пустоты) и размером междоузельного атома. Наиболее существенную роль в типичных полупроводниках со структурами типа алмаза, сфалерита и вюртцита играют октаэдрические и тетраэдрические пустоты, находящиеся между атомами, центры которых расположены соответственно в вершинах октаэдра и тетраэдра.

Вакансии, возникающие за счет ухода атома из узла на поверхность кристалла или на какую-либо границу внутри кристалла, называют *дефектами Шоттки*.¹

Парный дефект (вакансия + междоузельный атом), возникающий за счет перехода атома из узла решетки в междоузлие, называют *дефектом Френкеля*.

¹В некоторых работах дефектом Шоттки называют пару из катионной и анионной вакансий. В этом случае дефект Шоттки уже парный дефект, то есть сложный [27].

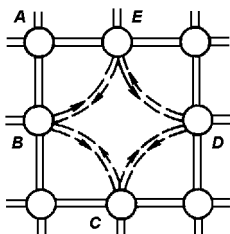


Рис. 3.2. Схематическое изображение образования «акцепторной связи» вокруг вакансии в двумерной решетке.

Точечные дефекты в кристаллах могут быть как электрически неактивными, так и электрически активными. В частности, вакансии и междоузельные атомы в полупроводниках могут находиться как в электрически неактивном, так и в электрически активном состоянии (одно- и многозарядном). Рассмотрим, например, процесс образования электрически активной вакансии. При образовании вакансий в ковалентном кристалле возникают оборванные связи. Перераспределение ковалентных связей у атомов, окружающих вакансию, схематически показано на рис. 3.2. Можно предположить, что неспаренный электрон атома E взаимодействует с неспаренными электронами атомов B и D. В результате один валентный электрон осуществляет связь не между двумя, а между тремя атомами. Такая связь оказывается нестабильной и может вести себя как акцепторная. Процесс ионизации вакансии при этом описывается схемой $V_A \rightleftharpoons V_A^- + h^+$. Аналогично для процесса ионизации междоузельного атома — $A_i \rightleftharpoons A_i^+ + e^-$.

Таким образом, в элементарных полупроводниках, не содержащих примесей, могут существовать четыре вида простых точечных дефектов, а именно: электрически неактивные вакансии и междоузельные атомы основного вещества, заряженные вакансии и междоузельные атомы основного вещества.

Образование точечных дефектов всех типов требует затраты энергии. Энергия образования одной вакансии или междоузельного атома, как правило, имеет порядок нескольких электронвольт, но энергия образования междоузельного атома немного (в 1.5–2 раза) больше, чем энергия образования вакансии. Эта энергия тратится на разрыв связей при образовании точечного дефекта и на искажение решетки вокруг дефекта, вызванное смещением атомов из равновесных положений. Рассматривая кристалл как сплошную упругую среду, можно показать, что упругие напряжения вокруг точечного дефекта убывают пропорционально $1/r^3$, где r — расстояние от дефекта. Энергия образования точечного дефекта

может быть снижена за счет объединения простых точечных дефектов в группы или *комплексы*. Например, процесс взаимодействия между вакансией и междоузельным атомом, ведущий к их взаимной аннигиляции, наиболее энергетически выгоден. Но и объединение двух вакансий в одну *дивакансию* также энергетически выгодно, так как при этом уменьшается число оборванных связей.

К аналогичному выводу приводит анализ возможности образования комплекса, содержащего вакансию и атом примеси. В случае электрически неактивных дефектов образование такого комплекса уменьшает энергию упругих искажений решетки и поэтому энергетически выгодно. В случае заряженных примеси и вакансии дополнительный выигрыш энергии возникает еще и из-за кулоновского взаимодействия между заряженной вакансией и ионом примеси.

Вероятность образования комплексов (сложных дефектов) из простых точечных дефектов тем выше, чем больше их энергия связи (то есть связанный с комплексообразованием выигрыш в энергии) и чем выше концентрация одиночных дефектов (то есть чем больше вероятность их встречи). Свойства сложных точечных дефектов, как правило, отличны от свойств образующих их простых дефектов.

Простые точечные дефекты в химических соединениях более разнообразны, чем в элементарных веществах.

Помимо вакансий и междоузельных атомов, в соединениях важную роль играют так называемые *антиструктурные дефекты* — дефекты, которые образуются при взаимном обмене местами атомов элементов, образующих соединение. Такие дефекты наиболее часто возникают в соединениях, в которых размеры и электроотрицательность образующих их атомов близки, то есть роль ионной составляющей связи невелика. Антиструктурное разупорядочение наблюдается, например, в Bi_2Te_3 . В ионных кристаллах антиструктурные дефекты практически не встречаются.

Особенностью соединений является также образование точечных дефектов при отклонении состава от стехиометрического в области гомогенности (см. гл. 4). Отклонение от стехиометрии приводит к образованию вакансий в большинстве полупроводников (образуются твердые растворы вычитания), но могут образовываться и междоузельные атомы. Например, в GaAs избыток Ga создает вакансии As , а избыток As приводит к появлению вакансий Ga и, кроме того, междоузельных атомов As . Однако в этом случае вакансию и междоузельный атом нельзя объединить в один парный дефект, поскольку вакансии Ga и междоузельный атом As не могут аннигилировать.

Следует иметь в виду тот факт, что при образовании дефектов кристалл в целом остается электронейтральным. Выполнение этого условия обеспечивается образованием равного количества положительно и отрицательно заряженных дефектов, образованием электрически неактивных сложных дефектов или же образованием свободных электронов или дырок. Нейтрализация дефектов решетки с помощью электронов и дырок тем более вероятна, чем больше их в разрешенных зонах, то есть чем уже запрещенная зона. В кристаллах с широкой запрещенной зоной более вероятна нейтрализация точечных дефектов путем образования равного количества противоположно заряженных дефектов (вторичные дефекты). Например, в некоторых кристаллах избыток анионов B^- приводит к тому, что катионные вакансии становятся донорами A^+ , а избыток катионов — к тому, что анионные вакансии становятся акцепторами. При этом образуется такое количество вторичных дефектов, которое необходимо для обеспечения электронейтральности кристалла.

Точечные дефекты могут перемещаться по кристаллу путем диффузии, взаимодействовать друг с другом и с другими дефектами (см. ниже).

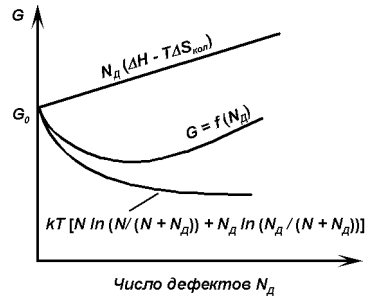
Основные механизмы образования точечных дефектов

1. Тепловые колебания атомов. Любой кристалл, даже самый совершенный, можно представить идеальным (бездефектным) только при абсолютном нуле температуры. При $T > 0$ К тепловое движение атомов приводит к возникновению точечных дефектов — вакансий, междоузельных атомов, дефектов Френкеля и т.д. При любой отличной от нуля температуре имеется некоторая равновесная концентрация дефектов, при которой полная энергия кристалла $G = H - TS$ минимальна (здесь G — потенциал Гиббса, H — внутренняя энергия или энтальпия, S — энтропия). Действительно, рассмотрим, например, образование вакансий в идеальном кристалле при $T > 0$. Пусть кристалл содержит N узлов в единице объема, в каждом из которых при 0 К расположен атом, и потенциал Гиббса системы в этом состоянии равен G_0 . Пусть при повышении температуры за счет тепловых флуктуаций возникает N_d дефектов Шоттки в единице объема кристалла и теперь N атомов и N_d вакансий размещается по $(N + N_d)$ узлам. Образование N_d вакансий вызывает повышение G на

$$\Delta G = G - G_0 = \Delta H - T\Delta S = N_d W - T(N_d \Delta S_{\text{кол}} + \Delta S_{\text{конф}}), \quad (3.1)$$

где W — энергия образования одной вакансии в кристалле; $\Delta S_{\text{кол}}$ — изменение энтропии, связанное с отклонением от упорядоченного расположения атомов, вызванным тепловыми колебаниями атомов около узлов

Рис. 3.3. Зависимость полной энергии кристалла от концентрации вакансий.



решетки; $\Delta S_{\text{конф}}$ — изменение энтропии, связанное с распределением атомов и дефектов по узлам решетки. Величины W и $\Delta S_{\text{кол}}$ всегда положительны,² но, как правило, $W > \Delta S_{\text{кол}}$ в 6–8 раз.

Изменение $\Delta S_{\text{конф}}$ может быть записано в виде

$$\Delta S_{\text{конф}} = k \ln(P_1/P) = k \ln \frac{(N + N_d)!}{N!N_d!}, \quad (3.2)$$

где P_1 и P — число способов, которым может быть осуществлено размещение N атомов и N_d вакансий по $(N + N_d)$ узлам и размещение N атомов по N узлам ($P = 1$) кристаллической решетки в единице объема кристалла, соответственно. Используя разложение Стирлинга,³ получим

$$\begin{aligned} \Delta S_{\text{конф}} &= k \ln \frac{(N + N_d)!}{N!N_d!} = -k \left[N \ln \frac{N}{N + N_d} + N_d \ln \frac{N_d}{N + N_d} \right], \\ \Delta G &= N_d(W - T\Delta S_{\text{кол}}) + kT \left[N \ln \frac{N}{N + N_d} + N_d \ln \frac{N_d}{N + N_d} \right]. \end{aligned} \quad (3.3)$$

Таким образом, из рис. 3.3 видно, что полная энергия G кристалла уменьшается при возникновении в кристалле вакансий, потом достигает минимума при некоторой их равновесной концентрации и затем снова возрастает [17].

²Энтропия колебаний кристалла, содержащего N атомов ($3N$ осцилляторов) равна $S_{\text{кол}} = 3R_0 [\ln(kT/h\nu) + 1]$, где ν — частота колебаний атомов, R_0 — универсальная газовая постоянная, k — постоянная Больцмана. Если атомы изменяют частоту колебаний от ν до ν_1 , то $\Delta S_{\text{кол}} = 3R_0 \ln(\nu/\nu_1)$. Отсюда $\Delta S_{\text{кол}} > 0$, если $\nu_1 < \nu$ и $\Delta S_{\text{кол}} < 0$, если $\nu_1 > \nu$. Наличие вакансии (любого дефекта) приводит к уменьшению частоты колебаний атомов, находящихся вблизи дефекта, а значит $\Delta S_{\text{кол}} > 0$ и ΔG уменьшается.

³Для больших значений N_d формула Стирлинга, используемая для вычисления энтропии системы с дефектами, может быть записана в виде $\ln N_d! \approx N_d (\ln N_d - 1)$.

Равновесная концентрация дефектов Шоттки $N_{\text{д.равн}}$ может быть определена из условия

$$\partial G / \partial N_{\text{д}} = W - T \Delta S_{\text{кол}} + kT \ln [N_{\text{д.равн}} / (N + N_{\text{д.равн}})] = 0 \quad (3.4)$$

Отсюда с учетом того, что $N \gg N_{\text{д.равн}}$ и $\Delta S_{\text{кол}}$ мало, получим

$$N_{\text{д.равн}} = N \exp(-W/kT). \quad (3.5)$$

Аналогичным образом может быть получена и формула для температурной зависимости концентрации дефектов Френкеля с учетом парного характера взаимодействия:

$$N_{\text{д.равн}} = (NN_i)^{1/2} \exp(-W/2kT), \quad (3.6)$$

где $N_{\text{д.равн}}$ — равновесная концентрация дефектов, N — число атомов в узлах в единице объема, N_i — число междоузельных позиций в единице объема, W — энергия образования дефекта в кристалле, k — постоянная Больцмана. Множитель $1/2$ в показателе экспоненты появился вследствие того, что в кристалле одновременно возникает два дефекта: атом в междоузлии и вакансия.

Следовательно, при любой конечной температуре кристалл содержит равновесные вакансии и междоузельные атомы, причем их концентрация определяется энергией дефектообразования и температурой кристалла (является экспоненциальной функцией температуры).

Повышенная, по сравнению с равновесной, концентрация дефектов при низких температурах может получаться в результате закалки (быстрого охлаждения) кристалла. Закалка может резко изменить механические и электрические свойства материала. Экспериментальные данные показывают, что резкое охлаждение материала от высоких температур до комнатной сохраняет некоторую часть избыточных тепловых вакансий. Чем больше скорость охлаждения и чем совершеннее кристалл, тем большая часть вакансий сохраняется в метастабильном состоянии. При сильно неравновесных условиях охлаждения возможно пересыщение кристалла вакансиями, тогда они могут объединяться и образовывать поры, перерастающие иногда в «отрицательные» кристаллы.⁴

2. Пластическая деформация. Вакансии и междоузельные атомы могут возникать и исчезать в процессе движения и пересечения дислокаций, образующихся при пластической деформации кристалла (см. ниже).

⁴Отрицательный кристалл — это пора в кристалле, имеющая ту же форму, что и сам кристалл.

3. Облучение. При облучении материала частицами высокой энергии (например, быстрыми электронами с энергией ≈ 100 кэВ и выше) также возникают точечные дефекты. Такие дефекты называют *радиационными дефектами* [28]. Однако пороговая энергия образования этих дефектов составляет от 6 до 30 эВ, что выше энергии образования тепловых дефектов. Например, энергия образования радиационного дефекта Френкеля в кремнии равна 20.4 эВ при комнатной температуре. Это, по-видимому, обусловлено сложными лавинными процессами генерации дефектов первично выбитыми атомами.

4. Отклонения от стехиометрии. Такие отклонения приводят, как уже говорилось выше, к образованию вакансий в большинстве полупроводниковых соединений. Стехиометрические вакансии, как и тепловые, являются равновесными. Но если концентрация тепловых вакансий зависит от температуры, то концентрация стехиометрических в пределах области гомогенности (однородности) зависит только от состава соединения.

Таким образом, первый и четвертый механизмы образования точечных дефектов приводят к возникновению равновесных дефектов, а второй и третий — неравновесных, возникающих в процессе роста кристалла из-за несовершенства технологии, из-за его пластической деформации, при термообработке или в результате радиационного воздействия.

О концентрации и состоянии точечных дефектов в кристалле можно судить по изменению плотности кристалла или параметров решетки, по проводимости (в случае электрически активных дефектов), по спектрам поглощения света и т. д.

3.1.2. Дислокации

Дислокации — линейные дефекты кристаллической решетки типа обрыва или сдвига атомных плоскостей, нарушающие правильность их чередования. Энергия образования дислокаций существенно выше энергии образования точечных дефектов, поэтому они не могут существовать в измеримых концентрациях как термодинамически устойчивые дефекты. Они легко образуются при выращивании монокристаллов или эпитаксиальных пленок, сопровождающемся термическими, механическими и концентрационными напряжениями, приводящими к пластической деформации кристалла. Часть дислокаций может сохраняться в кристалле даже после самого тщательного отжига. Более подробно вопрос о причинах возникновения дислокаций будет рассмотрен при обсуждении методов выращивания монокристаллов и эпитаксиальных пленок (см. гл. 6

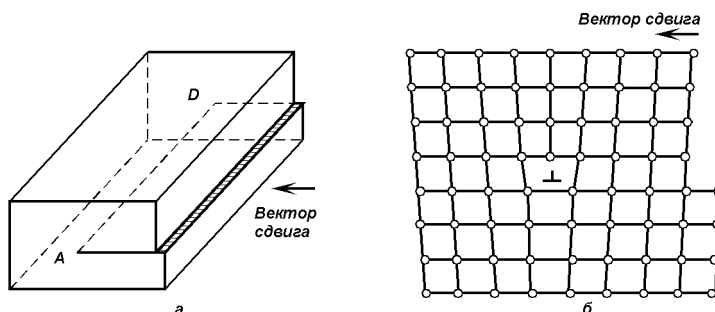


Рис. 3.4. *а* — Схема образования краевой дислокации. *б* — Расположение атомов вокруг краевой дислокации в простой кубической решетке.

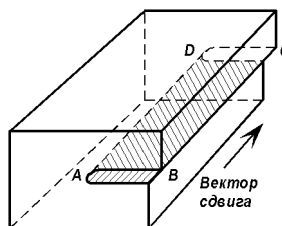
и гл. 9), так как он имеет огромное значение для технологии полупроводников.

Поскольку дислокации при небольших концентрациях распределены по объему кристалла неравномерно, то нарушения свойств кристалла, обусловленные их присутствием, локализованы в небольших объемах, окружающих дислокации. Следовательно, кинетика процессов, протекающих вблизи этих дефектов, будет иной, чем в объеме кристалла, где дислокации отсутствуют.

Различают простые и смешанные дислокации. К простым дислокациям относят *краевые* и *винтовые*.

Рассмотрим кристалл, в котором в результате приложения внешних сил одна часть начинает смещаться относительно другой по некоторой плоскости (плоскости скольжения) вдоль некоторого выбранного направления (направления сдвига), то есть кристалл подвергается пластической деформации сдвига путем скольжения (рис. 3.4,*а*). Перемещения атомов по плоскости скольжения не являются равномерными, так как сдвиг на единицу скольжения происходит только в одной части кристалла, а в остальной части сдвиг полностью отсутствует. Перпендикулярную направлению сдвига линию AD , расположенную внутри кристалла и являющуюся границей между этими двумя частями кристалла, называют *линией краевой дислокации* и обозначают $\perp$. Расположение атомов внутри деформированного скольжением кристалла в плоскости, перпендикулярной линии AD , имеет вид, изображенный на рис. 3.4,*б*. Атом, расположенный непосредственно над знаком дислокации $\perp$, имеет окружение, отличающееся от окружения атомов в бездефектной части кристалла. Все происходит так, как если бы в объеме кристалла появилась

Рис. 3.5. Схема образования винтовой дислокации.



одна «лишняя» атомная плоскость, край которой обрывается внутри кристалла. *Краевая дислокация*, таким образом, представляет собой дефект, при котором одна из атомных плоскостей обрывается внутри кристалла по линии краевой дислокации AD , перпендикулярной направлению сдвига. Краевые дислокации образуются вдоль плоскостей скольжения. Условно подразделяют краевые дислокации на положительные и отрицательные. Положительная дислокация соответствует случаю, когда «лишняя» атомная плоскость находится сверху над знаком $\perp$ (рис. 3.4,б) и в верхней половине кристалла действуют сжимающие напряжения, в нижней — растягивающие; а отрицательная — случаю, когда «лишняя» атомная плоскость находится снизу под знаком $\perp$, то есть верхняя половина кристалла растянута, а нижняя сжата.

Сдвиг одной части кристалла относительно другой, возникающий под влиянием внешних воздействий, может деформировать кристалл таким образом, что его можно представить состоящим из атомных плоскостей, закрученных в виде винтовой лестницы, ось которой и образует *линию винтовой дислокации* AD (рис. 3.5 и рис. 3.6). Винтовая дислокация обозначается $\otimes$. Линия винтовой дислокации характеризуется тем, что она параллельна направлению сдвига. При каждом обходе вокруг нее мы поднимаемся или опускаемся на одно межплоскостное расстояние (рис. 3.6). Выход винтовой дислокации на поверхность кристалла заканчивается незарастающей ступенькой. В отличие от краевой дислокации, винтовая дислокация не имеет «лишних» плоскостей и может образовываться при сдвиге по любой атомной плоскости, проходящей через линию дислокации AD , то есть она не определяет однозначно плоскость скольжения. Различают правые и левые винтовые дислокации, причем направление вращения играет ту же роль, что и знак у краевых дислокаций.

Следует отметить, что в кристаллах полупроводников в чистом виде краевая и винтовая дислокации встречаются редко. Как правило, они по-

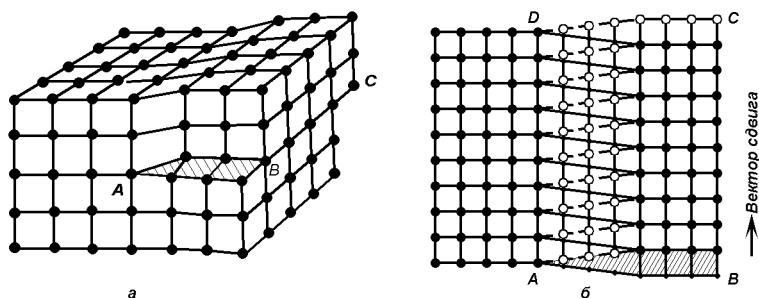


Рис. 3.6. *а* — Винтовая дислокация. *б* — Расположение атомов на винтовой поверхности вдоль винтовой дислокации.

являются одновременно, приводя к образованию смешанных форм, при этом линия результирующей дислокации (граница между сдвинутой и не сдвинутой частями кристалла) не обязательно должна быть прямой, она может представлять собой произвольную кривую. Криволинейную дислокацию произвольной формы, которую можно представить как состоящую из различных комбинаций «чистых» краевой и винтовой дислокаций, называют *смешанной*. На рис. 3.7 показана смешанная дислокация в простой кубической решетке.

Основной характеристикой дислокации является вектор Бюргерса (вектор сдвига) ***b***. *Вектор Бюргерса* — это мера искажений решетки, обусловленных присутствием дислокации. Для его определения строят замкнутый контур в кристалле с дефектом и контур, проходящий через те же атомы, в кристалле без дефекта. Проведем в решетке, содержащей краевую дислокацию, замкнутый контур *A-B-C-D-A* вокруг этой дислокации, начав его из произвольно взятого узла *A* и откладывая против часовой стрелки определенное число межатомных расстояний (рис. 3.8). Если построить тот же контур в решетке без дислокации, откладывая такое же число межатомных расстояний, то контур окажется незамкнутым. Вектор ***b***, который необходимо добавить, чтобы замкнуть контур, и есть вектор Бюргерса. Величина разрыва контура характеризует сумму всех упругих смещений решетки, накопившихся в области вокруг дислокации. Для примера, изображенного на рис. 3.8 (простая кубическая решетка), вектор Бюргерса по абсолютной величине равен расстоянию между соседними атомами и ориентирован перпендикулярно линии дислокации.

Аналогичное построение для винтовой дислокации в простой кубиче-

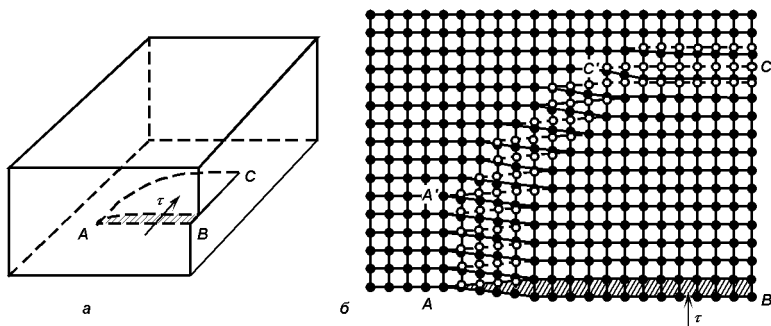


Рис. 3.7. *а* — Сдвиг, создавший смешанную дислокацию. *б* — Смешанная дислокация в простой кубической решетке. Краевая плоскость скольжения совпадает с плоскостью чертежа; черные кружки — атомы под плоскостью скольжения, белые — атомы над ней; заштрихованный участок — ступенька, образовавшаяся в результате сдвига.

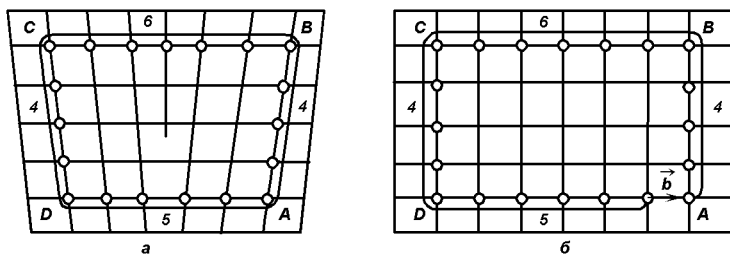


Рис. 3.8. Контур и вектор Бюргера краевой дислокации.

ской решетке показывает, что и в этом случае $\mathbf{b}$ равен расстоянию между соседними атомами (это шаг винта в направлении линии дислокации), но направлен вдоль линии дислокации.

Вектор Бюргерса — наиболее инвариантная характеристика дислокации. Он остается постоянным вдоль всей линии любой дислокации и сохраняется при ее движении [27]. Символ вектора Бюргерса показывает его направление и абсолютную величину. Общее обозначение дислокации в кубической решетке имеет вид $\frac{a}{n}\langle uvw \rangle$, где $\langle uvw \rangle$ — символ кристаллографического направления вектора Бюргерса, a — параметр решетки, а $\frac{a}{n}(u^2 + v^2 + w^2)^{1/2}$ — его абсолютная величина.⁵ Так, дислокация, вектор Бюргерса которой направлен по $\langle 110 \rangle$ и имеет длину $\frac{a\sqrt{2}}{2}$, описывается символом $\frac{a}{2}\langle 110 \rangle$.

Теперь можно дать определение простых дислокаций через вектор Бюргерса. *Краевой дислокацией* называют дислокацию, вектор Бюргерса $\mathbf{b}$ которой перпендикулярен линии краевой дислокации. *Винтовой дислокацией* называют дислокацию, вектор Бюргерса $\mathbf{b}$ которой параллелен линии винтовой дислокации. В общем случае смешанной дислокации вектор Бюргерса может иметь иные направления относительно линии дислокации.

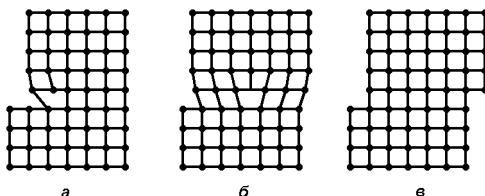
Дислокации окружены полями упругих напряжений. Область над линией краевой дислокации, содержащая лишнюю полуплоскость, испытывает напряжения сжатия, область под линией дислокации — напряжения растяжения. Вокруг винтовых дислокаций существует поле сдвиговых (касательных) напряжений. Величина напряжений убывает обратно пропорционально расстоянию от линии дислокации [27].

Свойства дислокаций

1. Дислокации не могут обрываться внутри кристалла и должны либо замыкаться сами на себя, либо разветвляться на другие дислокации, либо выходить на поверхность кристалла. Это свойство является следствием того, что линия дислокации представляет собой кривую, вдоль которой вектор Бюргерса $\mathbf{b}$ остается постоянным. Для узла разветвляющихся дислокаций справедлива теорема, аналогичная теореме Кирхгофа для разветвляющихся линий токов: если все дислокации соединяются в одной точке пересечения, то сумма их векторов Бюргерса должна быть равна нулю, например, $\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 = 0$. Разветвляясь и снова сливаясь, дислокации образуют в кристалле плоские и пространственные сетки, определяющие мозаичную структуру кристалла.

⁵Обозначение векторов Бюргерса в других решетках см. в [27].

Рис. 3.9. Схема пластического скольжения, осуществляемого путем перемещения краевой дислокации.



2. *Энергия дислокации* — это энергия, которая затрачивается на искажение решетки при образовании дислокации. Она складывается из упругой энергии искажений решетки и неупругой энергии ядра дислокации.⁶ Оценка упругой энергии, приходящейся на одно межатомное расстояние вдоль дислокации в структуре типа алмаза, показала, что $E_{\text{уп}} \sim Gb^2$, где G — модуль сдвига, и составляет 4–5 эВ [29]. Если учесть, что реальная длина дислокаций превышает десятки межатомных расстояний, и учесть неупругую энергию ядра дислокации ($E_{\text{неуп}} \approx 0.1E_{\text{уп}}$), то ясно, что даже при температурах, близких к температуре плавления, дислокации не могут образовываться в результате тепловых флуктуаций. Таким образом, дислокации не являются равновесными дефектами.

3. Под действием внешних сил дислокации могут перемещаться скольжением и переползанием, при этом дислокации передвигаются как вдоль плоскостей скольжения,⁷ так и перпендикулярно им соответственно. Движение дислокаций связано с пластической деформацией кристалла. Экспериментальные данные показывают, что дислокации могут двигаться со скоростями от 10^{-9} м/с до 10^{-3} м/с в зависимости от материала и приложенного напряжения. Однако скорость дислокации в кристалле не может быть больше скорости звука, так как перемещение дислокации есть перемещение волны упругой деформации.

Перемещение дислокаций в плоскости скольжения называется *скольжением*. Такое движение дислокаций можно в определенном смысле представить как следствие пластической деформации кристалла, связанное с коллективным перемещением атомов. В результате скольжения одной дислокации через весь кристалл происходит пластический сдвиг на одно межатомное расстояние. Однако в каждый данный момент сдвиг происходит не по всей плоскости скольжения одновременно, а путем

⁶В области радиусом в несколько межатомных расстояний ($\approx 5\text{--}10 \text{ \AA}$), называемой ядром дислокации, искажения решетки настолько велики, что не могут быть описаны в рамках теории упругости.

⁷Плоскость скольжения для простейшего случая прямолинейной дислокации определяется как плоскость, в которой лежат вектор Бюргерса дислокации и линия дислокации.

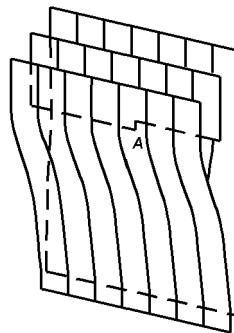
последовательного перемещения атомов, находящихся у линии дислокации. Так, например, при движении краевой дислокации по плоскости скольжения (рис. 3.9) разрываются и пересоединяются связи между атомами, лежащими у линии дислокации, и при этом эти атомы совершают небольшие перемещения. В результате лишняя полуплоскость, занимавшая определенное положение в кристаллической решетке, соединяется с атомной плоскостью, находящейся под плоскостью скольжения, а соседняя атомная плоскость становится теперь лишней полуплоскостью. Таким образом, поскольку скольжение происходит путем последовательного перемещения атомов на небольшие расстояния, то для движения дислокаций в плоскости скольжения достаточно внешних напряжений намного меньших, чем напряжения, необходимые для пластической деформации совершенного кристалла без дислокаций.

Плавное движение возможно и для винтовой дислокации. Однако все плоскости, которые содержат винтовую дислокацию, содержат и вектор Бюргерса, поэтому винтовая дислокация, в отличие от краевой, может скользить в любом направлении. Следовательно, все плоскости, содержащие винтовую дислокацию, являются плоскостями плавного скольжения.

Переползание дислокаций происходит, как правило, перпендикулярно плоскости ее скольжения и осуществляется или присоединением вакансий (приток вакансий), или присоединением атомов (приток междоузельных атомов) к краю полуплоскости, при этом полуплоскость смещается на одно межатомное расстояние (рис. 3.10).⁸ На рис. 3.10 представлена краевая дислокация, линия дислокации которой переходит с одной плоскости скольжения на другую, расположенную на одно межплоскостное расстояние выше. Когда вакансия подходит к ступеньке, последняя смещается на одно межплоскостное расстояние, а сама вакансия исчезает. Аналогично поглощаются и междоузельные атомы. Смещение дислокации происходит в противоположных направлениях при поглощении вакансии или междоузельного атома. Оба процесса требуют диффузионного перемещения вакансий или междоузельных атомов к дислокации. Такое движение, носящее диффузионный характер, является результатом стремления системы к уменьшению свободной энергии за счет уменьшения упругой энергии решетки (см. п. 7). Наличие незаполненных (ненасыщенных) связей у атомов полуплоскости облегчает отрыв атомов и вакансий от дислокации или присоединение междоузельных атомов и вакансий к дислокации. Так как скорость диффузии быстро уменьшается с понижением температуры (см. гл. 8), то переползание (в отличие

⁸Обратный процесс — испускание точечных дефектов краевой дислокацией — является основным источником термодинамически равновесных точечных дефектов.

Рис. 3.10. Схема переползания краевой дислокации при поглощении вакансий. Буквой *A* обозначена ступенька, к которой присоединяется вакансия.



от скольжения) происходит с заметной скоростью только при достаточно высоких температурах.

Рассмотренный механизм переползания дислокации применим к любой дислокации, содержащей краевую компоненту. Однако винтовая дислокация не имеет «лишней» полуплоскости, поэтому переползание для винтовой дислокации в общепринятом смысле невозможно. Тем не менее, если линия винтовой дислокации скручивается в спираль, то такая спираль имеет краевую компоненту и, следовательно, получает возможность переползать. Переползание в этом случае вызывает расширение спирали в радиальном направлении. Таким образом, переползание винтовых дислокаций превращает их в геликоидальные (изогнутые по цилиндрической спирали с осью вдоль вектора Бюргерса).

4. Так как поля упругих напряжений вокруг отдельных дислокаций могут перекрываться, то дислокации могут взаимодействовать друг с другом (притягиваться, отталкиваться), если это взаимодействие приводит к уменьшению упругой энергии кристалла. Так, если сумма энергий двух дислокаций больше (меньше) энергии дислокации, образующейся при слиянии этих дислокаций в одну, то эти дислокации будут притягиваться (отталкиваться) и объединяться (распадаться) в одну (на две). Например, при сближении двух дислокаций, расположенных в одной плоскости скольжения, с одинаковыми, но противоположно направленными векторами Бюргерса сжатие и растяжение кристалла по обе стороны от плоскости скольжения взаимно компенсируются, а при сближении двух дислокаций с одинаково направленными векторами Бюргерса — увеличивается. Поэтому параллельные винтовые и параллельные краевые дислокации, лежащие в общей плоскости или в одной плоско-

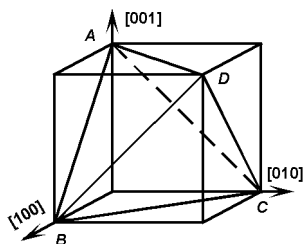
сти скольжения, взаимодействуют так же, как заряженные нити: разноименные притягиваются, а одноименные отталкиваются с силой, обратно пропорциональной расстоянию. В общем случае взаимодействие дислокаций носит более сложный характер, зависящий от взаимной ориентации векторов Бюргерса дислокаций, однако обычно действует упрощенное правило: две дислокации притягиваются, если их вектора Бюргерса составляют тупой угол, и отталкиваются, — если острый.

5. Число различных типов дислокаций в реальных кристаллах конечно. Каждая решетка Браве имеет свой, присущий только ей, набор возможных дислокаций с определенными векторами Бюргерса. Основная причина появления этого свойства заключается в том, что в общем случае вектор Бюргерса — это векторная сумма векторов трансляций решетки, взятых с целочисленными коэффициентами. Поэтому величина и направление возможных $\mathbf{b}$ ограничены рядом дискретных значений, определяемых структурой решетки. С другой стороны, как было показано в п. 2, упругая энергия искажений решетки при наличии дислокаций пропорциональна квадрату вектора Бюргерса. Поэтому для каждого типа кристаллической решетки существуют дислокации с наименьшими векторами Бюргерса, которые обладают наименьшей энергией, наиболее энергетически устойчивы против расщепления и наиболее подвижны. Как правило, такими дислокациями являются *полные дислокации*, то есть дислокации, вектор Бюргерса которых равен вектору трансляции решетки. Так, в кристаллах с о.ц.к. решеткой минимальной энергией обладают дислокации с векторами Бюргерса $(a/2)\langle 111 \rangle$, в кристаллах с г.ц.к. решеткой — с векторами Бюргерса $(a/2)\langle 110 \rangle$. Эти дислокации наиболее распространены в названных решетках. В качестве примера рассмотрим дислокации в структуре типа алмаза.

В структуре типа алмаза энергетически наиболее выгодны дислокации с вектором Бюргерса $(a/2)\langle 110 \rangle$. Любое сложное перемещение в решетке можно рассматривать как сумму последовательных трансляций в направлениях $\langle 110 \rangle$. Поэтому линии простых дислокаций должны быть направлены вдоль какого-либо из направлений $\langle 110 \rangle$. В решетке алмаза возможны 9 различных типов дислокаций, из которых три являются простыми, а остальные — сложными (см. табл. 3.1). Покажем это.

Рассмотрим тетраэдр, образованный направлениями $\langle 110 \rangle$ и вписанный в куб (рис. 3.11). Вектор Бюргерса $(a/2)\langle 110 \rangle$ направлен вдоль ребра BC тетраэдра. Каждое ребро такого тетраэдра представляет собой линию простой дислокации, а суммирование двух ребер тетраэдра приводит к образованию линии сложной дислокации из линий двух простых дислокаций. Сложные дислокации возникают, когда линия дислокации

Рис. 3.11. Тетраэдр с ребрами вдоль направлений $\langle 110 \rangle$.



последовательно меняет свою ориентировку от одного из направлений $\langle 110 \rangle$ к другому и в общем случае оказывается непараллельной ни одному из направлений типа $\langle 110 \rangle$. Различные комбинации ребер при суммировании дают весь возможный набор сложных дислокаций.

Из табл. 3.1 следует, что к простым дислокациям в структуре типа алмаза относятся: винтовая (линия дислокации совпадает с вектором Бюргерса) (рис. 3.12,а), краевая с плоскостью скольжения $\{100\}$ (линия дислокации перпендикулярна вектору Бюргерса) (рис. 3.12,в) и 60-градусная дислокация с плоскостью скольжения $\{111\}$ (вектор Бюргерса образует угол 60° с линией дислокации) (рис. 3.12,б). Остальные дислокации — сложные.

Дислокации в структурах типа сфалерита, NaCl, вюртцита подробно рассмотрены в [27].

6. Характерной особенностью краевых и 60-градусных дислокаций является то, что атомы, образующие край атомной полуплоскости, имеют ненасыщенные (оборванные) связи, то есть эти дислокации электрически активны в отличие от электрически неактивных винтовых дислокаций. Оборванные связи вносят вклад в энергию дислокаций. Однако основная часть энергии дислокации и в этом случае приходится на энергию упругих искажений. По проведенным оценкам энергия одной разорванной связи составляет 0.7 эВ для германия и 1.2 эВ для кремния [17], то есть в несколько раз меньше, чем упругая энергия, приходящаяся на одну атомную плоскость, пересекаемую дислокацией.

7. Взаимодействие дислокаций с точечными дефектами.

Стремление к уменьшению свободной энергии кристалла вызывает эффективное взаимодействие дислокаций с точечными дефектами и прежде всего с примесными атомами. В результате этого взаимодействия атомы примеси распределяются в решетке неравномерно, как правило, группируясь вблизи дислокаций и образуя так называемые *атмосферы Коттрелла*.

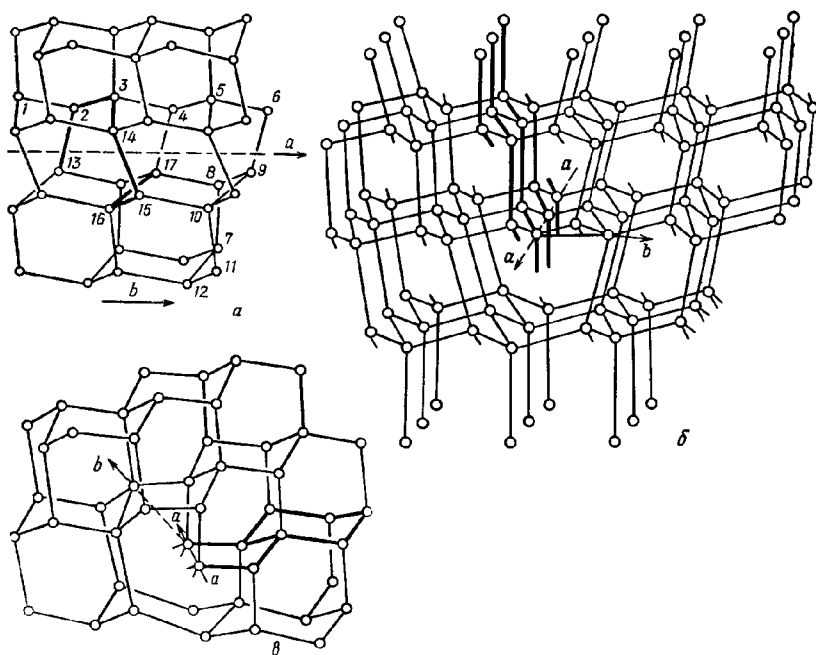


Рис. 3.12. Простые дислокации в решетке алмаза: *a* — винтовая; *б* — 60-градусная с плоскостью скольжения $\{111\}$; *в* — краевая с оборванными связями и с плоскостью скольжения $\{100\}$. *a* — *a* — линия дислокации; *b* — вектор Бюргерса.

Таблица 3.1. Возможные типы дислокаций в решетке алмаза [27].

NN	Линия дислокации	Символ линии дислокации	Угол между линией дислокации и вектором Бюргерса	Плоскость скольжения	Число оборванных связей
<i>Простые</i>					
1	BC	$\langle 110 \rangle$	0°	—	0
2	AB, AC, DB, DC	$\langle 110 \rangle$	60°	$\{111\}$	1.41
3	AD	$\langle 110 \rangle$	90°	$\{100\}$	2.83 или 0
<i>Сложные</i>					
4	$BC + AC, BC + BA$				
	$BD + BC, DC + BC$	$\langle 211 \rangle$	30°	$\{111\}$	0.82
5	$AC + AB, DC + BC$	$\langle 211 \rangle$	90°	$\{111\}$	1.63
6	$AD + BD, DA + BA$				
	$AD + CD, DA + CA$	$\langle 211 \rangle$	$73^\circ 13'$	$\{311\}$	2.45 или 0.82
7	$AB + DB, AC + DC$	$\langle 211 \rangle$	$54^\circ 44'$	$\{110\}$	1.63 или 0
8	$AC + DB, AB + DC$	$\langle 100 \rangle$	90°	$\{110\}$	2.0 или 0
9a	$AD + BC, AD + CB$	$\langle 100 \rangle$	45°	$\{100\}$	2.0 или 0
9b	$AC + BD, AB + CD$	$\langle 100 \rangle$	45°	$\{100\}$	2.0 или 0

Различают три типа взаимодействия дислокаций с точечными дефектами: упругое взаимодействие I рода (размерное); упругое взаимодействие II рода (взаимодействие по модулю упругости); электрическое взаимодействие (кулоновское).

1). Упругое взаимодействие I рода обусловлено полями упругих напряжений вокруг дислокаций и вокруг примесного атома. Знак напряжений вокруг примесного атома зависит от соотношения радиусов атомов основного вещества r_0 и примеси r . В случае примеси замещения при $\Delta r = r - r_0 > 0$ возникают упругие напряжения радиального сжатия, при $\Delta r < 0$ — растяжения. В случае примеси внедрения напряжения всегда сжимающие. Атом, создающий растягивающие напряжения, будет стремиться в сжатую область вокруг дислокации, в то время как атом, создающий сжимающие напряжения, — в растянутую. Взаимодействие этого типа связано с краевыми дислокациями. Упругая энергия E_I такого взаимодействия равна работе, совершаемой упругими силами при замене атома основного вещества примесным атомом: $E_I \sim Gbr_0^3 \Delta r \sin \theta / r_0 R$, где G — модуль сдвига, R и $\sin \theta$ — сферические координаты атома примеси (краевая дислокация находится в начале координат) [17]. Знак отношения $\Delta r / r_0$ показывает, куда стремится атом примеси: в растянутую или сжатую область решетки. Характерная энергия этого взаимодействия в полупроводниках составляет ≈ 0.5 эВ.

Эту же работу надо затратить и для отрыва примесного атома от дислокации. Расчеты показывают, что уже примерно на 3–5 межатомных расстояниях энергия $E_I \approx kT$. Это значит, что дальше этого расстояния от дислокации «облако» атомов примеси (атмосфера Коттрелла) расщепляется тепловым движением. Чем сильнее тепловое движение, тем меньше концентрация атомов примеси в облаке.

2). Упругое взаимодействие II рода обусловлено тем, что примесный атом или вакансия представляют собой малые области с упругими постоянными, иными, чем у матрицы. В этом случае энергия взаимодействия E_{II} между дислокацией и точечным дефектом пропорциональна: $E_{II} \sim (\Delta G)b^2/R^2$, где R — расстояние от точечного дефекта до дислокации [17]. В отличие от первого взаимодействия, второе сказывается лишь на очень малых расстояниях; по порядку величины оно составляет $E_{II} \approx 0.2$ эВ. Это взаимодействие вызывает увеличение концентрации вакансий вокруг дислокаций.

3). Электрическое взаимодействие проявляется главным образом в полупроводниковых и ионных кристаллах. Оборванные связи дислокаций обычно действуют как акцепторы. В материале n -типа эти связи захватывают электроны и тем самым создают кулоновское взаимодействие

между дислокацией и положительно заряженными ионами. Величина такого взаимодействия $E_{эл} \sim f e^2/a$, где f — доля свободных оборванных связей; a — расстояние между оборванными связями вдоль линии дислокации; e — заряд электрона. Максимальное значение $E_{эл}$ при комнатной температуре ≈ 0.02 эВ.

Влияние дислокаций на некоторые физические свойства кристалла

Дислокации влияют не только на механические свойства кристаллов (пластичность, прочность),⁹ для которых наличие дислокаций является определяющим, но и на другие физические свойства кристаллов, например, на электрические.

Дислокации подобно примесным атомам и собственным точечным дефектам могут создавать дополнительные электронные состояния в запрещенной зоне. Например, при наличии краевой дислокации атомы, образующие край атомной полуплоскости (линию дислокации), имеют ненасыщенные связи. Неспаренные электроны атомов с оборванными связями приводят к появлению дополнительных состояний. При этом возможны два случая:

а) захват дополнительного электрона на ненасыщенную связь с возможным образованием разрешенного состояния электрона в запрещенной зоне (дислокация проявляет акцепторные свойства);

б) разрыв ненасыщенной связи и превращение неспаренного электрона в электрон проводимости (дислокация проявляет донорные свойства).

Таким образом, краевые дислокации по электрическим свойствам оказываются сходными с примесными центрами. Однако между ними имеются и существенные различия. Дислокации образованы непрерывной цепочкой атомов основного вещества (среднее расстояние между ними очень мало вдоль дислокации), и поэтому их нельзя считать изолированными дефектами как в случае примесных атомов. Отдельные дислокации можно представить как одномерные кристаллы, которые в запрещенной зоне могут создавать энергетические дислокационные зоны. Следует заметить, что при наличии примесных атмосфер Коттрелла вблизи дислокации могут возникать и дискретные локальные уровни.

Краевую дислокацию в определенном смысле можно рассматривать как заряженную линию. Если, например, проявляются акцепторные свойства дислокации, то в материале n -типа она будет заряжена отрицатель-

⁹Наличие дислокаций делает кристалл более пластичным. Для упрочения сплавов необходимо уменьшить скорость движения дислокаций путем их торможения, закрепления и т. д. [16].

но и окружена цилиндрическим слоем индуцированного положительного заряда. Локальные области пространственного заряда, образующиеся вокруг дислокаций, могут приводить к рассеянию носителей заряда, тем самым снижая их подвижности, причем подвижность вдоль и перпендикулярно дислокации оказывается существенно различной.

Электрическая активность дислокаций, как правило, отрицательно сказывается на свойствах полупроводниковых приборов, например, вызывает преждевременный пробой в областях прибора, где дислокация пересекает $p-n$ -переход. Дислокации оказывают существенное влияние и на время жизни свободных носителей. В чистых кристаллах нередко именно они ограничивают времена жизни неравновесных носителей заряда.

Кроме того, дислокации определяют концентрацию точечных дефектов в полупроводниковых монокристаллах, так как являются их «источником» и «стоком». Области увеличенных межатомных расстояний являются потенциальными «стоками» междоузельных атомов и «источниками» вакансий в кристалле, а области сжатых межатомных расстояний — наоборот (см. п. 3). Точечные дефекты появляются также и при пересечении отдельно движущихся дислокаций, и при аннигиляции дислокаций, движущихся в параллельных плоскостях скольжения.

Дислокации увеличивают скорость диффузии атомов в кристалле (см. гл. 8), ускоряют эффекты «старения» материала и другие процессы, протекающие с участием диффузии. Сгущение облаков Коттрелла вокруг дислокаций может привести к образованию включений второй фазы. В некоторых случаях дислокации могут играть определяющую роль в процессах роста кристаллов (см. гл. 4).

Методы наблюдения дислокаций

Существует достаточно много экспериментальных методов наблюдения дислокаций. Например, за дислокациями можно наблюдать с помощью электронного микроскопа с высокой разрешающей способностью, с помощью рентгеновской топографии. Однако особенно широкое распространение при изучении дислокаций получили методы избирательного травления и декорирования. Метод избирательного травления основан на том, что вблизи дислокаций энергия связи атомов гораздо слабее, чем в недеформированной решетке. Поэтому места выхода дислокаций на поверхность кристалла травятся специально подобранным травителем быстрее, чем окружающая дислокацию поверхность. В результате такого травления на поверхности кристалла возникают ямки травления. Подсчет

их позволяет определять одну из важнейших для полупроводниковых кристаллов характеристик — плотность дислокаций, выражаемую числом ямок травления на квадратный сантиметр. Расположение на поверхности кристалла и форма ямок травления дают возможность исследовать свойства дислокаций. По фигурам травления можно проследить за образованием и движением дислокаций в процессе пластической деформации. Недостатком метода травления является то, что с его помощью определяются не все дислокации, а лишь те, которые выходят на поверхность образца или образуют петли на незначительной глубине от поверхности. Кроме того, ямки травления могут возникать не только в местах выхода дислокаций на поверхность кристалла, но и в местах скоплений точечных дефектов и включений второй фазы. Поэтому важна правильность идентификации ямок травления.

Метод декорирования основан на том, что дислокации могут служить местами стока примесных атомов, что позволяет их выявлять. Для этого диффузией вводят в кристалл примеси, которые осаждаются на линиях дислокаций, декорируют их. Это делает возможным их наблюдение с применением, например, инфракрасной микроскопии. В отличие от метода травления метод декорирования позволяет наблюдать дислокационную структуру не только на поверхности, но и в глубине кристалла. Существенным недостатком этого метода является необходимость нагревания и выдержки образцов при повышенных температурах, что влияет на количество и распределение дислокаций. Кроме того, к недостаткам этого метода следует отнести невозможность наблюдения за изменением дислокационной структуры.

В полупроводниковых монокристаллах, выращенных в обычных условиях, плотность дислокаций колеблется в пределах $N_D = 10^4\text{--}10^6\text{ см}^{-2}$. Путем отжига можно понизить эту плотность до $10^3\text{--}10^4\text{ см}^{-2}$. Методы выращивания монокристаллов, практически не содержащих дислокаций, сложны и разработаны только для немногих материалов, в частности для германия и кремния и некоторых соединений $A^{III}B^V$ (см. гл. 7).

3.1.3. Двумерные дефекты

Под двумерными дефектами понимают такие нарушения в идеальном расположении атомов в кристалле, которые обладают большой протяженностью в двух измерениях при незначительной (в несколько межатомных расстояний) протяженности в третьем.

Малоугловые границы. Бюргерс предположил, что граница соединения двух монокристаллических блоков, разориентированных друг от

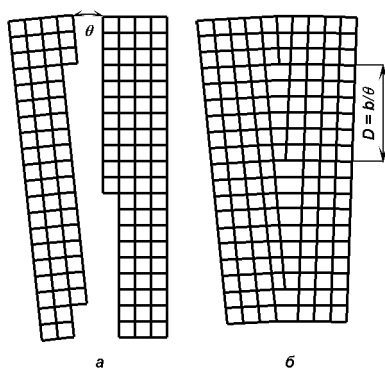


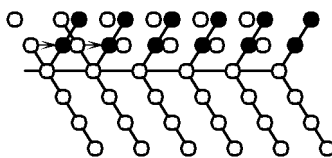
Рис. 3.13. Дислокационные малоугловые границы.

носителем угла на небольшой угол, состоит из совокупности дислокаций. Экспериментальные исследования подтвердили дислокационную модель Бюргерса для границ зерен с малым углом разориентировки. Договорились поэтому совокупность дислокаций, которые образуются при сращивании различным образом ориентированных монокристаллических блоков с малым углом разориентировки ($2-5^\circ$), называть *малоугловой границей* (МУГ). На рис. 3.13 показана простейшая дислокационная модель малоугловой границы для случая, когда краевые дислокации, выстроенные в ряд, образуют границу блоков, симметрично разориентированных на угол θ . Согласно этой модели МУГ представляет собой простую наклонную границу, состоящую из совокупности одинаковых краевых дислокаций, расположенных на расстоянии $D \sim b/\theta$ одна над другой (b — длина вектора Бюргерса) в параллельных плоскостях скольжения. Анализ полей напряжений взаимодействующих дислокаций показал, что такая система одинаковых краевых дислокаций образует дислокационную сетку, которая наиболее устойчива в кристалле [27]. Заметим, что область упругого искажения решетки вблизи МУГ не распространяется далеко в глубь кристаллических блоков и ограничена в основном слое, толщина которого равна расстоянию между дислокациями D .¹⁰

Чем больше угол разориентации, тем чаще расположены дислокации на границе между блоками. Наконец, при большой разориентации $\theta > 10-12^\circ$ малоугловые границы превращаются в обычные границы меж-

¹⁰Если соседние блоки повернуты один относительно другого на угол θ вокруг оси, перпендикулярной к плоскости границы блока, то граница состоит из совокупности винтовых дислокаций и называется границей кручения. Границы блоков общего типа включают в себя и наклон (совокупность краевых дислокаций) и кручение (совокупность винтовых дислокаций).

Рис. 3.14. Схема возникновения двойниковой границы в кристалле.



ду зернами. *Границы зерен* в поликристаллическом слитке образуются в результате одновременного разрастания нескольких кристаллических зародышей (см. гл. 4). В большинстве случаев растущие зерна кристаллов имеют беспорядочную кристаллографическую ориентацию.

Двойники, дефекты упаковки. В условиях, неблагоприятных для скольжения, пластическая деформация кристаллов может приводить к образованию не дислокаций, а двойников или дефектов упаковки. *Двойники и дефекты упаковки* представляют собой дефекты, при наличии которых происходят нарушения кристаллографической ориентации частей кристалла из-за нарушения порядка чередования атомных плоскостей. Рассмотрим эти дефекты, например, в плотноупакованной г.ц.к. решетке. В этой решетке последовательность укладки плотноупакованных плоскостей $\{111\}$ — это последовательность типа $ABCABCABC\dots$. В случае двойникового дефекта в г.ц.к. решетке слои $\{111\}$ чередуются в последовательности $ABCA$ BCB $ACBA\dots$. Расположение слоев слева от слоя C в выделенном блоке представляет собой зеркальное отражение расположения слоев справа от него. На рис. 3.14 показано двойникование на примере двумерной решетки. Светлыми кружками на нем представлены положения атомов до двойникового сдвига, черными — после двойникового сдвига. Величина смещения атомов в каждой из параллельных плоскостей пропорциональна расстоянию данной плоскости от плоскости двойникового сдвига. Так как при двойниковании сохраняются неизменными все межатомные расстояния, а углы между связями изменяются только в одном граничном слое, то возникновение двойниковой границы требует небольшой энергии. Двойникование полупроводниковых кристаллов является очень распространенным дефектом структуры. Как правило, этот дефект вызывается теми же причинами, что и образование дислокаций, и часто порождается скоплениями последних. Возникновению двойников способствуют также скопление примесей на фронте кристаллизации, отклонение направления роста от направления температурного градиента и механические напряжения (см. гл. 6).

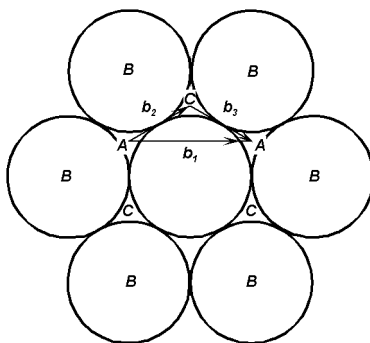
В случае дефекта упаковки последовательность типа $ABCABCABC\dots$

имеет вид $ABCA\boxed{C}ABC$, то есть дефект представляет собой как бы тонкий двухатомный слой гексагональной упаковки в г.ц.к. решетке. Нарушение правильного чередования слоев решетки может быть создано путем добавления или изъятия слоя; соответственно различают дефект упаковки типа внедрения и типа вычитания. Приведенный пример иллюстрирует дефект упаковки типа вычитания. Одной из возможных причин, приводящих к появлению дефектов упаковки, является пластическая деформация, происходящая путем сдвига, например, в г.ц.к. решетке по плоскости $\{111\}$. На рисунке 3.15 показан слой атомов B в плоскости (111) , по которому происходит скольжение в г.ц.к. решетке. Полному сдвигу слоев, лежащих выше этой плоскости, соответствует дислокация с вектором Бюргерса $(a/2)\langle 110 \rangle$; однако такой сдвиг энергетически невыгоден. Энергия полной дислокации с вектором Бюргерса $(a/2)\langle 110 \rangle$ может понизиться вследствие ее распада в плоскости (111) на две дислокации с векторами Бюргерса в направлениях $\langle 121 \rangle$ в той же плоскости. Дислокации с векторами Бюргерса меньшими, чем параметр решетки, называются *частичными*; а комплекс, состоящий из двух частичных дислокаций, связанных между собой дефектом упаковки, называют *растянутой дислокацией*. Проследим, как меняется порядок чередования атомов при распаде полной дислокации на две частичные с помощью рисунка 3.15. Видно, что трансляция AA переводит атомы в одинаковые положения, не меняя их чередования. Иначе говоря, если полная дислокация b_1 проходит через кристалл, вызывая пластический сдвиг на b_1 , то структура кристалла вдоль плоскости скольжения полностью восстанавливается, а на поверхности кристалла появляется ступенька шириной b_1 . В противоположность этому трансляция b_2 или b_3 , то есть сдвиг на вектор Бюргерса частичной дислокации, нарушает этот порядок, переводя лунку A в C или C в A . Таким образом, вместо прямого скольжения $A \rightarrow A$ реализуется зигзагообразное скольжение $A \rightarrow C \rightarrow A$, и в результате меняется порядок положений слоев $\{111\}$, то есть образуется дефект упаковки.

3.1.4. Объемные нарушения

Зональные упругие напряжения. При макроскопически неоднородном характере поля внешних сил, действующих на кристалл, в нем могут возникать напряжения, которые называются *зональными*. Эти напряжения, уравнивающиеся в макрообъемах кристалла, существенно влияют на его свойства, а при определенных условиях могут вызывать пластическую деформацию, трещинообразование и даже разрушение кристалла.

Рис. 3.15. Схема распада полной дислокации $\mathbf{b}_1 = (a/2)\langle 110 \rangle$ на две частичные дислокации с векторами $\mathbf{b}_2 = \mathbf{b}_3 = (a/6)\langle 112 \rangle$ в плоскости плотнейшей упаковки (111).



В качестве примера рассмотрим некоторые наиболее распространенные случаи возникновения зональных напряжений.

Зональные напряжения возникают при выращивании кристаллов из-за наличия термических напряжений. При росте кристалла из жидкой фазы отвод тепла от его наружных слоев обычно происходит быстрее, чем от слоев, находящихся в середине. Более горячие внутренние слои мешают наружным слоям сократиться. В результате возникают термоупругие напряжения, сжимающие во внутренних объемах и растягивающие во внешних. Если эти напряжения превысят предел текучести¹¹ материала, то они вызовут пластическую деформацию, которая полностью или частично снимет термические напряжения. Зональные термические напряжения возникают также и при быстром охлаждении выращенного кристалла от температуры роста до комнатной температуры (см. гл. 6).

Зональные напряжения возникают и на границе подложки и эпитаксиальной пленки, выращиваемой из материала, отличного от материала подложки. Помимо упругой деформации, в эпитаксиальной пленке может происходить и пластическая деформация, вызванная зональными напряжениями. Наиболее существенными факторами, влияющими на процесс пластической деформации, оказываются толщина пленки и разность коэффициентов термического расширения материалов пленки и подложки. Более подробно этот вопрос будет рассмотрен в разделах, касающихся технологии выращивания пленок (см. гл. 9).

Поры. *Поры* — это объемные пустоты макроскопических размеров в кристаллической решетке. К появлению пор приводят скопления многих

¹¹Предел текучести — это напряжение, вызывающее деформацию, равную 0.2% от длины или другого размера кристалла.

вакансий. Поры могут образовываться в ходе выращивания кристаллов или при радиационном облучении кристаллов.

Включения второй фазы. *Включения второй фазы* — это частицы макроскопических размеров, состав которых отличается от состава основного вещества. В кристаллах могут встречаться включения второй фазы двух типов: включения одного из компонентов соединения, образующиеся при выращивании монокристаллов (например, Pb в PbTe), и включения макроскопических частиц инородной фазы, образующиеся при выращивании легированных кристаллов (например, CdTe в кристаллах PbTe, легированном кадмием).

3.2. Примеси

В этом разделе мы рассмотрим группу дефектов, которая связана с введением в полупроводник чужеродных атомов — примесей. Сначала мы приведем существующую классификацию примесей, а затем рассмотрим конкретные примеры всех типов примесей в элементарных полупроводниках и полупроводниковых соединениях.

По характеру вхождения в кристаллическую решетку полупроводников примеси можно разделить на примеси замещения и примеси внедрения. Возможны ситуации, когда атомы примеси одновременно являются и атомами замещения, и атомами внедрения. В общем случае, как правило, чем меньше разность ковалентных радиусов атомов примеси и атомов основного вещества и чем меньше разность их электроотрицательностей, тем больше вероятность размещения примесного атома в узле решетки полупроводника и тем больше предельная растворимость примеси в основном веществе (см. гл. 7).

Примеси могут быть электрически активными и неактивными. Электрически активные примеси влияют на электрические свойства полупроводника, а электрически неактивные — нет. К электрически активным относят примеси, которые при введении их в полупроводник образуют уровень или систему уровней в запрещенной зоне полупроводника¹² и при увеличении температуры от 0 К могут ионизоваться с изменением концентрации свободных носителей заряда в полупроводнике. К электрически неактивным примесям относят примеси, которые не образуют

¹²В узкозонных полупроводниках некоторые примеси могут образовывать и так называемые квазилокальные (резонансные) уровни, которые расположены вблизи краев зон в области разрешенных состояний. Такие примеси также относятся к электрически активным. Например, примесь In в PbTe образует донорный уровень, лежащий на 0.07 эВ выше края зоны проводимости.

таких уровней и не ионизируются.

Электрически активные примеси принято подразделять на доноры, акцепторы и амфотерные примеси.

Донор — это структурный дефект в кристаллической решетке полупроводника, способный отдавать электроны в зону проводимости или другим примесным центрам. Доноры, отдавая электроны, не участвующие в образовании химической связи, в зону проводимости, увеличивают концентрацию свободных электронов и уменьшают концентрацию дырок.¹³ В полупроводнике, содержащем донорные примеси, электрический ток переносится преимущественно электронами (электроны являются основными носителями заряда, а дырки — неосновными), обуславливая примесную электронную проводимость. Такой полупроводник называется полупроводником *n*-типа.

Акцептор — структурный дефект в кристаллической решетке полупроводника, способный присоединять к себе электроны из валентной зоны или с других примесных центров. Акцепторы, захватывая электроны из валентной зоны, увеличивают концентрацию дырок и уменьшают концентрацию свободных электронов. В полупроводнике, содержащем акцепторные примеси, электрический ток переносится преимущественно дырками (основными носителями заряда в нем являются дырки, а неосновными — электроны), обуславливая примесную дырочную проводимость. Такой полупроводник называется полупроводником *p*-типа.

Примеси, которые в кристалле могут отдавать или принимать только один электрон и, следовательно, находиться только в двух различных зарядовых состояниях, называют простыми. Таким образом, простой донор в основном (неионизованном) состоянии — это примесь, с которой при сформировавшихся валентных связях связан один электрон; а простой акцептор в основном состоянии — это примесь, с которой при сформировавшихся валентных связях связана одна дырка. Если же примесные атомы в основном состоянии при сформировавшихся валентных связях связывают несколько электронов или дырок, то примесь называют многозарядным донором или акцептором. Поведение примеси в полупроводнике при неизменных внешних условиях в значительной степени определяется ее положением в периодической системе.

Примеси в полупроводниках характеризуются энергетическим положением уровня E_i (основное состояние) для простых доноров и акцепторов или системы альтернативных энергетических уровней (основное

¹³Увеличение электрически активными примесями концентрации носителей одного знака и уменьшение другого связано с тем, что в примесных полупроводниках соблюдается соотношение $np = n_i^2 = p_i^2$.

и заряженные состояния примеси) в случае многозарядных примесей.¹⁴ Такие уровни являются локальными, так как занимающие их электроны расположены вблизи примесных центров. Простые доноры и акцепторы, поведение которых в первом приближении описывается водородоподобной моделью [6], образуют, как правило, мелкие уровни в запрещенной зоне, расположенные вблизи краев зоны проводимости или валентной зоны, соответственно. Многозарядные доноры и акцепторы могут образовывать глубокие уровни в запрещенной зоне, расположенные вдали от соответствующих разрешенных зон.

Примеси в полупроводниках ионизируются в результате теплового возбуждения, освещения или какого-либо другого внешнего воздействия. В отсутствие внешних воздействий при температуре вблизи нуля атомы примесей в полупроводниках не ионизованы.¹⁵ Энергия ионизации примеси зависит от силы связи валентных электронов примесного атома с его ядром и с электронами других атомов в кристалле. Обычно валентные электроны примесного атома, не принимающие участия в образовании химической связи в кристалле, ионизируются в первую очередь при внешних воздействиях. Энергия ионизации примеси экспериментально может быть определена из измерений температурной зависимости постоянной Холла (термическая энергия ионизации), спектральной зависимости коэффициента поглощения света и фотопроводимости (оптическая энергия ионизации). Установлено, что энергия ионизации примеси зависит от концентрации примеси и в общем случае уменьшается с ее ростом.

Термин «амфотерность» происходит от греческого слова «амфотерос», что означает двойственный. Это понятие применительно к примеси в полупроводниках впервые ввел Данлэп (1955 г.) при анализе электрических свойств германия, легированного золотом. Оказалось, что примесь золота в одном и том же кристалле — германии — проявляет себя и как донор, и как акцептор. Под амфотерностью элементов принято понимать их способность образовывать положительные и отрицательные ионы. Если рассматривать амфотерные примеси в полупроводниках, то оказывается, что таких примесей не так уж много, но все же достаточно, чтобы выделить их в самостоятельный класс. К *амфотерным примесям* в полупроводниках относят примеси, которые в одном и том же матери-

¹⁴В этой главе, если специально не оговорено, рассматриваются полупроводники с одной примесью и пренебрегается эффектами компенсации [6]. Эти эффекты могут проявляться при наличии в легированном кристалле собственных точечных дефектов и (или) одновременно нескольких примесей.

¹⁵В сильно легированных полупроводниках ситуация может быть иной: примеси могут оставаться ионизованными и при нулевой температуре [6].

але могут вести себя и как доноры, и как акцепторы. В настоящее время все известные амфотерные примеси в полупроводниках можно сгруппировать по кристаллохимическому принципу, то есть по их размещению в кристаллической решетке. Выделяют пять типов амфотерных примесей в полупроводниках: узельные, междоузельные, диссоциативные, катионно-анионные и ассоциативные амфотерные центры.

Примесные атомы, способные быть и донорами и акцепторами в *одной* из кристаллических позиций — в узле или междоузлии — отнесем к первым двум типам и будем их называть *амфотерным узельным* или *амфотерным междоузельным* центром, соответственно.

Если примесные атомы, размещенные в узлах, являются акцепторами, а в междоузлиях — донорами, то такие примесные атомы назовем узельно-междоузельными или *диссоциативными амфотерными* центрами.¹⁶

Примесные центры четвертого типа, проявляющие амфотерность при замещении различных по своей природе компонентов (катионы или анионы) полупроводниковых соединений, будем называть *катионно-анионными амфотерными* центрами.

Наконец, как всякие другие, амфотерные примеси могут образовывать сложные ассоциированные дефекты при взаимодействии друг с другом или с другими точечными дефектами. Более того, в кристалле могут возникать такие ассоциаты точечных дефектов, которые проявляют амфотерность, хотя каждый из компонентов ассоциата в отдельности и не является амфотерным. Подобные сложные амфотерные центры, являющиеся продуктами взаимодействия различных точечных дефектов, будем называть *ассоциативными амфотерными* центрами.

В элементарных полупроводниках проявляются амфотерные примесные центры первых трех типов, а в полупроводниковых соединениях — все пять типов амфотерных центров.

Для первых двух типов амфотерных центров характерно проявление амфотерности в одной и той же кристаллической позиции. В этом случае примесный центр должен иметь возможность как отдать электрон из какой-либо своей электронной оболочки в зону проводимости полупроводника, так и захватить электрон на эту оболочку. Очевидно, что обе эти возможности у одного и того же атома, первоначально находящегося в зарядовом состоянии Z_0 , могут быть реализованы только, если его «работающая» электронная оболочка заполнена не полностью. А это могут быть лишь атомы *d*- или *f*-элементов. Экспериментальные данные подтверждают это правило. Но, как показывает опыт, во многих случаях примеси переходных металлов не проявляют свойства амфотерности. Примеси же *f*-атомов часто вообще электрически нейтральны.

Амфотерные примесные центры третьего типа — диссоциативные — встречаются чаще всего в элементарных полупроводниках. Из общих соображений ясно, что для равновероятного размещения примесного атома в узле и междоузлии требуются примерно равные энергии их введения в обе позиции. Однако вычисление этих энергий наталкивается на ряд трудностей, связанных прежде всего с различием взаимодействий примесного атома с окружающими атомами кристалла. В узле примесный атом вступает в химическую связь с окружающими атомами кристалла. В междоузлии же электроны окружающих атомов лишь отталкивают электроны примесного атома и поэтому химических связей между атомами окружения и примесным центром не возникает.

Катионно-анионные амфотерные центры возможны лишь в полупроводниковых соединениях. Размещение одного и того же примесного атома в различных подрешетках соеди-

¹⁶Заметим, что в принципе могут быть и такие диссоциативные амфотерные центры, которые проявляют, наоборот, в узлах донорный характер, а в междоузлиях — акцепторный.

нения реализуется только для тех атомов, которые имеют промежуточную валентность по сравнению с валентностями компонентов соединения. Наиболее наглядно это проявляется для элементов четвертой группы периодической системы в полупроводниках $A^{III}B^V$. В этом случае атом C^{IV} , размещаясь в катионном узле A^{III} , проявляет донорные свойства, а занимая анионный узел B^V — акцепторные. При этом проявляется известное правило валентности $\Delta V = \pm 1$, справедливое для водородоподобных примесей замещения. Однако и здесь при прогнозе, будет ли та или иная примесь амфотерной в полупроводниковом соединении, возникают определенные трудности, так как упомянутое выше сходство оказалось лишь внешним, поскольку тщательное исследование показало, что амфотерные примеси четвертого типа не являются, строго говоря, водородоподобными.

Вопросы о том, какие именно примеси будут проявлять электрическую активность в том или ином полупроводнике и какие из них окажутся амфотерными, требуют специального анализа. К сожалению, ответы на эти вопросы нельзя получить исходя из простых представлений о свойствах примесного атома, основанных на близости геометрических (атомных или ионных радиусов) и электрохимических (электроотрицательностей) характеристик примесного атома и той кристаллохимической позиции, которую он занимает в полупроводнике. Ни тот, ни другой критерий не могут быть использованы для атомов переходных металлов (случай амфотерных узельных и амфотерных междоузельных центров) из-за неприменимости представления о радиусах и электроотрицательностях, как о постоянных атомных характеристиках [30]. Эти критерии оказываются неприменимы и для амфотерных диссоциативных примесей по тем же причинам, что и в предыдущем случае. Прогнозирование проявления амфотерных примесных центров всех типов в полупроводниках возможно только на основе строгой теории о узельной и междоузельной растворимости примесей в полупроводниках. Современному состоянию этой проблемы, различным подходам к ее решению посвящена монография [31].

К электрически неактивным примесям в полупроводниках относятся изоэлектронные примеси. Довольно часто электрически неактивными примесями в полупроводниках бывают такие примеси, как водород, кислород, азот, а также примеси некоторых элементов, размещающиеся в междоузлиях кристаллической решетки полупроводника. Примеси подобного рода не изменяют концентрацию носителей заряда, однако могут влиять на времена жизни свободных носителей заряда и вносить вклад в излучательную рекомбинацию.

Изоэлектронные примеси — это примесные атомы из той же подгруппы таблицы Менделеева, что и атомы основного элементарного полупроводника или компонента соединения, имеющие с ними одинаковое число валентных электронов и размещающиеся в узлах решетки основного вещества (например, олово в германии и кремнии; висмут в фосфиде галлия).

3.2.1. Примеси в элементарных полупроводниках

Поведение примесей в элементарных полупроводниках рассмотрим на примере их поведения в германии и кремнии. Напомним, что главным образом это поведение определяется положением примесей в периодиче-

Таблица 3.2. Термическая энергия ионизации примесей элементов VA, IIIA подгрупп и Li в Si и Ge [32].

Кремний				Германий			
Доноры		Акцепторы		Доноры		Акцепторы	
Атом	E_i , эВ	Атом	E_i , эВ	Атом	E_i , эВ	Атом	E_i , эВ
Li	0.033	B	0.045	P	0.0120	B	0.0104
P	0.044	Al	0.057	As	0.0127	Al	0.0102
As	0.049	Ga	0.065	Sb	0.0096	Ga	0.0108
Sb	0.039	In	0.16			In	0.0112
Bi	0.069	Tl	0.26			Tl	0.01

ской системе элементов.

Электрически активные примеси

Элементы IIIA и VA подгрупп встраиваются в решетку германия и кремния, преимущественно замещая атомы основного вещества, и при этом ведут себя в соответствии с своей валентностью. Атом элемента VA подгруппы отдает четыре валентных электрона на образование химической связи, а один его электрон может быть переведен в зону проводимости. Атом элемента IIIA подгруппы отдает три валентных электрона на образование химической связи и может присоединить один электрон, что приведет к образованию дырки в валентной зоне. Таким образом, элементы VA подгруппы ведут себя как простые доноры, а элементы IIIA подгруппы как простые акцепторы, образуя мелкие энергетические уровни в запрещенной зоне. Термически определенные значения энергии ионизации примесей элементов VA и IIIA подгрупп, а также лития в слаболегированных Si и Ge приведены в табл. 3.2.

Из таблицы следует, что энергии ионизации примесей IIIA и VA подгрупп в германии отличаются друг от друга существенно меньше, чем в кремнии. В германии энергии ионизации примесей как из одной, так и из разных подгрупп мало отличаются друг от друга и близки к значению 0.01 эВ, предсказываемому водородоподобной моделью для примесных атомов [6] и [32]. Небольшое различие в энергиях ионизации разных примесей в этом случае показывает, что для примесей IIIA и VA подгрупп в германии водородоподобная модель дает неплохое приближение. В кремнии же ситуация иная: наблюдается большая разница в энергиях ионизации примесей даже из одной и той же подгруппы, чем в германии для примесей из разных подгрупп (см. табл. 3.2). Таким образом,

в кремнии более сильно проявляются отклонения в значениях энергии ионизации примесей из IIIA и VA подгрупп от значений, предсказываемых водородоподобной моделью, для этих примесей. Особенно это заметно в случае In и Tl, которые являются, по существу, глубокими акцепторными примесями. Это единственные стабильные одноуровневые примеси из известных глубоких примесей. Тем не менее, и в этом случае (за исключением In и Tl) энергии ионизации примесей небольшие, поэтому при температурах, близких к комнатной и выше, атомы этих примесей ионизованы почти полностью. Наличие отклонений в значениях энергии ионизации примесей из IIIA и VA подгрупп в Si и Ge от значений, предсказываемых водородоподобной моделью, обусловлено тем, что на межатомных расстояниях потенциал, создаваемый примесным ионом, существенно отличается от потенциала точечного заряда и зависит от химической природы примеси. Эта короткодействующая часть примесного потенциала создает дополнительное по отношению к величине энергии ионизации, предсказываемой водородоподобной моделью, смещение примесного уровня, называемое химическим сдвигом. Благодаря химическому сдвигу примесные уровни разных примесей отличаются друг от друга.

Как уже было сказано выше, элементы IIIA и VA подгрупп встраиваются в германии и кремнии преимущественно в узлы кристаллической решетки, однако некоторые из них могут также занимать и междоузельное положение. Очевидно, что расположенные в междоузлиях примесные атомы, ведут себя иначе, чем примесные атомы в узлах решетки. В этом случае концентрация основных носителей может отличаться от концентрации примеси в температурном интервале, соответствующем полной ионизации примеси. Например, примесь фосфора в кремнии преимущественно встраивается в узлы решетки, но часть атомов размещается в междоузлиях. Атомы фосфора в междоузлиях, по-видимому, электрически неактивны, поэтому концентрация электронов в Si, легированном P, в температурном интервале, соответствующем полной ионизации примеси, немного ниже, чем полная концентрация вводимого в кремний фосфора.

Практическая значимость примесей, дающих мелкие уровни энергии в запрещенной зоне полупроводника, заключается в том, что они имеют невысокие энергии ионизации (например, в германии при температурах ~ 10 К атомы этих элементов практически полностью ионизованы) и хорошо растворяются в легируемом материале (имеют высокие коэффициенты растворимости, см. гл. 7, и малые коэффициенты диффузии, см. гл. 8), поэтому, вводя их в полупроводники, можно в широких пре-

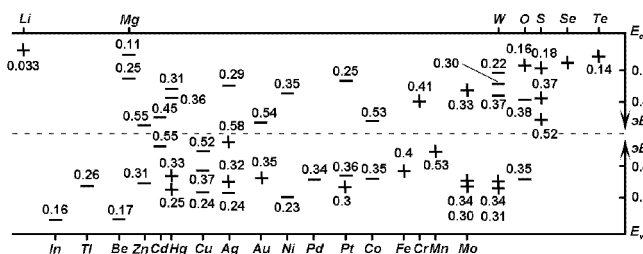


Рис. 3.16. Энергетические уровни примесей в Si. Символы + и — обозначают донорные и акцепторные уровни соответственно [32].

делах менять концентрацию электронов и дырок в них (от собственной до 10^{21} см^{-3}) и их проводимость. Подобные примеси называются легирующими.

Поведение примесей в кремнии и германии усложняется по мере «удаления» их в периодической системе от элементов IVA подгруппы. Большинство примесей из I, II, VI групп и переходные металлы в слабелегированных кремнии и германии являются примесями замещения, хотя некоторые из них могут размещаться в решетке основного вещества и (или) в междоузлиях. Например, примесь меди в германии. Атомы Cu входят как в узлы, так и в междоузлия в решетке Ge. Магний, кальций, стронций и барий в кремнии и германии также могут размещаться как в узлах, так и в междоузлиях кристаллической решетки [17].

Примеси I, II, VI групп и переходные металлы образуют несколько альтернативных глубоких локальных уровней энергии в запрещенной зоне кремния и германия. Глубокие состояния, как правило, возникают, когда основной вклад в энергию связи дает не электрическое притяжение, ослабленное диэлектрической проницаемостью (водородоподобное приближение), а короткодействующий потенциал, который определяется химической природой примеси. Мелкие донорные состояния можно считать отщепившимися от зоны проводимости, а мелкие акцепторные состояния — от валентной зоны. Глубокие состояния принадлежат в равной мере обеим зонам и могут быть и донорными, и акцепторными. Положение экспериментально определенных энергетических уровней некоторых примесей из этих групп в запрещенной зоне кремния и германия приведены на рис. 3.16 и рис. 3.17, соответственно.

Можно считать, что поведение примесей замещения, образующих глубокие уровни в Si и Ge, соответствует их валентности, если предположить, что в случае элементов I, II групп и переходных металлов (на

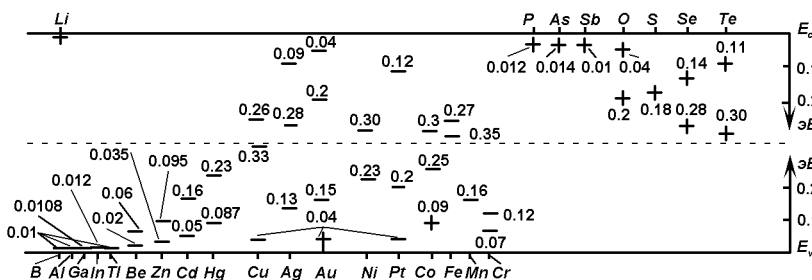


Рис. 3.17. Энергетические уровни примесей в Ge. Символы + и — обозначают донорные и акцепторные уровни соответственно [32].

самой внешней оболочке атомов из этой группы находится по два электрона ns^2) все валентные электроны поступают на образование связей тетраэдрического типа, а в случае элементов VI группы в образовании этих связей участвуют четыре валентных электрона и в зону проводимости могут быть переведены два валентных электрона. В решетке полупроводника в ионизованном состоянии эти примеси присутствуют в виде многозарядных ионов. Максимальное число локальных энергетических уровней, создаваемых примесью замещения в элементарном полупроводнике, как правило, равно разности между числом валентных электронов атома кремния (германия) и атома примеси. Например, Zn (IIB группа) в Ge (см. рис. 3.17). Валентность атомов Zn равна двум, для образования тетраэдрических связей в Ge этой примеси недостает еще двух электронов, что приводит к их захвату из валентной зоны и образованию двух дырок в ней, и поэтому атомы цинка в германии являются двухзарядными акцепторами. Аналогичные рассуждения показывают, что железо (VIII группа, валентность равна двум) в германии является двухзарядным акцептором, а медь (I группа, валентность равна 1) в Si является трехзарядным акцептором. Однако установлено, что Li (I группа) в решетке кремния и германия размещается в междоузлиях, что связано с его малым размером, и образует в запрещенной зоне один мелкий донорный уровень. Теллур Te (VI группа) имеет шесть валентных электронов, только четыре из которых принимают участие в образовании тетраэдрических связей в Ge, поэтому атомы теллура в германии образуют два донорных уровня. Эти рассуждения в основном подтверждены экспериментальными данными, приведенными на рис. 3.16 и рис. 3.17. Из этих рисунков и табл. 3.2 следует, что глубина залегания примесей II и III групп, как правило, увеличивается с ростом атомного номера. Аналогич-

ная тенденция наблюдается и для Cu, Ag и Au (IB подгруппа).

Следует отметить, что поведение переходных металлов из VII, VIII групп и с незаполненной $5d$ -оболочкой в Si и Ge изучено менее подробно, чем примесей остальных групп (см. рис. 3.16 и рис. 3.17). Чаще всего они образуют несколько глубоких уровней в запрещенной зоне. Как уже отмечалось выше, если эти примеси размещаются в узлах кристаллической решетки Si или Ge, то их поведение, как правило, соответствует их валентности — они являются двухзарядными акцепторами. Однако так бывает не всегда. Например, примесные атомы Fe и Mn в Ge размещаются узлах и образуют в соответствии со своей валентностью по два акцепторных уровня; однако в Si, занимая стабильные положения в узлах при наличии в решетке вакансий, эти примеси образуют донорные уровни [32] (см. рис. 3.16 и рис. 3.17). Кроме того, примеси указанных переходных металлов могут размещаться как в узлах, так и в междоузлиях решетки полупроводника. Например, установлено, что никель в кремнии размещается как в узлах, так и в междоузлиях, причем основная часть Ni попадает в междоузельное положение. В узловом положении Ni проявляет электрическую активность и является двухзарядным акцептором, в то время как электрическая активность междоузельного никеля не обнаружена [31] и [32]. Последнее связано, по-видимому, с образованием преципитатов при охлаждении материала [32]. Еще одной особенностью переходных металлов является то, что они могут быть амфотерными примесями в Si и Ge.

Амфотерными примесями в кремнии и германии, как уже обсуждалось выше, могут быть атомы с незаполненной d -оболочкой, способные быть донорами или акцепторами в одной из кристаллических позиций, или те, которые могут размещаться как в узлах, так и в междоузлиях кристаллической решетки и проявлять при этом донорные и акцепторные свойства в зависимости от своего расположения. К настоящему времени экспериментально установлены следующие данные [31].

Примесные атомы переходных металлов, находящиеся в полупроводнике в неионизованном состоянии, благодаря своей электрической активности способны под внешним воздействием отдавать электроны с d -оболочки в зону проводимости или захватывать на d -орбитали валентные электроны кристалла, образуя в валентной зоне дырку. Отдача или захват электронов идут с изменением электронной конфигурации d -оболочек, то есть с изменением их заполнения электронами. Исходная конфигурация d -атома определяется его кристаллохимическим состоянием. Она зависит от типа кристалла и той позиции в кристаллической решетке, которую занимает примесный атом. Правила, по которым происходит изменение электронной конфигурации d -атома при его размещении в кристалле полупроводника, определяются моделью, первоначально установленной на основе экспериментальных данных Людвигом и Вудбери [33], а затем теоретически обоснованной Ройциным и Фирштейном [34], и поэтому сокращенно называемой РФЛВ моделью. Согласно этой моде-

ли, развитой применительно к кремнию, принимается, что при замещении узлового атома в кристаллической решетке электроны $4s$ -оболочки и часть недостающих (до 4) электронов $3d$ -оболочки примесного d -атома идут на образование ковалентной связи с четырьмя ближайшими атомами кремния; в случае размещения примесного d -атома в междоузлии электроны из его $4s$ -оболочки отталкиваются электронами окружающих атомов кремния и переходят на d -оболочку, то есть в четырехвалентных кристаллах кремния и германия имеет место следующая перестройка электронной структуры d -атома: $3d^n 4s^m$ (свободный атом) $\rightarrow 3d^{n-(4-m)}$ — узел; $3d^n 4s^m$ (свободный атом) $\rightarrow 3d^{n+m}$ — междоузлие.

Большое число экспериментальных данных подтвердило применимость этой модели и для бинарных полупроводников $A^{III}B^V$ и $A^{II}B^{VI}$. Поэтому в общем виде результат модели можно представить в виде схемы: $3d^n 4s^m$ (свободный атом) $\rightarrow 3d^{n-(V-m)}$ — узел; $3d^n 4s^m$ (свободный атом) $\rightarrow 3d^{n+m}$ — междоузлие, где V — валентность замещаемого атома кристалла.

В основе модели РФЛВ лежат представления теории кристаллического поля (ТКП), предполагающей эквивалентность спектроскопических единиц, описывающих электронные термы, и координационных полиэдров, составляющих структуру кристалла. Полиэдр состоит из отрицательно заряженных ионов-лигандов (например, атомов кремния), находящихся в его вершинах, и расположенного в центре полиэдра положительно заряженного иона d -металла. В ТКП пренебрегается электронной структурой лигандов, то есть лиганды отождествляются с точечными электрическими зарядами, и их роль сводится только к созданию электрического кристаллического поля. Симметрия кристаллического поля определяется симметрией координационных полиэдров, составляющих структуру рассматриваемого кристалла. Для полупроводников A^{IV} , $A^{III}B^V$ и $A^{II}B^{VI}$ следует рассматривать только два типа симметрии: тетраэдрическую и для учета второй координационной сферы — октаэдрическую. Механизм преобразования электронных термов d -иона в кристаллическом поле рассмотрен в [30]. Подобная модель позволяет получить общую картину поведения всех уровней d -иона в кристаллическом поле: качественный характер расщепления уровней, их взаимное расположение и относительные энергетические зазоры между ними.

Однако модель кристаллического поля не дает ответа на вопросы: какова «привязка» системы расщепленных уровней к краям разрешенных зон и какой характер, донорный или акцепторный, имеют энергетические основные состояния образуемых уровней. Для ответа на эти вопросы необходимо рассмотрение примесного потенциала d -атома с учетом его резонансного характера и с учетом межэлектронных взаимодействий [31]. Результат этого рассмотрения позволяет определить правильное положение d -уровней в энергетическом спектре полупроводника и выяснить, у каких d -примесей следует ожидать проявления амфотерности. Было установлено, что такими примесями могут быть только те, у которых оба уровня, соответствующие переходам из исходного (неионизованного относительно решетки) состояния примесного атома d^n в ионизованное состояние с передачей электрона в зону проводимости (донорный переход) и в ионизованное состояние с захваченным электроном (акцепторный переход) оказываются в запрещенной зоне полупроводника. Для того, чтобы в запрещенной зоне могли «разместиться» оба таких уровня, требуется достаточно широкая запрещенная зона и большая степень перемешивания p -орбиталей электронов валентной зоны и d -орбиталей электронов примеси. Расчетные данные показали, что Co, Cu и Ni, находясь в узлах решетки Si, дают уровни в запрещенной зоне, а уровни Fe, Mn, Cr, V и Ti соответствуют их междоузельному расположению в Si. Из этих же данных следует, что амфотерность узельного типа следует ожидать только для примеси Co в Si, а в соединениях $A^{III}B^V$ амфотерностью обладают лишь примесь хрома в GaAs и примесь ванадия в GaP.

Из экспериментов по диффузионному насыщению образцов кремния

примесью кобальта (см. гл. 8) было установлено, что концентрации уровней амфотерного кобальта в *n*- и *p*-типе образцов близки, и отсюда следует вывод о том, что они принадлежат одному и тому же примесному состоянию кобальта. Кроме того, известно, что примесь кобальта в кремнии размещается как в узлах, так и в междоузлиях кристаллической решетки. Экспериментальные данные по растворимости электрически активного Со и его полной растворимости (см. гл. 7) указывают (разница на два порядка), что амфотерную составляющую кобальта, по-видимому, следует приписать узельному состоянию Со, имеющему в неионизованном относительно решетки состоянии электронную конфигурацию d^5 . При захвате электрона в *n*-Si амфотерный центр Со перейдет в состояние d^6 , а в *p*-Si при захвате дырки — в d^4 . Именно срединное заполнение *d*-оболочки неионизованного центра Со (d^5) и определяет его «неустойчивое» состояние, способное практически равновероятно принимать или отдавать электрон. Таким образом, атомы кобальта в кремнии оказываются амфотерными примесями узельного типа, хотя их растворимость в этом полупроводнике и невелика.

Примеси титана, ванадия, марганца в Si являются междоузельными амфотерными примесями. Относительно скандия и ртути вопрос до сих пор остается открытым.

Перейдем к обсуждению диссоциативных амфотерных примесей в Si и Ge. В качестве примера подробно рассмотрим наиболее изученную примесь этого типа: медь в германии. Большим числом исследований установлено, что атомы меди в германии растворяются в узлах и междоузлиях в сравнимых количествах, причем концентрация меди в обеих кристаллохимических позициях зависит от температуры. С этой особенностью связан сложный характер диффузии меди в этом материале (см. подробнее гл. 8). Зарядовое состояние примеси Cu в Ge можно прогнозировать, исходя из общей модели РФЛВ поведения *d*-примесей в полупроводниках (см. выше). Из нее следует, что размещаясь в узлах кристаллической решетки Ge, примесные атомы меди должны быть трехзарядными акцепторами; атомы меди в междоузлиях могут находиться только в состояниях Cu^0 и Cu^+ . Эта ситуация действительно реализуется на опыте. Экспериментально подтверждено, что атомы меди в узловом положении образуют три акцепторных уровня (см. рис. 3.17). Присутствие меди в междоузлиях кристаллической решетки германия доказывается многочисленными опытами по ее диффузии (см. гл. 8) и по кинетике распада твердого раствора Ge(Cu), однако никаких энергетических уровней в запрещенной зоне германия, связанных с Cu_i , не обнаружено. В то же время опыты по определению зарядового состоя-

нения реализуется только для тех атомов, которые имеют промежуточную валентность по сравнению с валентностями компонентов соединения. Наиболее наглядно это проявляется для элементов четвертой группы периодической системы в полупроводниках $A^{III}B^V$. В этом случае атом C^{IV} , размещаясь в катионном узле A^{III} , проявляет донорные свойства, а занимая анионный узел B^V — акцепторные. При этом проявляется известное правило валентности $\Delta V = \pm 1$, справедливое для водородоподобных примесей замещения. Однако и здесь при прогнозе, будет ли та или иная примесь амфотерной в полупроводниковом соединении, возникают определенные трудности, так как упомянутое выше сходство оказалось лишь внешним, поскольку тщательное исследование показало, что амфотерные примеси четвертого типа не являются, строго говоря, водородоподобными.

Вопросы о том, какие именно примеси будут проявлять электрическую активность в том или ином полупроводнике и какие из них окажутся амфотерными, требуют специального анализа. К сожалению, ответы на эти вопросы нельзя получить исходя из простых представлений о свойствах примесного атома, основанных на близости геометрических (атомных или ионных радиусов) и электрохимических (электроотрицательностей) характеристик примесного атома и той кристаллохимической позиции, которую он занимает в полупроводнике. Ни тот, ни другой критерий не могут быть использованы для атомов переходных металлов (случай амфотерных узельных и амфотерных междоузельных центров) из-за неприменимости представления о радиусах и электроотрицательностях, как о постоянных атомных характеристиках [30]. Эти критерии оказываются неприменимы и для амфотерных диссоциативных примесей по тем же причинам, что и в предыдущем случае. Прогнозирование проявления амфотерных примесных центров всех типов в полупроводниках возможно только на основе строгой теории о узельной и междоузельной растворимости примесей в полупроводниках. Современному состоянию этой проблемы, различным подходам к ее решению посвящена монография [31].

К электрически неактивным примесям в полупроводниках относятся изоэлектронные примеси. Довольно часто электрически неактивными примесями в полупроводниках бывают такие примеси, как водород, кислород, азот, а также примеси некоторых элементов, размещающиеся в междоузлиях кристаллической решетки полупроводника. Примеси подобного рода не изменяют концентрацию носителей заряда, однако могут влиять на времена жизни свободных носителей заряда и вносить вклад в излучательную рекомбинацию.

Изоэлектронные примеси — это примесные атомы из той же подгруппы таблицы Менделеева, что и атомы основного элементарного полупроводника или компонента соединения, имеющие с ними одинаковое число валентных электронов и размещающиеся в узлах решетки основного вещества (например, олово в германии и кремнии; висмут в фосфиде галлия).

3.2.1. Примеси в элементарных полупроводниках

Поведение примесей в элементарных полупроводниках рассмотрим на примере их поведения в германии и кремнии. Напомним, что главным образом это поведение определяется положением примесей в периодиче-

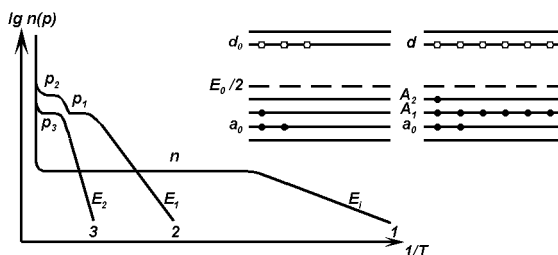


Рис. 3.18. Температурные зависимости концентрации электронов n в исходном германии и дырок p в $\text{Ge}(\text{Zn})$; E_i — энергия ионизации остаточных доноров; E_1 и E_2 — энергии ионизации первого и второго акцепторных уровней цинка, соответственно. На вставке — энергетические диаграммы $\text{Ge}(\text{Zn})$ с рабочим уровнем A_1 (слева) и $\text{Ge}(\text{Zn})$ с рабочим уровнем A_2 , $N_{\text{Zn}} < N_d < 2N_{\text{Zn}}$ (справа).

ственный ход температурных зависимостей концентрации электронов в исходном образце ($N_d^* = N_{d0} - N_{a0}$) и концентрации дырок в том же образце после легирования его цинком при выполнении условия $N_{\text{Zn}} \gg N_d^*$. В отличие от кривой $n(T)$ (кривая 1 на рис. 3.18), содержащей лишь одно широкое плато в области полной ионизации мелких примесных центров и область их неполной ионизации с $E_i \cong 0.01$ эВ, где E_i — энергия активации примеси, кривая $p(T)$ (кривая 2 на рис. 3.18) в легированном образце содержит два плато и область неполной ионизации примесных центров с энергией активации $E_1 = 0.03$ эВ. Соотношение $p_2/p_1 \cong 2$ позволяет считать, что, во-первых, оба плато обусловлены одной примесью, а, во-вторых, оба уровня — акцепторные. В самом деле, поскольку $N_{\text{Zn}} \gg N_d^*$, то уровень A_1 скомпенсирован очень слабо, а уровень A_2 — совсем не скомпенсирован. В таких состояниях только акцепторные уровни могут дать указанное выше соотношение p_2/p_1 , причем ясно, что $p_1 \cong N_{\text{Zn}}$, а $p_2 \cong 2N_{\text{Zn}}$.

Если легирование производится с целью выделения в качестве «рабочего уровня» A_1 , то Zn и заданная компенсирующая донорная примесь (например, сурьма) вводятся в соотношении $N_d^* < N_d < N_{\text{Zn}}$. Практически обе эти примеси вводятся в расплав одновременно с учетом их коэффициентов разделения. Из $p(T)$ (кривая 2 на рис. 3.18) определяются концентрации основной и компенсирующей примесей, а также степень компенсации рабочего уровня:

$$\begin{aligned} p_1 &= N_{\text{Zn}} - N_d; \quad p_2 = 2N_{\text{Zn}} - N_d; \quad N_{\text{Zn}} \cong p_2 - p_1; \\ N_d &= N_{\text{Zn}} = p_2 - 2p_1; \quad N_d/N_{\text{Zn}} = (p_2 - 2p_1)/(p_2 - p_1). \end{aligned} \quad (3.7)$$

Получение кристаллов германия с частично скомпенсированным уровнем A_1 необходимо для изготовления детекторов ИК-излучения с красной границей чувствительности $\hbar\nu \geq 0.03$ эВ ($\lambda \leq 30$ мкм). Область рабочих температур таких приборов $T \leq 50$ К.

Для выделения в качестве рабочего уровня A_2 Zn (детекторы излучений с красной границей чувствительности $\hbar\nu \geq 0.09$ эВ) необходимо введение кроме цинка мелкой донорной примеси в количестве $N_{\text{Zn}} < N_d < 2N_{\text{Zn}}$. Тогда уровень A_1 окажется полностью скомпенсированным, таким образом, исключенным из процессов ионизации дырок в валентную зону, а уровень A_2 будет ионизоваться в соответствии с величиной E_2 и степенью его компенсации. В области полной ионизации уровня, то есть на плато $p(T)$ (кривая 3 на рис. 3.18)

$$p_3 = 2N_{\text{Zn}} - N_d, \quad (3.8)$$

а в области неполной ионизации уровня, когда $p(T) \ll p_3$,

$$p(T) = \frac{(2N_{Zn} - N_d)N_v}{g(N_d - N_{Zn})} \exp\left(-\frac{E_2}{kT}\right). \quad (3.9)$$

Здесь первая предэкспоненциальная дробь называется коэффициентом компенсации уровня A_2 , а разности в числителе и знаменателе есть концентрации некомпенсированных $N_{A_2}^0 = N_{Zn}^-$ и скомпенсированных $N_{A_2}^- = N_{Zn}^-$ уровней, соответственно. Необходимо обратить внимание на то, что в системе с несколькими энергетическими уровнями при ионизации уровня с номером большим единицы понятия «зарядовое состояние уровня» и «зарядовое состояние атома» примеси становятся различными. Использование (3.8) дает возможность из измеренных $p(T)$ (кривая 3 на рис. 3.18) определить концентрации основной и компенсирующей примесей и степень компенсации уровня $N_{A_2}^-/N_{A_2}^0 = N_{Zn}^-/N_{Zn}^-$. Однако, в отличие от кривой 2 (рабочий уровень A_1), в этом случае для нахождения указанных величин необходимо определить $p(T)$ в области экспоненциального спада (3.9), энергию ионизации уровня, а также знать фактор вырождения уровня g . Кроме того, для определения самой концентрации носителей из измеренной постоянной Холла необходимо знать связывающий их холл-фактор.

Важной особенностью двухуровневой системы, дающей в температурной зависимости концентрации носителей два плато, является возможность определения концентраций основной и компенсирующей примеси без использования области неполной ионизации примеси. Тогда, располагая зависимостью $p(1/T)$ в области частичной ионизации примеси и найденной из нее энергией активации уровня A_1 , можно найти фактор вырождения уровня A_1 . Такие измерения были проведены в p -германии, легированном ртутью, а также золотом. Полученные значения $g \cong 2$ показали, что вырождение донорного и первого акцепторного уровня ртути отражает спиновое вырождение валентной зоны.

Для определения холл-фактора r_x , связывающего постоянную Холла R_H с концентрацией носителей соотношением $n(p) = r_x/eR_H$, необходимо измерение зависимости постоянной Холла от магнитного поля. Известно, что в «слабом» магнитном поле B , когда выполняется условие $\mu^2 B^2 \ll 1$, холл-фактор определяется типом рассеяния носителей заряда в решетке [6]. В области сильных полей, удовлетворяющих условию $\mu^2 B^2 \gg 1$, холл-фактор $r_x = 1$. Таким образом, измерения $R_H(B)$ в указанной области магнитных полей дают возможность и измерения $r_x(B)$, и использования правильных значений r_x при измерениях в заданном магнитном поле. Измерения $R_H(B)$ необходимы и в случаях прецизионных определений примесных параметров, и в тех случаях, когда речь идет об исследованиях неоднородных кристаллов — в последнем случае холл-факторы могут быть искажены неоднородностью.

Электрически неактивные примеси

К электрически неактивным примесям в германии и кремнии относятся изоэлектронные примеси, а также такие примеси, как водород, кислород, азот.

Для германия и кремния изоэлектронными примесями являются элементы из IVA подгруппы — Pb, Sn и C. Примеси подобного рода не изменяют концентрацию носителей заряда, однако могут влиять на времена жизни свободных носителей заряда и вносить вклад в излучательную рекомбинацию.

Водород, кислород и азот, не будучи изоэлектронными, как правило, электрически неактивны в элементарных полупроводниках. Введение этих примесей в Ge и Si не приводит к увеличению концентрации свободных носителей заряда, хотя они и обладают высокой растворимостью (см. гл. 7). Для водорода одной из причин подобного поведения является его способность к размещению в междоузлиях в виде нейтральных атомов или молекул H_2 из-за его малого размера. Введенный обычными способами легирования азот образует нейтральные комплексы Ge_3N_4 , Si_3N_4 (нитрида германия (кремния)). Однако путем ионного легирования с последующим отжигом удается вводить азот в Si так, что он размещается атомарно в узлах решетки, образуя при этом донорный уровень. Наличие примеси азота в решетке кремния приводит к появлению характерной полосы инфракрасного поглощения около 10.6 мкм [32].

Особое место среди примесей в кремнии и германии занимает кислород, который является остаточной, в большинстве случаев вредной примесью; его концентрация зависит от способа получения кристалла. О поведении кислорода в Si и Ge известно, что он в решетке основного вещества присутствует либо в атомарном виде, либо образует комплексы типа $Si(Ge)O_x$. При этом, если атомы кислорода размещаются в междоузлиях, то они, по-видимому, электрически неактивны. Донорными же свойствами обладают некоторые комплексы $Si(Ge)O_x$. Больше всего кислорода обычно содержат монокристаллы кремния (германия), выращенные из кварцевых контейнеров (содержание кислорода в кремнии составляет $\sim 10^{18} \text{ см}^{-3}$, см. гл. 5).¹⁷ Обычно при выращивании из расплава большая часть атомов кислорода размещается в междоузлиях и образует электрически неактивные комплексы SiO_2 , хотя в малых концентрациях могут образовываться и электрически активные комплексы более высокого порядка. Процессы, связанные с нагревом кристалла (термообработка, диффузия), могут приводить к перераспределению по концентрации различных типов кремний(германий)-кислородных комплексов. Это перераспределение происходит в тех случаях, когда распределение кислородных комплексов неравновесно (твердый раствор $Si(O)$ находится в пересыщенном состоянии при температуре обработки). Изменение концентрации электрически активных комплексов приводит к изменению электрических свойств кристалла. В частности, при низкотемпературной (300–500°C) обработке в кристаллах Si образуются термодоноры SiO_4 в измеримых концентрациях, которые устойчивы при 430°C; комплексы

¹⁷Концентрацию кислорода в кремнии обычно измеряют по интенсивности инфракрасного поглощения на длине волны 9 мкм, соответствующей характеристическим колебаниям связи Si–O.

более низкого порядка при этих температурах распадаются. При повышении температуры выше 500°C комплексы SiO_4 распадаются и могут появляться включения второй фазы, которые в свою очередь распадаются при температурах прогрева материала выше 1200°C .

3.2.2. Примеси в полупроводниковых соединениях

Рассмотрим поведение примесей в полупроводниковых соединениях на примере соединений типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$. Поведение примесей в соединениях типа $\text{A}^{\text{III}}\text{B}^{\text{V}}$ так же, как и в элементарных полупроводниках, определяется положением примеси в периодической системе, однако оно становится более сложным из-за усложнения строения основного вещества. В соединениях возрастает число различных позиций, которые могут занимать примесные атомы.

В соединениях, образованных сравнительно сильно различающимися по электроотрицательности элементами, примесные атомы, при прочих равных условиях, замещают те атомы соединения, для которых разность электроотрицательностей между атомом компонента соединения и примесью меньше. В тех случаях, когда величина ΔX атомов компонентов соединения невелика, а примесь имеет промежуточную электроотрицательность, определяющую роль играет геометрический фактор: разность ковалентных радиусов примеси и компонентов соединения.

На рис. 3.19 приведены положения экспериментально определенных энергетических уровней примесей из разных групп периодической таблицы в запрещенной зоне слаболегированного GaAs [32].

Элементы II и VIA подгрупп в соединениях $\text{A}^{\text{III}}\text{B}^{\text{V}}$, как правило, замещаются в узлах решетки, причем атомы элементов второй группы замещают атомы элемента-катиона А, а атомы элементов шестой подгруппы — элемента-аниона В. При этом в соответствии с правилом валентности элементы II группы являются простыми акцепторами, а элементы VIA подгруппы — простыми донорами. Например, цинк или кадмий в InP и GaAs замещают In или Ga, соответственно, и образуют мелкий акцепторный уровень (каждый) в этих соединениях. Селен или теллур в GaP и GaAs замещают P или As, соответственно, и образуют мелкий донорный уровень (см. рис. 3.19). Примеси этих групп в соединениях $\text{A}^{\text{III}}\text{B}^{\text{V}}$, как правило, являются легирующими примесями. Особое место среди примесей VIA подгруппы в соединениях $\text{A}^{\text{III}}\text{B}^{\text{V}}$ занимает кислород. Например, кислород в GaAs. Экспериментально найдено, что кислород в GaAs замещает As и является глубоким донором [32].

Легирование полупроводниковых соединений примесями, образующи-

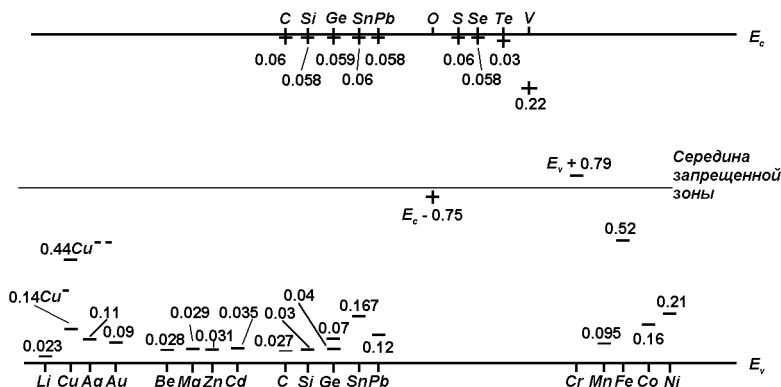


Рис. 3.19. Энергетические уровни примесей в GaAs. Символы + и — обозначают донорные и акцепторные уровни соответственно [32].

ми мелкие донорные и акцепторные уровни, используется для изменения в широких пределах концентрации носителей заряда и типа проводимости в них. Например, легирование арсенида галлия цинком используется при создании диодов и транзисторов.

Элементы IIIA и VA подгрупп, располагаясь преимущественно в соответствующих узлах кристаллической решетки соединений $A^{III}B^V$, являются электрически неактивными примесями. Это изоэлектронные примеси; они не влияют на концентрацию свободных носителей заряда, но влияют на их времена жизни, то есть на рекомбинацию (например азот, висмут, сурьма в GaAs или GaP).

Элементы IB подгруппы (Cu, Ag, Au), замещая катионы в соединениях $A^{III}B^V$, согласно правилу валентностей должны быть двухзарядными акцепторами. В некоторых случаях это правило находит экспериментальное подтверждение. Например, для примесных атомов Cu в GaAs. Атомы меди могут входить в решетку GaAs как в узлы, так и в междоузлия. Размещаясь в узлах Ga, атомы Cu образует два акцепторных уровня в запрещенной зоне GaAs, а входя в междоузлия, образуют однозарядные доноры [32]. Однако для Ag и Au в GaAs обнаружено только по одному акцепторному уровню (см. рис. 3.19). Элементы IA подгруппы в соединениях $A^{III}B^V$ могут размещаться как в узлах, так и в междоузлиях кристаллической решетки полупроводника, проявляя при этом разную электрическую активность. Например, установлено, что литий в InSb размещается в междоузлиях и ведет себя как простой донор, а в GaAs

размещается как в междоузлиях, так и в узлах. Размещаясь в междоузлиях, литий ведет себя как простой донор, а размещаясь в узлах — как двухзарядный акцептор. Однако при охлаждении температуры от температуры легирования (выше 500 К) до комнатной происходит компенсация, приводящая к образованию высокоомных кристаллов *p*-типа и в итоге наблюдается только один акцепторный уровень для Li в GaAs (см. рис. 3.19).¹⁸

Поведение переходных металлов из VI, VII и VIII групп в соединениях $A^{III}B^V$, как и в случае элементарных полупроводников, изучено менее подробно, чем элементов уже рассмотренных групп. Указанные переходные металлы в соединениях $A^{III}B^V$ ведут себя преимущественно как глубокие акцепторы. Для этих примесей экспериментальные данные качественно согласуются с правилом валентностей, если они размещаются в узлах металлической подрешетки. Например, в GaAs (см. рис. 3.19) Mn, Co и Ni, обладающие стабильной валентностью +2, при замещении ими Ga образуют однозарядные акцепторы. Хотя стабильные валентности Fe и Cr равны +3, они могут иметь также валентность +2 и, следовательно, встраиваясь в узлы галлия, быть в GaAs однозарядными акцепторами. Ванадий обычно обладает стабильной валентностью +4, и это согласуется с его ролью донора в GaAs.

Амфотерными примесями в соединениях $A^{III}B^V$ могут быть атомы с незаполненной *d*-оболочкой, способные быть донорами или акцепторами в одной из кристаллических позиций, или те, которые могут размещаться как в узлах, так и междоузлиях кристаллической решетки и проявлять при этом донорные и акцепторные свойства в зависимости от своего расположения, а также атомы IVA подгруппы.

Анализ поведения примесей с незаполненной *d*-оболочкой в соединениях $A^{III}B^V$ показал (см. выше), что амфотерностью обладает лишь примесь хрома в GaAs и ванадия в GaP [31]. Обсудим возможность появления свойства амфотерности у примеси в соединениях $A^{III}B^V$ на примере примеси хрома в GaAs. Многочисленные экспериментальные исследования свидетельствуют о размещении примесных атомов хрома в

¹⁸Компенсация обусловлена влиянием ионов Li^+ . В соответствии с законом действующих масс присутствие междоузельных ионов Li^+ приводит к увеличению концентрации вакансий галлия V^- на несколько порядков. При этом в кристалле могут образовываться дефекты Li^+V^- , Li^- , Li^{2-} , Li^+Li^{2-} , где Li^+ — междоузельный ион, а Li^- и Li^{2-} — замещающие ионы. Поскольку подвижность донорного лития в решетке GaAs высока, то при охлаждении от высокой температуры до комнатной избыточный Li^+ легко образует преципитаты. Избыточные менее подвижные акцепторы при этом остаются, обуславливая наблюдаемую компенсацию, *p*-тип проводимости и экспериментально наблюдаемую энергию активации Li в GaAs.

узлах галлия. В соответствии с моделью РФЛВ неионизованное (относительно решетки) состояние примеси хрома Cr^0 имеет электронную конфигурацию d^3 . Ионизованные состояния соответственно будут d^2 (донор) и d^4 (акцептор). Для того, чтобы Cr был амфотерен в GaAs необходимо, чтобы уровни, соответствующие донорному и акцепторному переходам из неионизованного состояния в ионизованные, располагались в запрещенной зоне. Все три электронные конфигурации, перечисленные выше, действительно, наблюдались в ЭПР при освещении образца с энергией кванта света, равной или большей ширины запрещенной зоны. Данные ЭПР позволяют представить процесс фотоионизации следующим образом. Квант света первоначально ионизует атом $\text{Cr}^{3+}(d^3)$, который при освещении захватывает электрон из валентной зоны и переходит в состояние $\text{Cr}^{2+}(d^4)$. Затем свободная дырка, образованная в валентной зоне, захватывается другим неионизованным центром $\text{Cr}^{3+}(d^3)$, переводит его таким образом в донорное состояние $\text{Cr}^{4+}(d^2)$. При выключении света происходит рекомбинация захваченных носителей и с помощью того же ЭПР можно наблюдать возврат к исходной концентрации центров d^3 .

Обсудим теперь катионно-анионные амфотерные примеси в соединениях $\text{A}^{\text{III}}\text{B}^{\text{V}}$ на примере примесей IV группы в GaAs. Поведение примесей из IV группы зависит от положения, занимаемого ими в решетке соединения GaAs. Рассмотрим некоторые возможные ситуации.

1. Атомы примеси размещаются преимущественно в подрешетке катиона и ведут себя при этом как простые доноры. Примером таких примесей является Sn [31] (рис. 3.19). Олово проявляет амфотерные свойства в наименьшей степени по сравнению с Si и Ge в GaAs. Вплоть до очень высоких концентраций олова образцы GaAs(Sn) обладают n -типом проводимости. Однако в то же время олово имеет глубокий акцепторный уровень в запрещенной зоне, причем, по-видимому, этот уровень обусловлен атомами олова, замещающими As в узлах решетки. Амфотерность олова в арсениде галлия окончательно была установлена после изучения эффекта Мессбауэра на ядрах ^{119}Sn при их имплантации в GaAs.

2. Примесные атомы размещаются преимущественно в подрешетке аниона и при этом ведут себя как простые акцепторы. Примером таких примесей является Pb в GaAs [32]. Свинец образует глубокий акцепторный уровень в запрещенной зоне GaAs (рис. 3.19).

3. Примесь входит в обе подрешетки с близкими концентрациями, амфотерные свойства примеси в этом случае проявляются наиболее ярко. При этом отношение концентраций донорных и акцепторных уровней N_d/N_a зависит от состава кристалла, концентрации самой амфотерной

примеси, температуры и давления летучего компонента при получении. Например, Si и Ge в GaAs [31]. Примесные атомы Si и Ge, размещаясь в узлах галлия, играют роль мелких простых доноров, а размещаясь в узлах мышьяка, — роль акцепторов.

Ярким примером амфотерной примеси, имеющей большое практическое значение при изготовлении светодиодов из GaAs, является примесь кремния. Было выяснено, что при выращивании GaAs из расплава кремний входит в узлы галлия и образует мелкий донорный уровень. Однако при выращивании GaAs из раствора в галлии (см. гл. 6) в условиях низкого давления паров мышьяка кремний размещается преимущественно в узлах мышьяка и образует акцепторный уровень. Использование этих особенностей вхождения примеси кремния в GaAs позволило, выращивая эпитаксиальную пленку методом жидкостной эпитаксии (см. гл. 9) при тщательном контроле температуры роста и условий охлаждения, в одном технологическом процессе получить $p - n$ -переход. Из структур, полученных таким образом, изготавливаются светодиоды с высоким внешним квантовым выходом люминесценции (до 6%) [32].

4. Атомы примеси локализуются в двух соседних узлах решетки соединения, образуя нейтральную донорно-акцепторную пару. Например, при легировании GaAs кремнием появляются донорно-акцепторные пары $[\text{Si}_{\text{Ga}}^+ \text{Si}_{\text{As}}^-]$.

Размещаясь в соединениях $\text{A}^{\text{III}}\text{B}^{\text{V}}$ в соответствии с пп. 3 и 4, элементы IVA подгруппы выступают в качестве частично самокомпенсирующих примесей.

Получение чистых полупроводниковых соединений $\text{A}^{\text{III}}\text{B}^{\text{V}}$ технологически более сложно, чем получение чистых элементарных полупроводников Ge и Si (см. гл. 5). Так, например, и кристаллы, и эпитаксиальные пленки GaAs в зависимости от метода их получения содержат разное количество разных остаточных примесей. В объемных кристаллах GaAs чаще всего обнаруживаются остаточные примеси, образующие мелкие донорные и акцепторные уровни. В тех случаях, когда необходимо получить высокоомный (полуизолирующий) материал для подложек, используется специальное легирование глубокими примесями, например, хромом или кислородом, причем хром является в GaAs глубоким акцептором, а кислород — глубоким донором ($\rho \sim 10^8$ Ом · см при 300 К при легировании хромом и $\rho \sim 10^7$ Ом · см при 300 К при легировании кислородом). Следует отметить, что довольно часто в кристаллах GaAs, получаемых наиболее широко распространенными в промышленности методами, в качестве остаточной примеси присутствует кислород. Если в кристалле имеются еще и мелкие акцепторы, то кислород может

их компенсировать, в результате чего материал станет высокоомным и без специального легирования. Содержание кислорода в GaAs, безусловно, зависит от метода его выращивания и обработки.

В ряде случаев рост эпитаксиальных пленок GaAs может сопровождаться нежелательным легированием глубокими примесями или появлением уровней собственных дефектов. Наличие глубоких примесей в полупроводнике может приводить к захвату носителей заряда и уменьшению их времени жизни. Эти эффекты оказывают влияние на многие характеристики полупроводниковых приборов, изготовленных из этого материала, например, на токи утечки, шумы диодов и транзисторов; они проявляются при переходных процессах во всех транзисторах, диодах и источниках света на основе GaAs. В транзисторах, изготовленных из GaAs, эффекты захвата связываются с примесью кислорода.

Перечисление примесных дефектов в полупроводниках не будет полным, если не упомянуть о возможности комплексообразования. В разделе, посвященном собственным точечным дефектам, уже упоминалось о простых комплексах (это дефекты Френкеля, дивакансии и т. д.). Подобные комплексы могут образовываться и примесными атомами. Так, в GaAs и GaSb, легированных литием (см. выше), образуются комплексы вакансий V_{Ga}^- с междоузельными атомами лития $[Li^+V_{Ga}^-]$, ведущие себя в процессах рассеяния носителей заряда иначе, чем дефекты Li^+ и V_{Ga}^- . При значительных концентрациях Li в GaAs возникают комплексы типа $[Li_i^+Li_{Ga}^{2-}]$, которые ведут себя как однократно заряженные акцепторы. При сильном легировании арсенида галлия Se или Te могут образовываться комплексы или преципитаты, которые влияют на квантовый выход излучения и характер люминесценции. На электрические свойства кремния значительное влияние оказывает комплексообразование между междоузельным примесным кислородом и атомами основного вещества (см. выше). Образование комплексов из простых доноров, приводящих к появлению глубоких уровней в запрещенной зоне, проявляется в появлении сильных избыточных токов в туннельных диодах из Ge при их легировании фосфором или сурьмой.

Глава 4

Фазовые равновесия в полупроводниковых системах. Элементы теории образования фаз

В предыдущих главах мы рассмотрели особенности химической связи и дефекты в полупроводниках. Было показано, что анализируя электронную конфигурацию и величины кристаллохимических параметров взаимодействующих атомов, возникающую в кристалле химическую связь, можно прогнозировать получение полупроводников с заданными параметрами. Обсуждались также свойства и основные принципиальные причины возникновения дефектов (без технологических деталей), их влияние на некоторые параметры полупроводниковых кристаллов. Однако прежде чем перейти к обсуждению методов получения полупроводниковых материалов, нам необходимо познакомиться с термодинамическим подходом к описанию взаимодействия веществ и их фазовых превращений. Это позволит определить условия протекания фазового превращения, в результате которого образуется полупроводник, а также условия получения материала в стабильном состоянии при заданных внешних параметрах. Также необходимо иметь представление о процессе образования кристаллических зародышей и о механизмах их роста. Полученные сведения будут использованы в последующих главах, посвященных обсуждению методов получения полупроводниковых материалов, их очистке, легированию и способов управления ростом кристаллов.

4.1. Фазовые равновесия в полупроводниковых системах

4.1.1. Диаграммы фазового равновесия

Реальные химические вещества, с которыми практически приходится иметь дело, и даже сверхчистые кристаллы элементарных полупроводников Ge и Si всегда содержат остаточные примеси, то есть всегда представляют собой вещества, состоящие из нескольких химических элементов. Взаимодействие химических элементов, образующих данный материал, может быть весьма сложным. Конкретный результат этого взаимодействия зависит от кристаллохимической природы взаимодействующих элементов, их концентрации, а также от внешних факторов — температуры и давления.

Основным средством изображения результатов взаимодействия химических элементов или соединений, образующих данное вещество, являются диаграммы состояния системы. Диаграмма состояния показывает устойчивые состояния, то есть состояния, которые при данных условиях обладают минимумом свободной энергии. Поэтому диаграмма состояния также может называться диаграммой фазового равновесия, так как она показывает, какие при данных условиях существуют равновесные фазы. В соответствии с этим и изменения в состоянии системы, которые отражены на диаграмме, относятся к равновесным условиям, то есть при отсутствии переохлаждения или пересыщения в системе. Однако фазовые превращения не могут происходить в равновесных условиях (см. ниже), поэтому диаграмма состояния представляет собой теоретический случай. Тем не менее, роль диаграмм состояния в понимании характера и результатов взаимодействия различных химических веществ и прогнозировании этих результатов чрезвычайно важна, потому что именно характер взаимодействия определяет свойства получаемого материала. На практике диаграммы состояния используются для рассмотрения превращений при малых скоростях охлаждения или нагрева.

Диаграммой состояния системы называется геометрическое изображение равновесных фазовых состояний одно- или многокомпонентной термодинамической системы как функции параметров, определяющих эти состояния (концентрации, температуры, давления).

Определим некоторые понятия, используемые при описании диаграмм состояния.

Термодинамической системой называется тело макроскопических размеров (совокупность тел), между отдельными частями которого (меж-

ду которыми) возможен теплообмен и диффузия хотя бы одного из компонентов системы и для которого (которых) справедливы начала термодинамики.

Термодинамические системы делятся на *гомогенные* и *гетерогенные*.

Гомогенной называется термодинамическая система, внутри которой нет поверхностей раздела фаз, отделяющих друг от друга части системы, которые различались бы либо по кристаллической структуре, либо по своим физическим и химическим свойствам. *Гетерогенная* система состоит из частей, имеющих либо различную структуру, либо различные физико-химические свойства и отделенных друг от друга поверхностями раздела фаз. Примером гетерогенной системы может являться вода, находящаяся в равновесии с паром.

Фаза — это гомогенная система или система, представляющая собой совокупность одинаковых по кристаллической структуре и физико-химическим свойствам гомогенных систем, отделенных друг от друга поверхностями раздела. В приведенном выше примере фазами являются вода и пар, которые различаются, например, плотностью.

Поверхности раздела фаз представляют собой слои конечной толщины, в которых в направлении от одной фазы к другой изменяется, по крайней мере, один из параметров системы. Поверхности раздела фаз по отношению к граничащим фазам обладают избыточной энергией (энергия поверхностного натяжения).

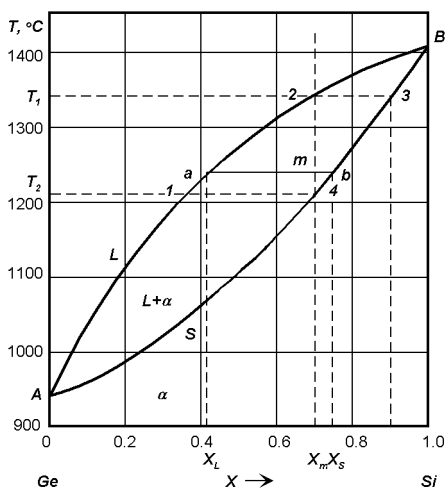
Для твердых тел важнейшим признаком фазы является ее кристаллическая решетка.¹ Каждая твердая фаза обладает своей, только ей присущей кристаллической решеткой, отличающейся от решеток других фаз либо типом, либо параметрами. Твердая кристаллическая фаза может быть получена в виде монокристалла или поликристалла, представляющего собой совокупность зерен или кристаллитов. Различно ориентированные в пространстве кристаллиты поликристалла отделены друг от друга поверхностями раздела в несколько атомных слоев (см. гл. 3). Очевидно, что границы зерен не являются межфазными границами.

Термодинамические системы могут быть одно- и многокомпонентными.

Компонентом системы называется часть системы, количество которой может изменяться независимо от количества других частей. В нашем случае компонентами системы могут быть химические элементы или соединения. Число компонентов системы, вообще говоря, может быть не

¹В принципе, твердая фаза может также быть аморфной или стеклообразной. Обе эти фазы характеризуются отсутствием дальнего порядка в расположении атомов, скорее напоминающая жидкость. Здесь мы будем рассматривать только кристаллические материалы.

Рис. 4.1. Диаграмма состояния системы Ge–Si.



равно числу различных химических элементов в системе. Например, вода (H_2O) состоит из водорода и кислорода, но это однокомпонентная система. На рис. 4.1 и рис. 4.2 показаны диаграммы фазового равновесия двух характерных двухкомпонентных (бинарных) полупроводниковых систем — Ge–Si и InSb–AlSb. Компонентами системы в первом случае являются Ge и Si, а во втором — InSb и AlSb, а не Sb, Al, In, так как количество In и Al в системе зависит от количества Sb, а количество InSb не зависит от количества AlSb. Поэтому *число компонентов системы* — это минимальное число химических веществ, необходимых для образования любой фазы данной системы.

Термодинамически равновесным состоянием системы называют такое ее состояние, при котором параметры этого состояния не меняются с течением времени и в системе отсутствуют потоки любого типа.

Равновесное состояние системы может быть однофазным, двухфазным и многофазным. При смешивании двух или более твердых фаз могут образовываться твердые растворы, соединения и *механические смеси*. Последнее реализуется, если эти фазы не взаимодействуют друг с другом. Фазами, образующими смесь, могут быть элементы, соединения или твердые растворы на их основе, а также аллотропические модификации одного и того же химического элемента (α - и β -олово и т.п.). Максимально возможное число фаз, находящихся в равновесии, определяется правилом фаз Гиббса [17]. Правило фаз устанавливает соотношение меж-

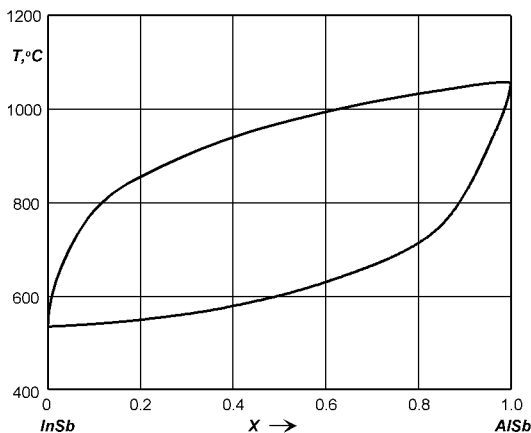


Рис. 4.2. Диаграмма состояния системы InSb–AlSb.

ду числами фаз, компонентов и степеней свободы системы:

$$c = k - f + 2, \quad (4.1)$$

где c — число степеней свободы системы, k — число компонентов системы, f — число фаз в системе.

Под *числом степеней свободы* системы понимают число внешних и внутренних параметров (температура, давление и концентрация), которое можно изменять без изменения числа фаз в системе. Если число степеней свободы равно нулю, то нельзя изменять внешние и внутренние параметры системы без того, чтобы это не вызвало изменения числа фаз. Если число степеней свободы равно единице, то возможно изменение в некоторых пределах одного из параметров и это не вызовет уменьшения или увеличения числа фаз.

Например, рассмотрим случай кристаллизации чистого вещества (элементарного полупроводника) при постоянном давлении. В этом случае правило Гиббса принимает вид $c = k - f + 1$.² Когда полупроводник находится в жидком состоянии, то есть $f = 1$, число степеней свободы равно 1 ($c = k - f + 1 = 1 - 1 + 1 = 1$). Температуру в данном случае можно изменять, не изменяя агрегатного состояния. В момент кристаллизации $f = 2$ (две фазы — твердая и жидкая), $c = k - f + 1 = 1 - 2 + 1 = 0$. Это значит, что две фазы находятся в равновесии при строго определенной

²Независимыми переменными в уравнении Гиббса являются концентрация, температура и давление. Если давление постоянно, то число переменных в уравнении уменьшится на единицу.

температуре (температуре плавления), и она не может быть изменена до тех пор, пока одна из фаз не пропадет (на графике температура–время появится площадка $T = \text{const}$, протяженность которой будет равна времени от начала и до конца кристаллизации). Источником поддержания постоянной температуры является в данном случае выделяющаяся *скрытая теплота кристаллизации*, равная разности теплосодержаний старой и новой фаз. По завершении кристаллизации в системе остается только одна твердая фаза, то есть температура снова может изменяться (уменьшаться) без изменения числа фаз.

Диаграммы состояния изображают фазовый состав системы при разных концентрациях компонентов X , температурах T и давлении P . Диаграммы состояния в общем случае являются пространственными. Размерность пространства зависит от числа независимых переменных, функцией которых является фазовый состав. Эти переменные и являются координатами, в которых строится диаграмма. Простейший тип фазовых диаграмм характеризует состояние чистого однокомпонентного материала в зависимости от давления и температуры, например, хорошо известная диаграмма состояния воды. Однако подобные однокомпонентные системы мы не будем рассматривать, а сразу перейдем к рассмотрению многокомпонентных систем, так как при получении полупроводников используются именно многокомпонентные диаграммы. Чаще всего такие диаграммы строят в координатах температура–концентрация ($T - X$). В этом случае для бинарных (двухкомпонентных) систем диаграммы изображаются на плоскости. Для тройных (трехкомпонентных) систем диаграммы строятся в трехмерном пространстве и т. д. Если кроме температуры переменным является также давление, то уже и для бинарных систем диаграммы становятся трехмерными ($P - T - X$ диаграммы). В дальнейшем мы будем рассматривать в основном только бинарные системы, построенные в координатах $T - X$. Однако в этой главе будут также рассмотрены и $P - T - X$ диаграммы некоторых полупроводниковых бинарных систем, имеющие большое практическое значение.

Обычно концентрация на диаграммах выражается в весовых или молярных долях одного из компонентов или в атомных процентах. Поэтому область изменения концентрации, отложенная на оси X , ограничена и простирается от нуля до единицы или до 100%. Для полупроводниковых систем наряду с диаграммами, построенными в линейном масштабе, иногда строят диаграммы, на которых концентрация какого-либо компонента откладывается в атомах на кубический сантиметр или в атомных процентах, но используется логарифмический масштаб. Это связано с тем, что, как правило, предельная растворимость (см. гл. 7) большин-

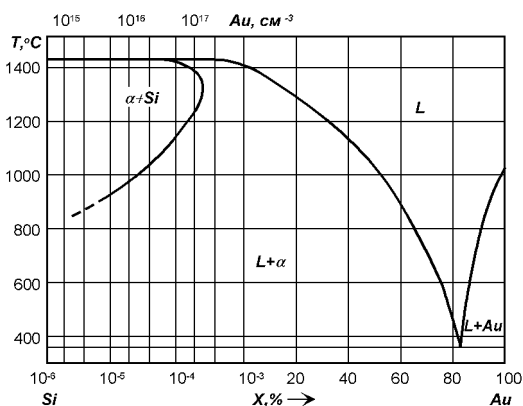


Рис. 4.3. Диаграмма состояния системы Si–Au с различными масштабами по оси концентраций (в области, примыкающей к полупроводнику, атомные проценты легирующего компонента отложены в логарифмическом масштабе, а далее концентрация в атомных процентах отложена в линейном масштабе).

ства элементов (примесей) в полупроводниках в твердом состоянии мала (менее 0.1 ат.%) и реально используемое легирование по концентрации составляет 10^{15} – 10^{19} атомов/см³, то есть 10^{-5} – 10^{-2} ат.% (см. рис. 4.3).

Фазовые диаграммы состояния дают информацию о характере фаз и фазовом составе системы при изменении концентрации одного или нескольких компонентов, температуры и давления. С помощью равновесных диаграмм состояния для данных условий можно определить: 1) число фаз в системе; 2) состав каждой фазы, ее природу (элементарное вещество, соединение, твердый раствор) и условия, при которых она образуется; 3) относительное количество каждой из фаз.

Фазовые диаграммы строятся на основе данных физико-химического анализа. В основе этого анализа лежит экспериментальное изучение зависимостей физических свойств от таких параметров, как концентрация, температура, давление. Знание этих зависимостей позволяет устанавливать природу фаз и границы их существования. Наиболее распространенными методами, используемыми для построения диаграмм состояния, являются термографические и dilatометрические методы. Их сущность заключается в том, что для сплава данного состава температуры фазовых превращений определяются по скачкообразному изменению энтальпии H (теплосодержания) или объема V системы, фиксируемому на кривых температура–время (температуру отмечают через определенные промежутки времени) или температура–объем в процессе охлаждения или нагревания сплава. Определив таким образом точки фазовых превращений для сплавов разного состава данной системы, можно построить всю диаграмму состояния. Этими методами определяют только фазовые превращения

первого рода. Эти переходы следует отличать от фазовых превращений второго рода (ферромагнитное–парамагнитное состояния, сверхпроводящее–несверхпроводящее, упорядоченное–неупорядоченное), сопровождаемых скачкообразным изменением коэффициента сжимаемости и теплоемкости. В этом случае строят диаграммы состав–свойство или для данного состава диаграммы температура–свойство и т. д.

4.1.2. $T - X$ диаграммы состояния бинарных систем

Диаграммы состояния систем с неограниченной растворимостью компонентов

Рассмотрим простейшую фазовую диаграмму с *неограниченной растворимостью* компонентов друг в друге на примере системы Ge–Si (рис. 4.1). Под термином *неограниченная растворимость* имеется в виду то, что компоненты Ge и Si системы неограниченно растворяются друг в друге в жидком и твердом состояниях и образуют *непрерывный ряд твердых растворов замещения*.³

Линию AaB на рис. 4.1, выше которой все сплавы находятся в жидком состоянии, называют *линией ликвидуса*. Линию AbB , ниже которой сплавы находятся в твердом состоянии (твердый раствор α), называют *линией солидуса*. Область, ограниченная линиями ликвидуса и солидуса, представляет собой двухфазную область, в которой сосуществуют жидкая и твердая фазы ($L + \alpha$).

Любая точка на диаграмме состояния, называемая *фигуративной*, характеризует состояние сплава. Проекция точки на ось X указывает состав сплава, а проекция на ось T — температуру, при которой он рассматривается.

Если фигуративная точка находится в однофазной области, то состав фазы совпадает с составом сплава. Если фигуративная точка t находится в двухфазной области, то состав фаз, находящихся в равновесии при данной температуре, определяется с помощью *коноды*. *Конода* — это отрезок изотермы, проходящий через данную фигуративную точку t и заключенный в пределах двухфазной области (отрезок amb на рис. 4.1). Составы фаз, находящихся в равновесии при данной температуре, определяются по абсциссам точек пересечения коноды с линиями ликвидуса

³При неограниченной растворимости любое количество атомов Ge может быть заменено атомами Si. Если увеличивать концентрацию атомов Si, то все больше и больше атомов Si будет находиться в узлах решетки вместо атомов Ge. Это будет происходить до тех пор, пока все атомы Ge не будут заменены атомами Si и, таким образом, как бы плавно совершится переход от Ge к Si.

(точка a — состав жидкой фазы X_L) и солидуса (точка b — состав твердой фазы X_S).

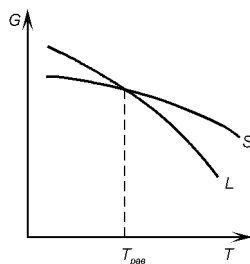
Количественное соотношение фаз, находящихся в равновесии при данной температуре, определяется с помощью правила рычага. Это правило гласит: отношение массовых количеств Q фаз, находящихся в равновесии при данной температуре, обратно пропорционально отношению отрезков коноды, заключенных между фигуративной точкой сплава и фигуративными точками, определяющими составы соответствующих фаз (см. рис. 4.1):

$$Q_L/Q_S = mb/ma. \quad (4.2)$$

Рассмотрим теперь кристаллизацию сплава Ge–Si, содержащего, например, 70% Si (X_m на рис. 4.1). Выше температуры T_1 сплав находится в жидком состоянии (L) и при охлаждении до T_1 не претерпевает фазовых превращений. Начиная с T_1 раствор оказывается пересыщенным кремнием, и из него начинает выделяться твердый раствор α , более богатый Si, чем исходная жидкость L . Поэтому в результате кристаллизации расплав обедняется Si. Первые выделяющиеся кристаллы имеют состав, отвечающий точке 3 (90% Si и 10% Ge). По мере охлаждения от T_1 до T_2 происходит дальнейшая кристаллизация сплава, в ходе которой количество жидкой фазы уменьшается, а твердой увеличивается. При T_2 исчезают последние капли расплава. В ходе кристаллизации от T_1 до T_2 составы жидкой и твердой фаз непрерывно изменяются. Если процесс кристаллизации идет с очень малой скоростью (так, что превращения успевают совершиться в соответствии с равновесной диаграммой состояния), то состав жидкой фазы изменяется по линии ликвидуса от точки 2 к точке 1, а состав твердой фазы — по линии солидуса от точки 3 к точке 4. К концу кристаллизации твердый раствор имеет состав, равный составу исходной жидкости.

Диаграммы с неограниченной растворимостью компонентов А и В друг в друге встречаются в системах, компонентами которых являются как химические элементы, так и соединения (см. рис. 4.1 и рис. 4.2). Для образования таких диаграмм оба компонента системы должны обладать одинаковым типом химической связи, одинаковым типом кристаллических решеток и атомами, размеры которых отличались бы не более чем на 10–15% (почему — см. ниже). Примерами полупроводниковых систем с неограниченной растворимостью являются системы Ge–Si, InAs–InP, GaSb–InSb, GaAs–GaP, GaAs–InAs, PbTe–SnTe, CdTe–HgTe и др. Твердые растворы этих полупроводников находят широкое практическое применение в качестве материалов для создания оптоэлектронных приборов.

Рис. 4.4. Зависимость термодинамического потенциала от температуры для жидкой и твердой фаз одинакового состава.



Растворы Ge–Si используются для создания современных СВЧ-транзисторов.

Типы фазовых диаграмм с неограниченной растворимостью компонентов

Фазовые диаграммы с неограниченной растворимостью компонентов друг в друге встречается относительно редко. Гораздо чаще приходится иметь дело с фазовыми диаграммами с ограниченной растворимостью компонентов друг в друге. Термодинамический анализ позволяет качественно объяснить переход от диаграмм состояния с неограниченной растворимостью к диаграммам состояния с ограниченной растворимостью компонентов друг в друге.

Как известно из термодинамики, термодинамически равновесному состоянию системы при данных условиях ($T = \text{const}$ и $P = \text{const}$) отвечает минимум термодинамического потенциала Гиббса

$$G = H - TS = U + PV - TS, \quad (4.3)$$

где H , T , S , U , P , V — энтальпия, температура, энтропия, внутренняя энергия, давление и объем термодинамической системы, соответственно.

Энтропия всегда положительна и увеличивается с повышением температуры, поэтому с ростом температуры потенциал Гиббса G всегда уменьшается. Кроме того, при фиксированной температуре для жидкой фазы энтропия больше, чем для твердой фазы того же состава, и поэтому кривая $G(T)$ для жидкой фазы снижается с ростом температуры резче, чем для твердой фазы (см. рис. 4.4). Таким образом, кривые $G(T)$ для твердой и жидкой фаз должны пересечься. Точка пересечения $T_{\text{равн}}$ двух кривых $G(T)$ для жидкой и твердой фаз соответствует температуре фазового превращения. Ниже и выше этой температуры стабильны разные фазы, а именно та, которая имеет меньшую величину G .

Рассмотрим концентрационную зависимость G в двухкомпонентной системе. Если компоненты А и В совсем не растворяются друг в друге и любой сплав данной системы представляет собой механическую смесь компонентов, то термодинамический потенциал любого сплава системы будет определяться аддитивно. Если же компонент В растворим в компоненте А, то вплоть до состава, отвечающего предельной растворимости (см. гл. 7), зависимость G от концентрации будет изображаться непрерывной (в силу принципа непрерывности⁴) кривой, обращенной выпуклостью вниз по крайней мере в области сильно разбавленных растворов.⁵ Второе утверждение чрезвычайно важно для получения особо чистых полупроводников. Из него следует, что при малых концентрациях компонента В величина dG/dX_B всегда отрицательна. Это означает, что растворение малого количества примеси в чистом веществе всегда сопровождается уменьшением термодинамического потенциала системы, то есть «загрязнение» чистого компонента примесью всегда энергетически выгодно. Поэтому для получения чистых полупроводников необходимо тщательно контролировать чистоту условий их выращивания.

Исходным пунктом для решения вопроса о результатах взаимодействия разных веществ при их смешении является знак изменения термодинамического потенциала системы при смешении. Рассмотрим этот вопрос на примере бинарной системы, компонентами которой являются элементарные вещества А и В. Для такой системы

$$\Delta G = \Delta H - T\Delta S, \quad (4.4)$$

где $\Delta H = H_{AB} - (H_A + H_B)$; $\Delta S = S_{AB} - (S_A + S_B)$; $\Delta G = G_{AB} - (G_A + G_B)$. Здесь G_{AB} , H_{AB} и S_{AB} — потенциал Гиббса, энтальпия и энтропия смешанной системы А+В, соответственно; G_A и G_B , H_A и H_B , S_A и S_B — потенциалы Гиббса, энтальпия и энтропия «чистых» систем А и В по отдельности, соответственно.

Величина и знак ΔG сложным образом зависят от ряда факторов, основными из которых являются два:

⁴Принцип непрерывности: при непрерывном изменении химического состава системы или других переменных факторов (T или P), не сопровождающемся изменением фазового состава, непрерывно и плавно меняются и свойства системы [17].

⁵Вблизи компонента А, то есть в области очень разбавленных растворов, $dG/dX_B = \text{tg } \alpha = \mu_B$, где μ_B — химический потенциал компонента В в растворе вещества А с концентрацией X_B , α — угол между касательной к кривой $G(X_B)$ и осью X . При $X_B \rightarrow 0$ $\mu_B = \mu_B^0 + RT \ln X_B \rightarrow -\infty$, где μ_B^0 — химический потенциал чистого компонента В, R — универсальная газовая постоянная. Таким образом, угол $\alpha \rightarrow 270^\circ$ и кривая $G(X_B)$ касается оси ординат (см. рис. 4.5,б; [17]).

1. химический, учитывающий степень химического сродства компонентов (их валентность, электроотрицательность, тип химической связи и тип решетки);

2. геометрический, учитывающий разницу в размерах атомов взаимодействующих компонентов $\Delta R = (R_A - R_B)/R_A$.

Усиление химического сродства компонентов уменьшает внутреннюю энергию системы U , и, следовательно, и H , так как $H = U + PV$. В этом случае при смешении H будет уменьшаться и поэтому $\Delta H < 0$.

Увеличение разницы в размерах взаимодействующих атомов ΔR вызывает рост U и, следовательно, рост H за счет увеличения энергии упругой деформации решетки, поэтому в этом случае $\Delta H > 0$.

Однако результат взаимодействия компонентов определяется ΔH лишь при относительно низких температурах. В соответствии с (4.4) при повышении температуры все большую роль начинает играть энтропийный член $T\Delta S$.

Согласно термодинамическим представлениям, чем сильнее разупорядоченна система (случайное распределение компонентов А и В по узлам кристаллической решетки), тем выше ее энтропия. Вследствие этого энтропия разных фазовых состояний (неупорядоченной системы $S_{\text{неуп.с}}$, упорядоченной системы $S_{\text{уп.с}}$, соединения S_c и механической смеси $S_{\text{м.с}}$) будет убывать в следующем порядке: $S_{\text{неуп.с}} > S_{\text{уп.с}} > S_c > S_{\text{м.с}}$.

Помимо разупорядоченности, связанной со случайным распределением атомов разного сорта по узлам решетки, нужно иметь в виду, что существует еще разупорядоченность, обусловленная колебаниями атомов около узлов решетки. Она тем больше, чем больше амплитуда колебаний и ниже частота колебаний. В твердых растворах к тепловым колебаниям динамического характера добавляются еще отклонения от идеальных узловых положений за счет разницы в размерах атомов А и В, то есть статические искажения. В формуле (4.3) поэтому энтропия состоит из части $S_{\text{конф}}$, связанной со случайным распределением атомов, и части $S_{\text{кол}}$, связанной с колебаниями, то есть с отклонениями от упорядоченного расположения атомов, вызванными динамическими и статическими искажениями.

Таким образом, при низких температурах преобладает роль ΔH , а при высоких — роль ΔS .

Очевидно, что неограниченная растворимость наблюдается в системах, для которых $\Delta H < 0$ как в жидком L , так и твердом S состояниях во всем интервале концентраций. Однако в зависимости от соотношения между абсолютными значениями $\Delta H(L)$ и $\Delta H(S)$ диаграммы с неограниченной растворимостью компонентов друг в друге имеют свои особен-

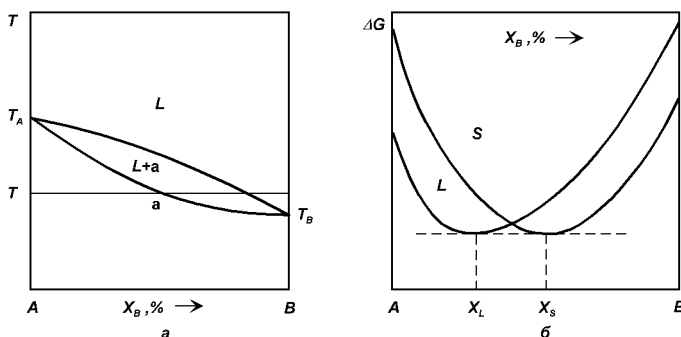


Рис. 4.5. *а* — Диаграмма состояния бинарной системы с неограниченной растворимостью компонентов друг в друге. *б* — Зависимости термодинамического потенциала твердых и жидких растворов от состава при некоторой температуре T такой, что $T_{\text{пл}}(A) > T > T_{\text{пл}}(B)$; X_L и X_S — концентрации сосуществующих при данной температуре T жидкой и твердой фаз, определенные по координатам точек касания общей касательной к кривым $\Delta G(L)$ и $\Delta G(S)$. Для бинарных систем условия равновесия двух фаз выражаются условием общности касательной к кривым $\Delta G(L)$ и $\Delta G(S)$ двух фаз, находящихся в равновесии [17].

ности. Проанализировав эти особенности, можно качественно объяснить причины перехода от диаграмм с неограниченной растворимостью к диаграммам с ограниченной растворимостью. Рассмотрим бинарную систему, в которой для определенности $T_{\text{пл}}(A) > T_{\text{пл}}(B)$, где $T_{\text{пл}}(A)$ и $T_{\text{пл}}(B)$ — температуры плавления веществ А и В.

1. Если при $T > T_{\text{пл}}(A)$ $\Delta H(L) < 0$ и $\Delta G(L)$ лежит ниже $\Delta G(S)$ во всем интервале концентраций, при $T_{\text{пл}}(B) < T < T_{\text{пл}}(A)$ кривые ΔG для жидкого и твердого растворов пересекаются только в одной точке во всем интервале концентраций (рис. 4.5,б),⁶ а при $T < T_{\text{пл}}(B)$ $\Delta H(S) < 0$ и $\Delta G(S)$ лежит ниже $\Delta G(L)$ во всем интервале концентраций, то мы будем иметь диаграмму с неограниченной растворимостью компонентов друг в друге (рис. 4.5,а).

2. Если же при температуре $T_1 > T^*$ $\Delta H(L) < 0$, $\Delta H(S) < 0$ и кривая $\Delta G(L)$ расположена ниже, чем $\Delta G(S)$ во всем интервале концентраций; а в интервале температур $T^* > T_2 > T_{\text{пл}}(A)$ зависимости ΔG для жид-

⁶Кривые зависимости термодинамического потенциала жидкой и твердой фаз от состава во всем интервале концентраций перемещаются при изменении температуры друг относительно друга. Их взаимное расположение определяется тем, какая из фаз при данной температуре обладает наименьшим термодинамическим потенциалом Гиббса, то есть является равновесной (см. выше).

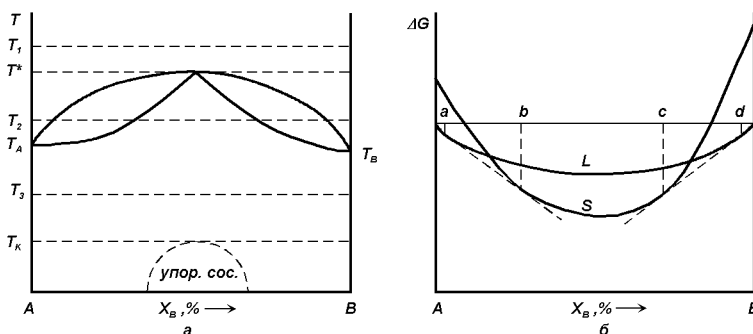


Рис. 4.6. *а* — Диаграмма состояния бинарной системы с неограниченной растворимостью компонентов друг в друге и максимумом температуры плавления. *б* — Зависимости термодинамического потенциала твердых и жидких растворов от состава при некоторой температуре T_2 такой, что $T^* > T_2 > T_{\text{пл}}(A)$.

кого и твердого раствора от состава будут пересекаться в двух точках (рис. 4.6,б) и в интервале концентраций между точками пересечения сродство компонентов А и В в твердом состоянии проявляется сильнее, чем в жидком $|\Delta H(L)| < |\Delta H(S)|$; при $T_{\text{пл}}(B) < T < T_{\text{пл}}(A)$ кривые ΔG для жидкого и твердого растворов пересекаются только в одной точке во всем интервале концентраций; а при температурах $T_3 < T_{\text{пл}}(B)$ кривая $\Delta G(S)$ расположена ниже, чем $\Delta G(L)$ во всем интервале концентраций, то на диаграмме состояния линии ликвидуса и солидуса примут вид кривых с общим максимумом (рис. 4.6,а).

С ростом абсолютной величины отрицательной по знаку $\Delta H(S)$ будет усиливаться тенденция к упорядочению твердого раствора. При высоких температурах этой тенденции будет препятствовать энтропийный вклад в $\Delta G(S)$ ($S_{\text{неуп.с}} > S_{\text{уп.с}}$). Следовательно, упорядочение будет проявляться, как правило, только при относительно низких температурах. Это означает, что должна существовать некоторая равновесная температура T_K , при которой термодинамические потенциалы упорядоченного и неупорядоченного твердого раствора равны $G_{\text{уп.с}} = G_{\text{неуп.с}}$. При температурах выше T_K твердый раствор неупорядочен, а при $T < T_K$ — упорядочен.

Упорядочение сопровождается резким изменением свойств твердого раствора, например, значительным изменением электро- и теплопроводности, пластичности и т. д. Процесс упорядочения является диффузионным процессом (превращение сопровождается перемещением атомов), поэтому медленное охлаждение способствует упорядочению. Температу-

ра T_K тем выше, чем больше $|\Delta H| = |H_{\text{уп.с}} - H_{\text{неуп.с}}|$. Выигрыш в энергии при упорядочении чаще всего связан с уменьшением энергии упругих искажений решетки твердого раствора, которое пропорционально квадрату ΔR [17], то есть T_K тем выше, чем больше ΔR . Но чем выше T_K , тем больше диффузионная подвижность и, следовательно, выше вероятность достижения равновесного упорядоченного состояния. Поэтому в тех случаях, когда ΔR велико, T_K высока и упорядочение достигается легко; в тех же случаях, когда ΔR мало, T_K низка и по кинетическим причинам упорядочение может не реализоваться или реализоваться лишь частично.

Степень упорядочения твердого раствора определяется отклонением числа ближайших соседей от предсказываемого теорией вероятности. Например, если твердый раствор В в А содержит 25 ат.% В, то первая (и любая другая) координационная сфера при отсутствии упорядочения должна содержать в среднем 75% атомов А и 25% атомов В. Если же атомы А окажутся окруженными в первой координационной сфере атомами В более чем на 25%, то твердый раствор будет упорядочиваться; если же более чем 75% соседей будут атомами А, то будет существовать тенденция к расслоению (или оно наступит) твердого раствора на две фазы. Таким образом, в твердых растворах возможно различное упорядочение в расположении атомов. В неупорядоченных твердых растворах замещения атомы разного сорта произвольно распределены в узлах кристаллической решетки, то есть отсутствует *дальний и ближний порядок*. В твердых растворах замещения с ближним порядком есть корреляция в расположении атомов разного сорта в области с конечным радиусом. В твердых растворах внедрения атомы одного компонента образуют регулярную кристаллическую решетку, атомы другого сорта беспорядочно распределены в междоузлиях этой решетки. В упорядоченном твердом растворе атомы компонентов образуют несколько вставленных друг в друга кристаллических подрешеток.

Степень упорядочения может быть оценена количественно с помощью дифракционных методов исследования.

Количественными критериями степени упорядочения служат *параметры дальнего и ближнего порядка*. Если в кристаллической решетке твердого раствора (бинарного) определенные узлы решетки заполняются преимущественно атомами определенного сорта (в этом случае распределение атомов разного сорта по узлам кристаллической решетки, существующее в одной из элементарных ячеек твердого раствора, повторяется во всех элементарных ячейках данного кристаллического блока), то возникающее распределение атомов называют *дальним порядком*. *Дальний порядок* — это корреляция в расположении атомов на макроскопических расстояниях. Под степенью (параметром) дальнего порядка η понимают величину

$$\eta = p - q, \quad (4.5)$$

где p — доля атомов сорта A , занимающих «правильные» положения в решетке, q — доля атомов того же сорта, занимающих «неправильные» положения в кристаллической решетке. Если $\eta = 1$, то твердый раствор полностью упорядочен (все «правильные» положения заняты правильно); если $\eta = 0$, то твердый раствор неупорядочен (нахождение атомов в «правильных» и «неправильных» позициях равновероятно).

Однако степень дальнего порядка не полностью определяет характер взаимного расположения атомов по узлам кристаллической решетки. Энергия взаимодействия между разными парами атомов, вообще говоря, различна, и поэтому каждый атом стремится окружить себя либо атомами другого сорта, либо одноименными атомами, причем при этом в твердом растворе может и не существовать дальнего порядка. Наличие корреляции в расположении атомов на расстояниях порядка нескольких межатомных расстояний называется *ближним порядком*. Только в простейших случаях ближний порядок распространяется на одну-две координационные сферы. Как правило, корреляция ощущается и на более далеких расстояниях. Степенью (параметром) ближнего порядка для i -й координационной сферы в неупорядоченных твердых растворах называют величину

$$\alpha_i = 1 - N_i^{AB}/N_{c_A c_B c_i} = 1 - p_i^{AB}/c_B, \quad (4.6)$$

где N_i^{AB} — число пар атомов A и B , находящихся на расстоянии r_i один от другого, $N_{c_A c_B c_i}$ — число пар атомов A и B , которые находились бы на этом расстоянии при полностью случайном распределении по узлам решетки, N — общее число атомов, c_A и c_B — концентрации атомов сорта A и B , c_i — число узлов в i -й координационной сфере, p_i^{AB} — вероятность нахождения атома B около атома A , расположенного в начале координат, на i -сфере. Из этого определения параметра ближнего порядка следует, что $\alpha_1 < 0$, если ближайшими соседями являются разноименные атомы, и $\alpha_1 > 0$, если ближайшие соседи — атомы одного сорта. Если $\alpha_1 = 1$, то сплав находится в состоянии расслоения, то есть представляет собой механическую смесь. При полном отсутствии ближнего порядка все $\alpha_i = 0$.

Параметры α_i связаны с характером распределения атомов при дальнем порядке, и эта связь позволяет ввести понятие о *сверхструктуре* (полностью упорядоченном твердом растворе). Действительно, при появлении дальнего порядка атомы разных сортов занимают определенные места в кристаллической решетке. Поэтому спектр параметров α_i , очевидно, должен в какой-то мере определять тип дальнего порядка. Эта связь проявляется в том, что зависимость параметров $\alpha(r_i)$ от радиуса координационной сферы оказывается различной для разных типов упорядочения, и это дает возможность по виду $\alpha(r_i)$ определять сверхструктуру.

Следует заметить, что для полупроводниковых систем появление сверхструктур не так характерно, как для металлических систем. Поэтому для простоты в качестве примера упорядоченных твердых растворов приведем упорядоченные твердые растворы в системе Fe–Al. В этой системе образуются две упорядоченные фазы (сверхструктуры) при 25% и 50% Al, отвечающие формулам Fe_3Al и FeAl . В неупорядоченном состоянии твердый раствор имеет о.ц.к. решетку со случайным распределением атомов по узлам кристаллической решетки. Вероятность наличия в данном узле решетки того или иного атома зависит от концентрации сплава. В упорядоченной фазе FeAl атомы Al располагаются в центрах всех малых кубов, как в решетке типа CsCl . Элементарная ячейка, как и в неупорядоченном растворе, равна малой о.ц.к. ячейке. В упорядоченной фазе Fe_3Al атомы Al занимают только центры четырех из восьми малых ячеек большого куба; центры остальных четырех малых ячеек остаются вакантными. Теперь элементарной будет уже не маленькая о.ц.к. ячейка, а большой куб, состоящий из восьми малых. Если определить α_i по формуле (4.6) для первых восьми координационных сфер, то окажется, что зависимость параметров $\alpha(r_i)$ от радиуса координационной сферы (последовательность значений $\alpha(r_i)$ для разных r_i) различна для разных типов упорядочения [36].

С увеличением абсолютной величины отрицательной по знаку $\Delta H(S)$ температура перехода T_K в упорядоченное состояние будет повышаться, но тип решетки твердого раствора, образующегося при $T < T_K$, будет такой же, что и твердого раствора, образующегося при $T > T_K$, то есть будет сохраняться, хотя периоды решетки могут изменяться (см. пример Fe–Al). Лишь в некоторых случаях может происходить незначительное искажение решетки. Наконец, при очень большой отрицательной величине $\Delta H(S)$ сильное химическое сродство может привести к образованию химического соединения со своей решеткой, отличающейся от решетки твердого раствора, уже при температуре кристаллизации. Таким образом, упорядоченные твердые растворы являются промежуточными фазами между химическими соединениями и твердыми растворами. При полном упорядочении твердые растворы напоминают химические соединения, потому что у них: а) соотношение чисел атомов компонентов фазы строго определенное: то есть соответствует стехиометрической пропорции, что может быть выражено формулой; б) расположение атомов в решетке упорядоченное. Однако упорядоченные твердые растворы могут быть причислены и к твердым растворам, так как у них сохранилась решетка растворителя.

3. Если силы химической связи разнородных атомов слабы и велика разница в размерах атомов взаимодействующих компонентов, то есть велика энергия упругой деформации при образовании твердого раствора, то в этом случае ΔH — величина положительная. Пусть $\Delta H(L) > 0$ и $\Delta H(S) > 0$, тогда, если при температуре $T_1 > T_{пл}(A)$ кривая $\Delta G(L)$ расположена ниже, чем $\Delta G(S)$ во всем интервале концентраций; при $T_{пл}(B) < T < T_{пл}(A)$ кривые ΔG для жидкого и твердого растворов пересекаются только в одной точке во всем интервале концентраций; но в интервале температур $T^* < T_2 < T_{пл}(B)$ кривые ΔG для жидкого и твердого растворов так же, как в предыдущем случае, пересекаются в двух точках (рис. 4.7,б) и в интервале концентраций между точками пересечения затруднено образование твердого раствора $\Delta H(S) > \Delta H(L)$; и при $T_3 < T^*$ кривая $\Delta G(S)$ расположена ниже, чем $\Delta G(L)$ во всем интервале концентраций, то линии ликвидуса и солидуса будут иметь общий минимум (рис. 4.7,а).

С увеличением величины положительного $\Delta H(S)$ образование связей А–А и В–В будет более выгодным по сравнению с образованием связей А–В, то есть будет проявляться тенденция к распаду твердого раствора и образованию механической смеси. При высоких температурах, близких к температуре, при которой линии ликвидуса и солидуса имеют общий минимум, этой тенденции препятствует энтропийный вклад в

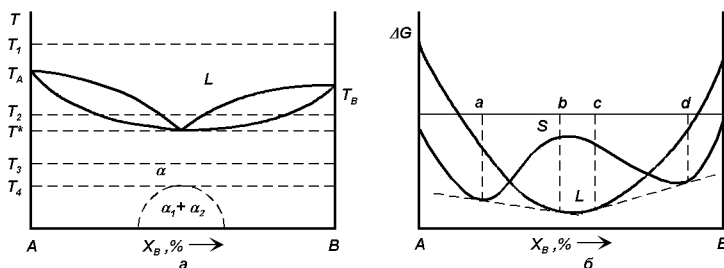


Рис. 4.7. а — Диаграмма состояния бинарной системы с неограниченной растворимостью компонентов друг в друге и минимумом температуры плавления. б — Зависимости термодинамического потенциала твердых и жидких растворов от состава при некоторой температуре T_2 такой, что $T^* < T_2 < T_{\text{пл}}(B)$.

$\Delta G(S)$, так как $S_{\text{тв.р}} > S_{\text{м.с.}}$. Сплавы поэтому будут кристаллизоваться в виде твердых растворов, но при более низких температурах твердый раствор может начать распадаться, превращаясь в механическую смесь твердых растворов меньшей концентрации ($\alpha_1 + \alpha_2$). Таким образом, при $\Delta H > 0$ при высоких температурах $G_{\text{тв.р}} < G_{\text{м.с.}}$, а при низких температурах $G_{\text{тв.р}} > G_{\text{м.с.}}$, и должна существовать некоторая температура T_4 , при которой $G_{\text{тв.р}} = G_{\text{м.с.}}$. Если $\Delta H(S)$ невелико, то эта температура низкая и распад твердого раствора не происходит. Однако с ростом $\Delta H(S)$ эта температура увеличивается, а температура общего минимума линий ликвидуса и солидуса одновременно понижается до тех пор, пока их значения не совпадут. Такая тенденция приведет к преобразованию диаграммы состояния из диаграммы состояния с неограниченной растворимостью компонентов друг в друге в диаграмму состояния с *эвтектическим превращением*, на которой уже при температуре кристаллизации жидкой фазы будет образовываться двухфазная смесь.

Диаграммы состояния систем с ограниченной растворимостью компонентов

Диаграммы состояния с эвтектическим превращением. Рассмотрим более подробно диаграмму с эвтектическим превращением (рис. 4.8). Она состоит из шести фазовых областей: области жидкой фазы L , двух двухфазных областей (жидкая фаза L + твердый раствор α и жидкая фаза L + твердый раствор β), двух однофазных областей твердых растворов α и β и двухфазной области сосуществования твердых растворов $\alpha + \beta$.

Линиями $A'e$ и $B'e$ на диаграмме обозначены линии ликвидуса. Ли-

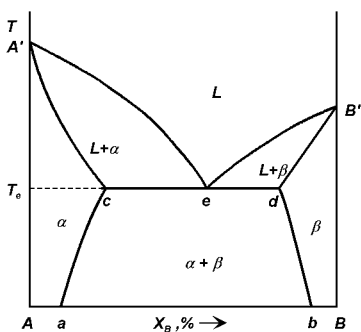


Рис. 4.8. Диаграмма состояния бинарной системы, образующей твердые растворы с эвтектикой.

нии $A'c$ и $B'd$ — линии солидуса при температурах, превышающих эвтектическую температуру T_e . Горизонтальная эвтектическая линия ced является линией солидуса, но не отвечает реально существующим фазам за исключением трех точек: c (эвтектическое превращение твердого раствора α), d (эвтектическое превращение твердого раствора β) и e (эвтектическое превращение жидкой фазы). Особенность этой линии состоит в том, что при соответствующей ей температуре T_e в сплавах состава от c до d в равновесии находятся три фазы: $L + \alpha + \beta$. Линии ac и bd — это линии твердых растворов предельной концентрации (ac — B в A , а bd — A в B), существующих при температурах ниже эвтектической.

При охлаждении сплавов составов от 0 до c мы сначала проходим двухфазную область $L + \alpha$, а затем получаем однофазный твердый раствор α . При дальнейшем охлаждении мы можем либо сохранить однофазный твердый раствор α , либо, в случае пересечения линии ac , получить двухфазную смесь $\alpha + \beta$, так как в точке пересечения сплав находится в состоянии предельно насыщенного раствора, а ниже начинается его распад с выделением фазы β . Во втором случае при дальнейшем охлаждении сплава в пределах двухфазной области одновременно изменяются как состав обеих фаз, так и их количественное соотношение. Состав фаз характеризуют точки пересечения горизонтальных линий с линиями ac и bd , ограничивающими двухфазную область; количественное соотношение фаз определяется правилом рычага. Кристаллизация сплавов составов от d до B происходит аналогично.

При охлаждении сплавов составов от c до d при температуре T_e (окончание кристаллизации) происходит распад жидкости состава e на два твердых раствора α и β с составами c и d по эвтектической реакции $L \rightarrow \alpha + \beta$. При дальнейшем охлаждении сплавов изменяется состав

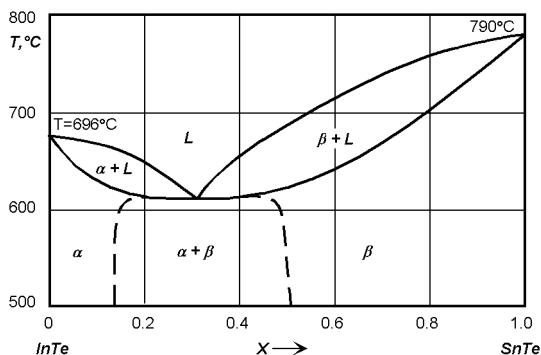


Рис. 4.9. Диаграмма состояния системы InTe–SnTe.

каждого твердого раствора α и β по кривым ac и bd соответственно, а количественное соотношение фаз определяется правилом рычага.

Основными особенностями диаграмм эвтектического типа являются: 1) уменьшение температуры плавления твердого раствора как со стороны компонента А, так и со стороны компонента В при их растворении друг в друге; 2) наличие эвтектического превращения, заключающегося в распаде жидкости при охлаждении на две твердые фазы.

Диаграммы эвтектического типа наблюдаются, например, в системах InTe–SnTe (рис. 4.9) и Si–Au (рис. 4.3).

Диаграммы состояния с перитектическим превращением. Если температура плавления твердых растворов уменьшается со стороны компонента А (или В), а со стороны компонента В (или А) повышается при их ограниченной растворимости друг в друге, то возможно *перитектическое превращение*, которое представляет собой процесс образования фазы за счет взаимодействия двух других, полностью или частично исчезающих при этом (жидкость реагирует с ранее выпавшими кристаллами и образует новый вид кристаллов): $L + \alpha \rightarrow \beta$. Реакция подобного типа называется *перитектической*.

Рассмотрим диаграмму состояния с перитектическим превращением (рис. 4.10). Линия $A'bB'$ — линия ликвидуса; линии $A'a$ и nB' — линии солидуса; линии ac и nd — линии твердых растворов предельной концентрации, существующих при температурах ниже перитектической температуры T_p . При температуре T_p в равновесии находятся три фазы: жидкая и две твердые фазы — $L + \alpha + \beta$. Горизонтальная линия ab на-

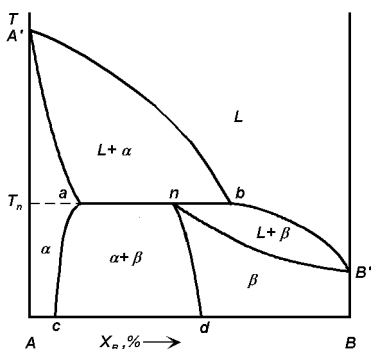


Рис. 4.10. Диаграмма состояния бинарной системы, образующей твердые растворы с перитектикой.

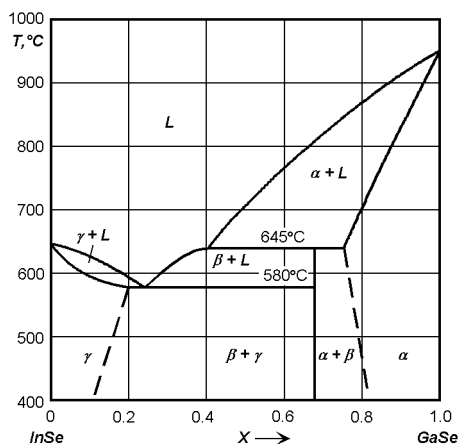
зывается перитектической горизонталью и ее точки не отвечают реально существующим фазам (как и на эвтектической горизонтали), кроме трех точек a , n и b .

При охлаждении сплавов составов от 0 до a кристаллизация происходит аналогично кристаллизации сплавов составов от 0 до c на диаграмме состояния с эвтектическим превращением.

При охлаждении сплавов, составы которых лежат в интервале от a до n , количество жидкой фазы при температуре T_p недостаточно для того, чтобы произошло взаимодействие со всей твердой фазой α . В итоге после перитектического превращения, кроме новой фазы β , сплавы будут содержать и избыточное количество кристаллов фазы α . При T_p состав жидкой фазы будет отвечать точке b (точка перитектического превращения жидкости), а состав кристаллов — точке a (точка перитектического превращения твердого раствора α). Эти обе фазы реагируют и дают третью фазу β , состав которой определяется точкой n (точка перитектического превращения твердого раствора β). Количественное соотношение фаз, участвующих в перитектической реакции, определяется правилом рычага. При дальнейшем охлаждении сплавы представляют собой механическую смесь твердых растворов $\alpha + \beta$, составы которых изменяются по кривым ac и nd соответственно, а количественное соотношение фаз определяется правилом рычага.

Аналогично для сплавов, составы которых лежат в интервале от n до b , количество фазы α при T_p мало по сравнению с количеством жидкой фазы. В результате перитектического превращения мы получим двухфазную систему $L + \beta$. При дальнейшем понижении температуры состав жидкой фазы изменяется по линии bB' , а состав образующихся кристаллов — по nB' . Количественное соотношение фаз определяется правилом

Рис. 4.11. Диаграмма состояния системы InSe–GaSe.



рычага. Важной особенностью системы с перитектикой является то, что при температуре T_n кристаллизация ряда составов сплава завершается не полностью. Для сплавов состава от n до b кристаллизация завершается на линии nB' при температурах ниже T_n и при этом образуется твердый раствор β .

Подобного типа диаграммы состояния редко встречаются в полупроводниковых системах, тем не менее, примером может служить система InSe–GaSe (рис. 4.11). По перитектической реакции в этой системе образуется β фаза (InGa_2Se_3).

Диаграммы состояния с химическими соединениями. В соответствии со сказанным выше при сильном химическом сродстве компонентов системы их взаимодействие приводит к образованию химического соединения. Диаграммы состояния с химическими соединениями различают по двум признакам: 1) по соотношению составов соединения и жидкой фазы, образующейся при плавлении этого соединения; 2) по способности соединения растворять в себе другие компоненты, в том числе и составляющие данное соединение, то есть по ширине области гомогенности (по величине области концентраций, в которой может существовать соединение, сохраняя свою однофазность).

1. *Конгруэнтно и инконгруэнтно плавящиеся соединения.* Соединения, для которых составы соединения и жидкой фазы, образующейся при плавлении этого соединения, совпадают, называются *конгруэнтно* плавящимися. Температура плавления таких соединений лежит на ли-

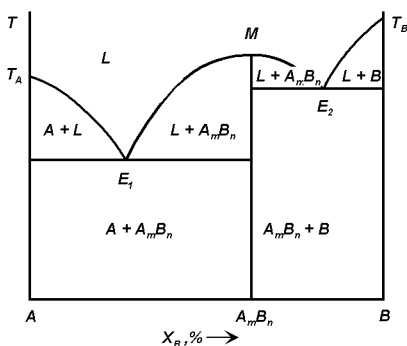
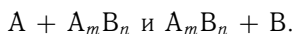


Рис. 4.12. Диаграмма состояния системы с одним конгруэнтно плавящимся соединением.

нии ликвидуса, а сама линия имеет вид кривой с максимумом (точка М), который в первом приближении соответствует составу соединения (рис. 4.12). В простейшем случае двухкомпонентной системы (компоненты системы А и В), в которой образуется только одно химическое соединение A_mB_n и твердые растворы отсутствуют, диаграмму состояния можно рассматривать как двойную систему: химическое соединение рассматривается как отдельный компонент, объединяющий две самостоятельные фазовые диаграммы с эвтектиками (на диаграмме состояния имеются две эвтектические точки E_1 и E_2):



Например, система InAs может быть разделена на системы In–InAs и InAs–As (рис. 4.13). Кристаллизация сплавов на этой диаграмме происходит совершенно аналогично кристаллизации сплавов, образующих механическую смесь кристаллов «чистых» компонентов. Отличие состоит только в том, что, кроме выделения кристаллов «чистых» компонентов А и В, происходит еще образование кристаллов химического соединения A_mB_n .

Если же соединение способно давать ограниченные твердые растворы с компонентами фазовой диаграммы (в случае полупроводниковых соединений растворимость в них компонентов соединения, как правило, очень невелика), то диаграмма состояния системы будет иметь вид, представленный на рис. 4.14. На этом рисунке изображены α и β твердые растворы на основе «чистых» компонентов А и В и γ твердый раствор на основе химического соединения A_mB_n .

В другом случае, когда составы жидкой и твердой фаз при плавлении не совпадают, соединения называются *инконгруэнтно* плавящими-

Рис. 4.13. Диаграмма состояния системы In–As с конгруэнтно плавящимся соединением InAs.

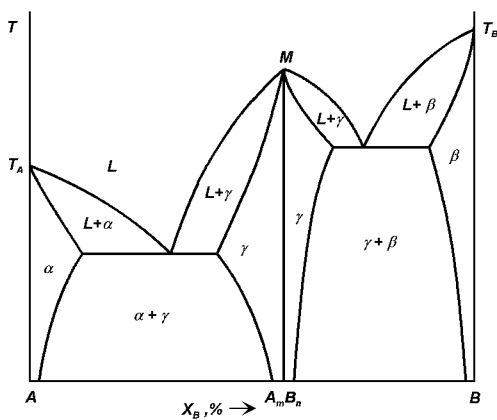
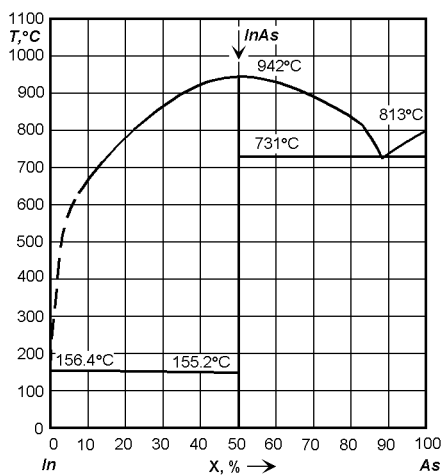


Рис. 4.14. Диаграмма состояния системы с одним конгруэнтно плавящимся соединением и ограниченными твердыми растворами с эвтектиками.

ся. Инконгруэнтно плавящиеся соединения характеризуются тем, что при плавлении твердая фаза превращается в две фазы — жидкую и твердую, составы которых лежат по обе стороны состава исходного соединения (соединение A_mB_n (точка J) и соединение $GaTe_3$ на рис. 4.15, *а, б*). Плавление инконгруэнтно плавящихся соединений идет как превращение перитектического типа, но плавится не новый твердый раствор, как при перитектической реакции, а химическое соединение. Температура плавления соединения при этом лежит ниже линии ликвидуса, в результате чего последняя не имеет максимума, связанного с соединением.

2. Ширина области гомогенности. Существуют соединения с широкой областью гомогенности (единицы–десятки атомных процентов), например, соединение Sb_2Te_3 в системе $Sb-Te$ (рис. 4.16). Существуют соединения с узкой областью гомогенности (сотые и тысячные доли процентов), например, соединение $InAs$ в системе $In-As$ (рис. 4.13). Полупроводниковые соединения группы $A^{IV}B^{VI}$ занимают промежуточное положение. Для них характерна ширина области гомогенности порядка 0.1 ат.%.

4.1.3. Кристаллизация в неравновесных условиях

Все, что до сих пор говорилось о фазовых превращениях, относилось к равновесным при данной температуре состояниям, то есть предполагалось, что скорость изменения температуры так мала, что при каждой температуре в системе успевает установиться равновесие. Однако установление равновесия в системе, осуществляемое путем диффузии, требует времени, продолжительность которого зависит от природы диффундирующих элементов, среды, температуры, размера и степени совершенства кристаллов и других факторов. В реальных условиях время, необходимое для достижения равновесия, может изменяться от долей секунд до многих сотен часов. Кроме того, как отмечалось в начале главы, фазовое превращение при температуре фазового равновесия невозможно, так как в этом случае нет энергетического выигрыша, стимулирующего это превращение. Поэтому равновесную диаграмму состояния следует рассматривать как предельный случай, при котором скорость изменения температуры настолько мала, что фазовое превращение совершается с бесконечно малой скоростью. Тем не менее, этот предельный случай крайне важен для изучения реальных условий кристаллизации.

Если скорость изменения температуры невелика, то есть если превращение происходит в условиях малых переохлаждений (в условиях близких к равновесным), то в этих случаях можно пренебречь небольшой

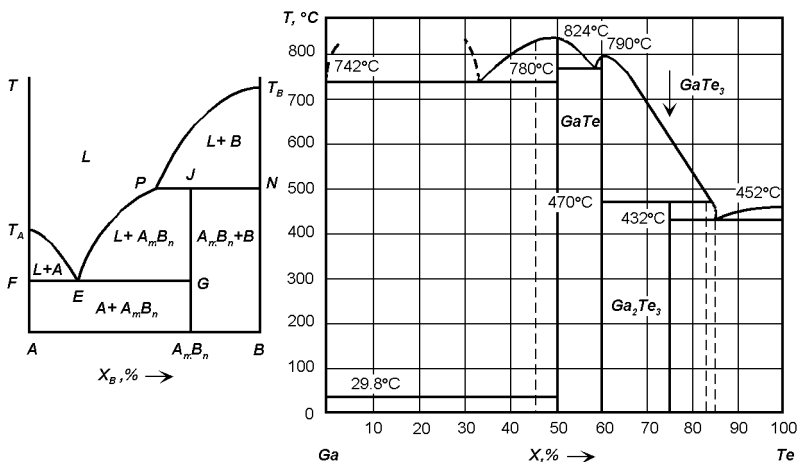


Рис. 4.15. *а* — Диаграмма состояния системы с инконгруэнтно плавящимся соединением. *б* — Диаграмма состояния системы Ga-Te с инконгруэнтно плавящимся соединением GaTe₃.

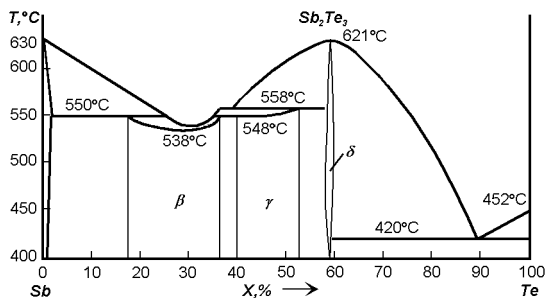


Рис. 4.16. Диаграмма состояния системы Sb-Te с широкой областью гомогенности у соединения Sb₂Te₃.

степенью переохлаждения и рассматривать превращение как совершающиеся при температурах и в последовательности, указанной на равновесной диаграмме состояния.

В условиях же быстрых изменений температур изменяется не только температура превращения, но и условия превращения, так как не успевают произойти диффузионные процессы, необходимые для осуществления превращений по типу равновесных. Для этих случаев диаграмма фазового равновесия уже не действительна.

Наиболее распространенным следствием отклонения от условий превращения, близких к равновесным, является *внутрикристаллическая ликвация*, которая заключается в том, что в системе образуется пространственно-неоднородное состояние: химический состав части кристалла, затвердевшей в начале кристаллизации, отличается от состава частей кристалла, закристаллизовавшихся в последнюю очередь. Ликвация является следствием того, что твердая фаза отличается по составу от жидкой для любой фигуративной точки, лежащей между линиями ликвидуса и солидуса, и следствием замедленности процессов диффузии в твердой фазе.

Рассмотрим подробнее это явление, а именно рассмотрим непрерывную кристаллизацию расплава состава C_0 при охлаждении в температурном интервале кристаллизации (рис. 4.17). В соответствии с равновесной диаграммой состояния при T_0 в нем должны появиться кристаллы состава a , при температуре T_1 — кристаллы состава b , при T_2 — состава c и т. д. Таким образом, с понижением температуры образуется твердая фаза, все более обогащенная элементом В. В условиях равновесной кристаллизации состав твердой фазы, отвечающий каждому значению температуры, должен быть однородным по объему. Для этого необходимо:

- а) чтобы успевало пройти диффузионное выравнивание состава в жидкой фазе;
- б) чтобы во всем объеме твердой фазы также успевало пройти диффузионное выравнивание состава.⁷

В реальных условиях диффузионное выравнивание состава в твердой фазе никогда полностью не происходит. Если принять, что в жидкой фазе и на границе раздела фаз выравнивание состава успевает пройти,⁸ а изменениями в составе твердой фазы вследствие диффузии пренебречь, то

⁷Помимо рассмотрения процессов в объемах двух фаз необходимо учитывать и процессы, происходящие на границе. Однако, как показывает опыт, на границе раздела фаз скорость установления равновесия не меньше таковой в жидкости и намного превышает соответствующую скорость для твердой фазы.

⁸Коэффициенты диффузии в жидкой фазе обычно на несколько порядков выше, чем в твердой фазе.

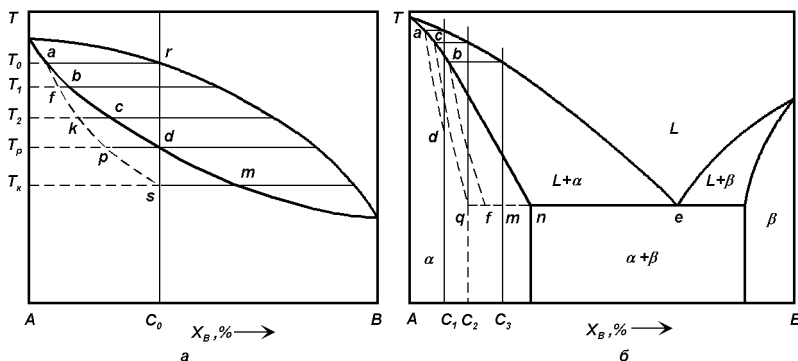


Рис. 4.17. Схема ликвации в сплавах систем с неограниченной (а) и ограниченной (б) растворимостью компонентов в твердом состоянии (ad , cq и bf — кривые среднего состава твердого раствора при неравновесной кристаллизации сплавов C_1 , C_2 и C_3 , соответственно; практически кристаллизация сплава C_2 завершается при температуре T_e и при этой температуре средний состав кристаллов будет C_2 как у расплава до начала кристаллизации; сплавы, составы которых лежат слева от C_2 , будут кристаллизоваться в однофазном состоянии, а сплавы, составы которых лежат справа, — в двухфазном, хотя сплавы из интервала C_2-n по равновесной диаграмме должны быть однофазными. Для составов, лежащих правее от C_2 , например, для состава C_3 при эвтектической температуре жидкость примет состав точки e , кристаллы — состав точки f и, следовательно, останется какое-то количество жидкости fm , которое затвердеет и даст эвтектику. Таким образом, при кристаллизации вместо однородного твердого раствора согласно равновесной фазовой диаграмме получается структура неоднородного твердого раствора + эвтектика).

неоднородный по составу кристалл будет характеризоваться средней концентрацией, которая всегда меньше равновесной концентрации, то есть лежит левее равновесной линии солидуса ad (рис. 4.17,а). Так, при температуре T_1 состав твердой фазы вместо равновесного b будет средним между b и a и соответствовать точке f , при T_2 — точке k и т.д. При равновесной температуре T_p еще будет оставаться какое-то количество жидкости (характеризуется отрезком pd), которое будет охлаждаться до температуры T_k пока линия as не пересечет линию rs . Таким образом, завершение кристаллизации произойдет не при равновесной температуре солидуса T_p , а при температуре T_k , когда средний состав твердой фазы, лежащий на линии as , сравняется с C_0 . Линию as называют неравновесным солидусом системы.

Так как скорость диффузии в твердой фазе мала, то по мере роста кристалла вновь нарастающие слои будут обогащаться компонентом В. Такое же обогащение было и в равновесном случае, однако там обогащались не только приграничные слои, но и весь объем твердой фазы. Это приводило к тому, что в конце кристаллизации состав кристалла становился равным составу исходного расплава. В результате неравновесной кристаллизации химический состав по длине кристалла оказывается неоднородным. Получающийся кристалл представляет собой ряд твердых растворов с непрерывно меняющимися концентрациями. Итак, мы пришли к выводу, сделанному вначале: неоднородность химического состава по длине кристалла связана с разницей в составах жидкой и твердой фаз, находящихся между собой в равновесии при данной температуре, и замедленностью процессов выравнивания состава (диффузии) в твердой фазе.

Следует иметь в виду, что при реальных скоростях, используемых при выращивании кристаллов, выравнивание состава жидкой фазы также, как правило, оказывается неполным (см. гл. 5).

Отклонения от равновесного состояния, еще более резкие, чем в системах с неограниченной растворимостью, наблюдаются в системах с эвтектическими и перитектическими превращениями. В таких системах однофазные согласно фазовым диаграммам сплавы могут оказаться двухфазными в условиях неравновесной кристаллизации (см. рис. 4.17,б).

Другим важным примером отклонения от равновесных условий кристаллизации является образование в процессе фазовых превращений *метастабильных фаз*. Так называют неравновесные фазы, представляющие собой переходную стадию между исходной фазой и равновесной. Свободная энергия у метастабильной фазы меньше, чем у исходной, но больше, чем у равновесной, однако превращение ее в равновесную тре-

бует определенной энергии активации, при отсутствии которой метастабильная фаза может долго существовать. С метастабильными фазами мы сталкиваемся при резком охлаждении твердых растворов (закалка), которые должны в равновесных условиях охлаждения испытывать эвтектические превращения; при распаде пересыщенных твердых растворов, если он осуществляется при низких температурах, когда диффузия и соответственно образование равновесных фаз затруднено. Для практических целей важно, что многие из метастабильных состояний обладают полезными, а иногда и уникальными свойствами. Поэтому оказывается необходимым установить природу этих неравновесных состояний, разработать принципиальные режимы термической и иной обработки, позволяющей их получить. Исходным моментом при решении этих задач является знание равновесных диаграмм состояния.

Знать диаграмму состояния необходимо также для выбора составов жидкой фазы и температурных условий кристаллизации. По фазовым диаграммам определяют температурную область плавления вещества, получают сведения о возможной степени ликвации компонентов, о режимах термической обработки, обеспечивающих нужный фазовый состав вещества и т. д., то есть роль диаграмм состояния для процессов кристаллизации огромна.

4.1.4. $P - T - X$ диаграммы состояния

Рассмотрим $P - T - X$ диаграммы для бинарных систем. Интенсивные работы по изучению $P - T - X$ диаграмм состояния показали, что использование высоких давлений (десятки и сотни тысяч атмосфер) в ряде случаев приводит к изменению типа диаграммы состояния, к резкому изменению температур фазовых и полиморфных превращений, к появлению новых фаз, отсутствующих в данной системе при атмосферном давлении. Так, например, диаграмма с неограниченной растворимостью в твердом состоянии при высоких температурах и распадом твердого раствора α на два твердых раствора $\alpha_1 + \alpha_2$ при низких температурах может с увеличением давления постепенно переходить в диаграмму с эвтектикой (см. рис. 4.18,а). На рис. 4.18,б показана диаграмма состояния системы Ga-P, в которой образуется полупроводниковое соединение GaP. В зависимости от давления это соединение может плавиться конгруэнтно или инконгруэнтно. Соответственно изменяется и вид двойной диаграммы $T - X$ на различных изобарических сечениях тройной $P - T - X$ диаграммы.

На практике объемные $P - T - X$ диаграммы строятся очень редко. Обычно фазовые превращения в трехмерных $P - T - X$ диаграммах ана-

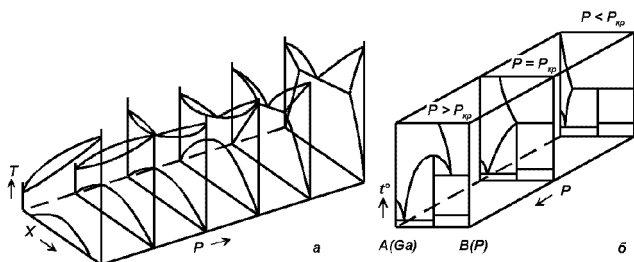


Рис. 4.18. *а* — P — T — X диаграмма; *б* — P — T — X диаграмма состояния системы Ga–P с конгруэнтно и инконгруэнтно плавящимся соединением GaP в зависимости от давления.

лизируют с помощью их проекций на плоскости P — T , T — X и P — X , а также различных сечений при постоянных значениях температуры или давления (см. рис. 4.18, *а*).

Заметим, что при анализе фазовых превращений в системе следует различать P — T — X фазовые диаграммы, в которых давление диссоциации $P_{\text{дис}}$ ⁹ мало и P на фазовой диаграмме — это внешнее давление и в которых давление диссоциации велико и P — это $P_{\text{дис}}$. В системах, компоненты которых обладают низким давлением диссоциации и в которых максимальная температура плавления смеси ниже самой низкой температуры кипения (в системе нет легколетучих компонентов), ролью газовой фазы при фазовых превращениях можно пренебречь. Если же давление диссоциации какого-либо из компонентов велико (система содержит легколетучие компоненты), то состав газовой фазы необходимо учитывать при температурах как выше, так и ниже ликвидуса.

Рассмотрим подробнее фазовые диаграммы $P_{\text{дис}}$ — T — X с высоким давлением диссоциации (фазовые диаграммы с легколетучими компонентами). Следует отметить, что внимание к ним повысилось в связи с возросшей ролью в полупроводниковой электронике соединений, содержащих летучие компоненты. Например, к ним относятся соединения $A^{\text{III}}B^{\text{V}}$, содержащие легколетучие компоненты фосфор и мышьяк, соединения $A^{\text{II}}B^{\text{VI}}$, содержащие ртуть, $A^{\text{IV}}B^{\text{VI}}$, содержащие серу, и т. д.

Все полупроводниковые соединения обладают более или менее протяженной областью гомогенности, то есть способны растворять в себе

⁹ $P_{\text{дис}}$ — равновесное для данных условий давление диссоциации всех фаз, находящихся в равновесии. При наличии в системе одного легколетучего компонента $P_{\text{дис}}$ — это равновесное давление диссоциации легколетучего компонента системы.

какой-либо из компонентов сверх стехиометрического состава или третий компонент.

Любые отклонения от стехиометрического состава сказываются на электрических свойствах (см. гл. 3). Поэтому для воспроизводимого получения кристаллов, содержащих летучий компонент, с заданными свойствами необходимо и воспроизводимое получение соединений заданного состава.

Однако летучесть одного из компонентов соединения приводит к отклонению от стехиометрического состава из-за образования вакансий — анионных или катионных — в зависимости от того, давление диссоциации какого компонента будет выше, и, соответственно, избытка другого компонента. Как уже обсуждалось в гл. 3, вакансии в ряде соединений могут создавать акцепторные или донорные уровни, тем самым влияя на физические свойства.

Энергия образования вакансий в позициях А и В практически никогда не бывает одинаковой, поэтому концентрация анионных и катионных вакансий также различна, а область гомогенности соединения оказывается несимметричной относительно стехиометрического состава. Соответственно практически для всех соединений максимум температуры плавления не соответствует сплаву стехиометрического состава.¹⁰

Предотвратить изменение состава соединения за счет летучести можно, если выращивать его из расплава или раствора при внешнем давлении летучего компонента, равном давлению диссоциации при температуре выращивания. Это условие и выбирают с помощью $P_{\text{дис}} - T - X$ диаграмм.

Давление диссоциации легколетучего компонента в сплавах сильно зависит от его состава, как правило, понижаясь с уменьшением концентрации этого компонента, как, например, для системы In–As (давление диссоциации мышьяка понижается почти на четыре порядка с уменьшением концентрации мышьяка в интервале от 100 до 20% [17]). В результате давление диссоциации летучего компонента в соединении оказывается намного меньше давления диссоциации над чистым компонентом при той же температуре.

Это обстоятельство используется в двухтемпературной схеме получения этого соединения. В одной печи создаются две температурные зоны.

¹⁰ Тем не менее, для соединений, в частности $A^{III}B^V$, с узкой областью гомогенности и большинства соединений, в частности $A^{IV}B^{VI}$, со средней шириной области гомогенности применяется понятие конгруэнтно плавящихся соединений, так как отклонения реальной температуры плавления соединения от температуры плавления соединения стехиометрического состава незначительны.

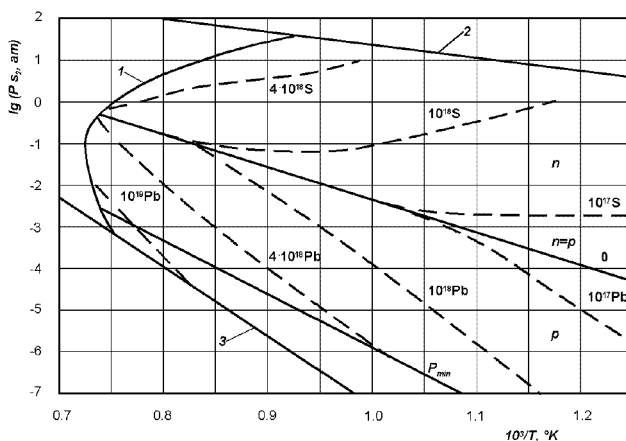


Рис. 4.19. $P_{\text{дис}} - T$ сечение $P_{\text{дис}} - T - X$ диаграммы состояния системы Pb-S. 1 — трехфазная линия; 2 — P_{S_2} чистой серы над PbS+S₂; 3 — P_{S_2} над PbS+Pb.

Одна имеет температуру T_1 , равную температуре кристаллизации соединения. Здесь помещают контейнер с расплавом. Во второй зоне помещают чистый летучий компонент соединения — As. Температура T_2 во второй зоне поддерживается равной температуре, при которой давление диссоциации летучего компонента в чистом виде равно давлению диссоциации этого компонента в соединении при температуре T_1 . В результате в первой зоне давление паров летучего компонента над соединением равно его парциальному давлению диссоциации в соединении, что предотвращает улетучивание этого компонента из расплава и обеспечивает кристаллизацию соединения заданного состава.

На рис. 4.19 приведена $P - T$ проекция фазовой диаграммы Pb-S. Сплошной линией показана линия трехфазного равновесия твердой, жидкой и газообразной фаз, ограничивающая область устойчивости твердого соединения; пунктиром — изоконцентрационные линии в пределах области гомогенности. Изоконцентрационные линии показывают составы с равным отклонением от стехиометрии (одинаковые составы) в сторону избытка свинца (проводимость n -типа) или в сторону избытка серы (проводимость p -типа), равновесные при данных значениях температуры и давления паров серы. Линия $n = p$ соответствует значениям температуры и давления P_{S_2} , при которых твердая фаза имеет строго стехиометрический состав. Она пересекает трехфазную линию при температуре, которая является температурой плавления соединения стехиометрического

состава. или в сторону избытка серы (проводимость p -типа) .

Как видно из рис. 4.19, температура плавления соединения стехиометрического состава ниже максимальной температуры плавления, которую имеет сплав с избытком свинца по сравнению с формульным составом. Видна резкая зависимость состава кристалла от парциального давления паров летучего компонента. В области высоких температур все кривые, соответствующие разным составам, приближаются к линии $n = p$. С понижением температуры разница между равновесными давлениями, соответствующими разным составам, увеличивается. Этим объясняется трудность получения сплава заданного состава непосредственно при кристаллизации, проходящей при высоких температурах. Поскольку кривые парциального давления для разных составов близки, небольшие случайные отклонения давления паров летучего компонента могут привести к ощутимому изменению состава твердой фазы.

Если же кристалл после выращивания подвергнуть длительному отжигу при более низких температурах и таком давлении, что изоконцентрационные линии для разных составов резко расходятся, то состав кристалла можно довести до заданного. Этим часто пользуются на практике.

4.2. Элементы общей теории образования фаз

В процессе получения любого материала в нем происходят переходы из одного фазового состояния в другое. Основными промышленными методами получения монокристаллов является кристаллизация из жидкой и конденсация из газовой фаз.

Кристаллизацией называется переход вещества из жидкой в твердую фазу. *Конденсацией* называется переход вещества из газообразной в жидкую или твердую фазы.

Фазовые превращения никогда не проходят одновременно во всем объеме материала. В начале процесса в разных местах исходной фазы образуются центры новой фазы — стабильные частицы новой фазы, которые затем разрастаются благодаря процессам переноса. Процесс роста кристаллов принято представлять состоящим из трех основных стадий: 1) возникновения неравновесного состояния — *переохлаждения* для исходной жидкой фазы или *пересыщения* для газовой фазы; 2) образования центров новой фазы; 3) увеличения их линейных размеров — собственно рост кристалла. При данной температуре жидкую фазу называют *переохлажденной*, если ее температура ниже температуры равновесного образования твердой фазы. При данной температуре газовую фазу назы-

вают *пересыщенной*, если ее давление превышает равновесное давление насыщенных паров жидкой или твердой фазы.

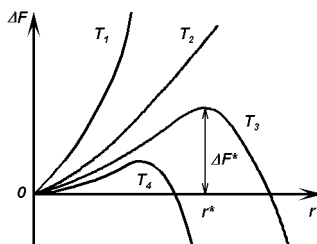
Создание неравновесного состояния — необходимое условие для самопроизвольного возникновения центров новой фазы. В равновесных условиях возможность их появления исключена. В различных точках газообразной или жидкой атомной или молекулярной системы, находящейся в термодинамическом равновесии, непрерывно возникают флуктуации, то есть отклонения величин некоторых параметров системы от их наиболее вероятных (средних) значений. Флуктуации плотности и концентрации в исходной фазе могут приводить к изменению фазового состояния, то есть к образованию зародышей новой фазы. Такие флуктуации сопровождаются изменением свободной энергии системы, однако энергия образующихся частиц (зародышей) новой фазы в системе, находящейся в термодинамически равновесном состоянии, значительно превышает энергию таких же частиц исходной фазы и флуктуационно возникшие частицы новой фазы быстро распадаются. Образование центров новой фазы в равновесных условиях оказывается энергетически не выгодным. Ниже будет показано, что создание неравновесного состояния для кристаллизации (конденсации) исходной фазы необходимо для сообщения системе дополнительной энергии, требуемой для того, чтобы процесс образования центров новой фазы сделать энергетически выгодным.

Движущей силой любого фазового превращения является стремление системы к уменьшению свободной энергии F . Для протекания фазового превращения необходимо, чтобы свободная энергия единицы объема исходной фазы была больше свободной энергии такого же объема новой фазы. Кроме того, следует учитывать, что при образовании частиц новой фазы обязательно возникают поверхности раздела между новой и исходной фазами. Эти поверхности обладают поверхностной энергией, что затрудняет превращение. Поэтому в общем виде суммарное изменение свободной энергии системы при фазовом превращении будет:

$$\Delta F = \Delta F_v + \Delta F_s, \quad (4.7)$$

где $\Delta F_v = F_v(S) - F_v(L)$ и $\Delta F_s = F_s(S)$ — изменения свободной энергии, связанные с объемной и поверхностной составляющими. Проанализируем более подробно необходимые для протекания фазовых превращений условия.

Рис. 4.20. Зависимость свободной энергии системы от радиуса зародыша новой фазы при различных температурах $T_1 > T_2 > T_3 > T_4$.



4.2.1. Образование центров новой фазы

Механизм образования центров новой фазы может быть гомогенным или гетерогенным.

Гомогенное образование центров новой фазы

Гомогенным называется образование зародыша новой фазы в объеме исходной фазы, сопровождающееся образованием всей поверхности, ограничивающей зародыш. Гиббс впервые показал, что зародыш новой фазы становится устойчивым лишь при условии, что его размер превосходит определенное критическое значение. Области новой фазы, размер которых меньше критического, называются *зародышами*, а размер которых больше критического, — *центрами новой фазы* (в случае кристаллизации центры новой фазы часто называют *центрами кристаллизации*). Френкель подробнее рассмотрел условия зарождения частиц новой фазы и определил размер критического зародыша. Найдём его для случая кристаллизации из жидкой фазы.

Пусть в жидкой фазе в результате флуктуаций в системе появился зародыш твердой фазы с объемом V и с поверхностью S , характеризуемой коэффициентом поверхностного натяжения σ . Тогда

$$\Delta F_v = \frac{V\rho}{M}(F'(S) - F'(L)) \text{ и } \Delta F_s = S\sigma, \quad (4.8)$$

где ρ и M — плотность и молекулярный вес зародыша, $F'(S)$ и $F'(L)$ — молярная свободная энергия твердой и жидкой фаз.

Зародыш твердой фазы будет устойчивым, если величина $\Delta F = \Delta F_v + \Delta F_s$ будет отрицательной.

Составляющая ΔF_s всегда положительна, так как при всех условиях создание новой поверхности раздела требует затраты энергии. Величина и знак ΔF_v зависят от соотношения между фактической температурой T

и $T_{\text{пл}}$. При $T_1 > T_{\text{пл}}$ (см. рис. 4.20) $F'(S) > F'(L)$ и, следовательно, $\Delta F > 0$ при любых размерах зародыша r . С приближением температуры T к $T_{\text{пл}}$ кривая $\Delta F(r)$ будет становиться все более полой, но ΔF будет монотонно расти. При температуре $T_3 < T_{\text{пл}}$ $\Delta F_v < 0$, так как $F'(S) < F'(L)$. Однако при малых r в суммарном изменении свободной энергии ΔF преобладает вклад положительной составляющей ΔF_s . По мере роста размера зародыша r абсолютная величина ΔF_v возрастает быстрее, чем ΔF_s , и при некотором критическом значении r^* кривая $\Delta F(r)$ достигает максимума. *Критическим зародышем* называется зародыш, размер которого определяется из условия максимума кривой $\Delta F(r)$ при $T < T_{\text{пл}}$. Таким образом, самопроизвольный (спонтанный) рост кристалла, сопровождающийся уменьшением свободной энергии системы, может происходить только при $T < T_{\text{пл}}$, то есть когда система находится в неравновесном состоянии, и для зародышей с размером r выше критического r^* , то есть при $r > r^*$.

Образование зародыша критического размера требует затраты энергии ΔF^* . Найдем ΔF^* и r^* . При этом будем считать, что форма зародыша остается постоянной в процессе его роста до r^* и, учтем, что $S = KV^{2/3}$, где K — коэффициент, учитывающий форму зародыша. Тогда

$$\Delta F = \frac{V\rho}{M}(F'(S) - F'(L)) + K\sigma V^{2/3}. \quad (4.9)$$

Максимальное значение изменения свободной энергии ΔF^* можно найти продифференцировав ΔF по V и приравняв производную нулю. Выполнив это, получим

$$\Delta F^* = \sigma S^*/3. \quad (4.10)$$

Таким образом, энергия, необходимая для образования зародыша критического размера, равна одной трети его поверхностной энергии. Остальные две трети этой энергии компенсируются выигрышем объемной составляющей свободной энергии.

Для зародыша кубической формы с длиной ребра $a = 2r$ ($V = 8r^3$ и $K = 6$) находим размер критического зародыша

$$r^* = \frac{2\sigma M}{[\rho(F'(L) - F'(S))]} \quad (4.11)$$

Полученное выражение не дает в явном виде зависимости критического размера зародыша от температуры, а точнее, от величины переохлаждения $\Delta T = T_{\text{пл}} - T$. Однако если учесть, что

$$\begin{aligned} \Delta F' &= F'(L) - F'(S) = [U'(L) - TS'(L)] - [U'(S) - TS'(S)] \\ &= U'(L) - U'(S) - T[S'(L) - S'(S)] = \Delta L - \Delta LT/T_{\text{пл}}, \end{aligned} \quad (4.12)$$

где $U'(L) - U'(S) = \Delta L$ — скрытая молярная теплота кристаллизации, а $S'(L) - S'(S) = \Delta L/T_{\text{пл}}$, так как при $T = T_{\text{пл}}$ $\Delta F = 0$ и $U'(L) - U'(S) = T_{\text{пл}}[S'(L) - S'(S)]$, то (4.11) преобразуется к виду

$$r^* = \frac{2\sigma M T_{\text{пл}}}{\rho \Delta L \Delta T}. \quad (4.13)$$

Таким образом, критический размер зародыша тем больше, чем больше коэффициент поверхностного натяжения σ на границе между новой и исходной фазами и чем меньше величина переохлаждения. Каждому значению переохлаждения соответствует своя величина равновесного критического зародыша. Зародыш с $r > r^*$ будет расти (появится центр новой фазы), а с $r < r^*$ — уменьшаться в своем размере. Заметим, что зародыш критического размера должен обладать свойствами макрофазы.

Выражение для размера критического зародыша, аналогичное приведенному выше для случая кристаллизации, может быть получено и для случая образования жидкой капли из пересыщенного пара с давлением P (пар характеризуется равновесным давлением P_0 при температуре T).¹¹ Изменение свободной энергии газа, обусловленное его изотермическим сжатием и энергией, затраченной на образование поверхности раздела жидкость–пар, для сферической капли радиусом r выражается формулой:

$$\Delta F = \Delta F_v + \Delta F_s = -\frac{4\pi r^3 \rho}{3M} R_0 T \ln(P_r/P_0) + 4\pi r^2 \sigma, \quad (4.14)$$

где P_r — давление пересыщенного пара над каплей радиуса r при температуре T , P_0 — равновесное давление пара при температуре T , ρ и M — плотность и молекулярный вес зародыша, R_0 — газовая постоянная.

Анализ выражения (4.14) показывает, что если $P_r/P_0 < 1$, то при любых размерах капли ΔF положительна, и капли, возникающие в результате флуктуаций, проявляют тенденцию к испарению. Если $P_r/P_0 > 1$, то при малых r преобладает второй член и значение ΔF положительно, а при достаточно больших r основную роль играет первый член и $\Delta F < 0$ (аналогично выше изложенному). Нетрудно показать, что ΔF^* составляет одну треть от работы образования поверхности зародыша критического размера при конденсации капли из пересыщенного пара (аналогично случаю кристаллизации), а критический радиус капли равен

$$r^* = \frac{2\sigma M}{\rho R_0 T \ln(P_r/P_0)}. \quad (4.15)$$

¹¹Вывод и выражение для размера критического зародыша для случая конденсации в твердую фазу будут аналогичны приведенным ниже для капель, если учесть форму кристаллического зародыша.

Капля размером $r < r^*$ будет испаряться, а размером $r > r^*$ будет расти.

Таким образом, рассмотрение начальных стадий процессов кристаллизации и конденсации показало, что для зарождения и роста кристаллов необходимо, чтобы исходная фаза была переохлаждена или пересыщена по отношению к возникающей в ней новой фазе. Пересыщение при этом выражается одним из следующих способов: $\Delta P = P - P_0$ — абсолютное пересыщение; $\delta = \Delta P/P_0$ — относительное пересыщение; $\alpha = P/P_0$ — коэффициент пересыщения. Жидкая фаза, как было показано выше, характеризуется переохлаждением $\Delta T = T_{\text{пл}} - T$.

Как уже отмечалось выше, образование зародышей новой фазы, требующее затраты энергии, обеспечивается за счет флуктуаций. Вероятность таких флуктуаций согласно статистике Больцмана равна

$$W \sim \exp(-\Delta F/kT). \quad (4.16)$$

Вероятность образования критического зародыша при кристаллизации из переохлажденного расплава нетрудно найти, подставляя размер критического зародыша (4.13) в формулу (4.10), а затем (4.10) в (4.16):

$$W \sim \exp \left[-\frac{32\sigma^3 T_{\text{пл}}^2 M^2}{(\Delta L \Delta T \rho)^2 kT} \right]. \quad (4.17)$$

Аналогичные действия дают выражение для конденсации капель из пересыщенного пара:

$$W \sim \exp \left[-\frac{16\pi\sigma^3 M^2}{3kT(R_0 T \rho \ln(P_r/P_0))^2} \right]. \quad (4.18)$$

Таким образом, вероятность возникновения зародыша критического размера растет с увеличением переохлаждения расплава или пересыщения пара, а размер критического зародыша уменьшается.

Рассчитаем скорость образования центров новой фазы в жидкой или газовой фазе в зависимости от переохлаждения или пересыщения. *Скоростью образования центров новой фазы* V ($\text{см}^{-3}\text{с}^{-1}$) называется число стабильных частиц новой фазы, образующихся в единицу времени в единице объема старой фазы.

Сначала найдем выражение для скорости образования центров новой фазы V в пересыщенном паре. Полагают, что скорость образования центров новой фазы в такой неравновесной системе равна средней скорости, с которой критические зародыши увеличивают свой размер. Предполагая, что концентрация критических зародышей равна $n_{\text{кр}}$,¹² мы получим

¹²Ниже концентрация критических зародышей определяется из условия флуктуационного равновесия с метастабильной исходной фазой.

следующее выражение для скорости образования центров новой фазы в единице объема газовой фазы:

$$V = \omega_{\text{кр}} n_{\text{кр}}, \quad (4.19)$$

где $\omega_{\text{кр}}$ — частота, с которой критические зародыши захватывают атомы или молекулы и становятся центрами новой фазы. Заметим, что в приведенном выражении не учитывается процесс, при котором центр новой фазы в результате распада снова становится критическим зародышем, а также не учитываются процессы объединения зародышей.¹³ Частота захвата частиц критическим зародышем при гомогенном зародышеобразовании капель может быть представлена в виде

$$\omega_{\text{кр}} = a_c q S^*, \quad (4.20)$$

где $q = P/(2\pi m k T)^{1/2}$ — плотность потока частиц с массой m из паровой фазы с давлением P и температурой T , $S^* = 4\pi r^{*2}$ — площадь поверхности зародыша критического размера, a_c — коэффициент конденсации, равный отношению числа конденсирующихся (остающихся на поверхности) частиц к числу ударившихся о поверхность зародыша частиц [37]. Тогда

$$V = a_c q S^* n_{\text{кр}}. \quad (4.21)$$

Пусть общее число зародышей новой фазы мало, то есть существует флуктуационное равновесие с исходной метастабильной фазой. Это предположение позволяет ввести представление о квазиравновесном распределении зародышей критического размера. Тогда согласно статистике Больцмана получаем $n_{\text{кр}} \approx n_1 \exp(-\Delta F^*/kT)$, где $n_1 = P/kT$ — концентрация частиц в газовой фазе (атомов или молекул, не входящих в зародыши).

Более точный анализ показывает, что расчет с использованием статистики Больцмана не вполне корректен (хотя и используется в качестве первого приближения), так как в процессе зарождения зародышей критического размера равновесное состояние не достигается. Уточненное выражение для скорости образования центров новой фазы V можно получить, анализируя образование и распад зародышей с помощью ряда бимолекулярных реакций типа $A + A_{i-1} \rightleftharpoons A_i$, описывающих процесс образования зародышей в газовой фазе путем последовательного присоединения частиц (A — частица, A_i — зародыш из i частиц). В результате анализа этих реакций рассчитывают стационарную концентрацию зародышей в условиях необратимого процесса, когда число зародышей критического размера может уменьшаться как путем их разрастания, так и путем их распада. Учет этих процессов приводит к возникновению в формуле (4.21) дополнительного сомножителя Z — «фактора неравновесности», имеющего вид

¹³ Такое предположение, вообще говоря, несправедливо. Однако далее возможность таких процессов будет учтена.

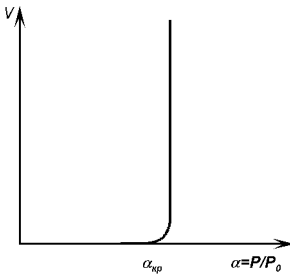


Рис. 4.21. Зависимость скорости гомогенного образования центров новой фазы (жидких капель) в пересыщенном паре от коэффициента пересыщения.

$$Z = (\Delta F^* / 3\pi k T i_{\text{кр}}^2)^{1/2}, \quad (4.22)$$

где $i_{\text{кр}}$ — число частиц в зародыше критического размера.

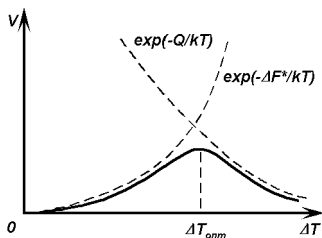
С учетом фактора неравновесности скорость образования центров новой фазы из пара будет определяться выражением:

$$V = 4\pi r^{*2} (a_c P / (2\pi m k T)^{1/2}) (\Delta F^* / 3\pi k T i_{\text{кр}}^2)^{1/2} n_1 \exp(-\Delta F^* / k T). \quad (4.23)$$

Анализ полученного выражения показывает, что, начиная с некоторого пересыщения, наблюдается резкое увеличение скорости зарождения центров новой фазы в зависимости от пересыщения. Это пересыщение называют *критическим* (рис. 4.21).

Иная зависимость скорости образования центров новой фазы от степени переохлаждения наблюдается при кристаллизации. Основное различие случаев зарождения центров новой фазы из расплава и газовой фазы состоит в следующем. В случае газовой фазы, как было показано выше, зарождение центров новой фазы развивается лавинообразно при некотором критическом пересыщении (рис. 4.21). Скорость же зарождения центров новой фазы в вязком расплаве при малых величинах переохлаждения (при температурах расплава немного ниже $T_{\text{пл}}$ материала) близка к нулю, с увеличением переохлаждения расплава она возрастает, достигает максимума при $\Delta T = \Delta T_{\text{опт}}$, затем снижается и при больших величинах переохлаждения обращается в 0 (рис. 4.22). Физическое объяснение такого различия состоит в том, что и на зарождение и на рост из расплава, кроме рассмотренного выше термодинамического фактора, сильное влияние оказывает подвижность атомов (скорость диффузии), которая с увеличением переохлаждения расплава значительно уменьшается. Рост скорости образования центров новой фазы при малых ΔT обусловлен тем, что вблизи $T_{\text{пл}}$ подвижность атомов велика, а вероятность возникновения критических зародышей растет с увеличением ΔT .

Рис. 4.22. Зависимость скорости гомогенного образования центров кристаллизации в единице объема расплава от степени его переохлаждения.



(см. 4.17). Уменьшение же скорости зарождения центров кристаллизации с увеличением ΔT (большие степени переохлаждения) вызвано тем, что при низких температурах подвижность атомов уменьшена, а тем самым уменьшена и способность системы к превращению. При очень больших ΔT скорость зарождения становится равной нулю, так как подвижность атомов уже недостаточна для того, чтобы осуществилась перестройка их из хаотического расположения в жидкости в правильное в кристалле.

Таким образом, в случае кристаллизации из расплава на скорость зарождения помимо термодинамического фактора существенное влияние оказывает и кинетический фактор. Этот фактор оказывает такое же влияние и на линейную скорость роста кристаллической фазы. Кинетический фактор учитывается с помощью множителя $\exp(-Q/kT)$, где Q — энергия активации процесса, контролирующего подвижность атомов. Тогда скорость образования центров кристаллизации в единице объема расплава будет

$$V \sim \exp(-Q/kT) \exp(-\Delta F^*/kT). \quad (4.24)$$

Вероятность возникновения критических зародышей, определяемая значением ΔF^* , растет с увеличением переохлаждения расплава. В то же время с ростом переохлаждения расплава понижается подвижность атомов, определяемая Q , особенно в вязких расплавах, и тем самым замедляется образование центров кристаллизации даже при сильных переохлаждениях. Поэтому для каждой жидкости существует оптимальное переохлаждение, при котором скорость возникновения центров кристаллизации максимальна (рис. 4.22).

При кристаллизации невязких жидкостей (например, расплавы кремния и германия) кинетический фактор играет меньшую роль, и в результате этого максимума в скорости образования центров может и не наблюдаться. В этом случае при достижении некоторого критического переохлаждения происходит резкое увеличение скорости образования цен-

тров кристаллической фазы и наступает быстрая кристаллизация всего объема жидкости.

Коротко изложенная теория позволяет оценить критическое пересыщение и оптимальное переохлаждение для начала гомогенного образования центров новой фазы. При кристаллизации из расплава $\Delta T_{\text{опт}}$ составляет $\sim 0.2T_{\text{пл}}$; в случае конденсации из газовой фазы критический коэффициент пересыщения $\alpha_{\text{кр}} = P_r/P_0 \approx 10$.

Однако наблюдаемые в реальных условиях значения переохлаждения или пересыщения во многих случаях оказываются намного ниже указанных величин. Это связано главным образом с наличием в реальных условиях факторов, снижающих величину составляющей поверхностной энергии в изменении свободной энергии при образовании зародыша новой фазы и соответственно уменьшающих энергию образования зародыша критического размера и критическое пересыщение или переохлаждение. Параметром, с помощью которого можно существенно изменять энергию образования критического зародыша, а следовательно, и необходимую величину пересыщения (переохлаждения), является поверхностное натяжение σ . Управлять процессом фазовых превращений, изменяя поверхностное натяжение σ , можно двумя путями: введением специальных примесей в исходную фазу или созданием готовых поверхностей раздела. Растворимые в расплаве примеси, располагающиеся в поверхностных слоях центров новой фазы и уменьшающие их свободную поверхностную энергию, называются *поверхностно-активными примесями*. В случае наличия готовых поверхностей раздела зародышеобразование происходит по гетерогенному механизму.

Гетерогенное образование зародышей новой фазы

Гетерогенным называется образование зародышей новой фазы в исходной фазе на готовых поверхностях раздела, сопровождающееся образованием лишь части поверхности зародыша. К гетерогенным фазовым превращениям на готовых поверхностях раздела относятся кристаллизация на затравках при выращивании монокристаллов, на подложках при эпитаксиальном росте, на нерастворимых твердых частицах в расплаве. К гетерогенным превращениям по существу относятся и все случаи роста зародышей, возникших гомогенным путем.

Рассмотрим процесс гетерогенного зародышеобразования на примере конденсации насыщенного пара на гладкую плоскую чистую подложку.

Поскольку свободная поверхностная энергия жидкости изотропна, то зародыш, образующийся на плоской подложке и находящийся в равнове-

ном состоянии, в общем случае имеет форму шарового сегмента с радиусом кривизны поверхности r , либо форму монослойного диска (рис. 4.23). При этом равновесный угол смачивания ϑ определяется из условия равновесия натяжений трех поверхностей раздела: $\gamma_{sv} = \gamma_{cs} + \gamma_{cv} \cos \vartheta$, где γ_{sv} , γ_{cs} и γ_{cv} — удельные свободные поверхностные энергии границ раздела подложка–пар, конденсат–подложка и конденсат–пар соответственно.

Изменение свободной энергии при образовании на подложке куполообразного зародыша с радиусом кривизны r при гетерогенном зарождении, так же как и при гомогенном зародышеобразовании, можно представить в виде двух составляющих — ΔF_v и ΔF_s . Объемная составляющая записывается в виде

$$\Delta F_v = -\frac{4\pi r^3}{3} \frac{\rho h o}{M} \frac{(2 + \cos \vartheta)(1 - \cos \vartheta)^2}{4} R_0 T \ln \frac{P_r}{P_0}. \quad (4.25)$$

Поверхностная составляющая изменения свободной энергии дается выражением:

$$\begin{aligned} \Delta F_s &= \Sigma \gamma_i s_i = \gamma_{cv} s_{cv} + \gamma_{cs} s_{cs} - \gamma_{sv} s_{sv} \\ &= 4\pi r^2 \gamma_{cv} \frac{(1 - \cos \vartheta)}{2} + \pi r^2 (\gamma_{cs} - \gamma_{sv}) \sin^2 \vartheta \\ &= 4\pi r^2 \gamma_{cv} \frac{(2 + \cos \vartheta)(1 - \cos \vartheta)^2}{4}, \end{aligned} \quad (4.26)$$

где $4\pi r^2 \gamma_{cv} \frac{(1 - \cos \vartheta)}{2}$ — поверхность шарового сегмента, $\pi r^2 \sin^2 \vartheta$ — площадь соприкосновения капли с подложкой, а $\gamma_{cs} - \gamma_{sv} = -\gamma_{cv} \cos \vartheta$. В результате всех подстановок и вычислений изменение свободной энергии системы при гетерогенном образовании зародыша будет определяться выражением [37]:

$$\Delta F_{\text{гет}} = \Delta F_{\text{гом}} \frac{(2 + \cos \vartheta)(1 - \cos \vartheta)^2}{4}. \quad (4.27)$$

Как следует из (4.27), свободная энергия при гетерогенном зародышеобразовании отличается от свободной энергии при гомогенном только множителем, зависящим от ϑ . Из этого следует, что выражение для r^* при гетерогенном зарождении дает такую же зависимость от пересыщения, что и при гомогенном образовании зародышей.

Случай $\vartheta = 180^\circ$ соответствует полной несмачиваемости подложки жидкостью, что происходит, когда взаимодействие между частицами подложки и жидкости мало по сравнению с взаимодействием частиц в жидкости. В этом случае нет никакого энергетического выигрыша и, следо-

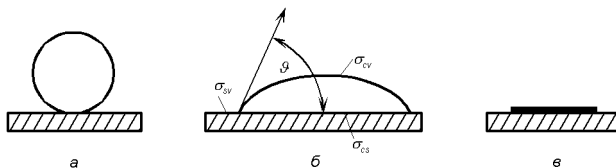


Рис. 4.23. Зависимость формы зародыша жидкой фазы при гетерогенном зарождении от равновесного угла смачивания ϑ : а — $\vartheta = 180^\circ$; б — $0 < \vartheta < 180^\circ$; в — $\vartheta = 0^\circ$.

вательно, подложка не оказывает каталитического действия на процесс зародышеобразования (рис. 4.23,а).

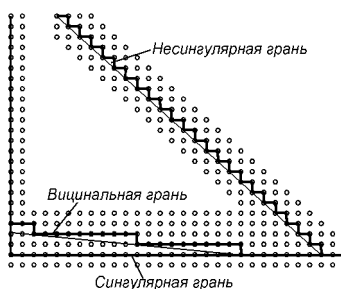
Случай $\vartheta = 0^\circ$ соответствует полной смачиваемости. В этом случае силы межатомного взаимодействия, структура, состав капли и подложки практически идентичны (рис. 4.23,в). На подложке при этом образуются двумерные зародыши, равновесной формой которых в простейшем случае является монослойный диск (зародыши, образующиеся на плоскости). В этом случае получается максимальный выигрыш в работе образования зародыша. Из выражения (4.27) следует, что $\Delta F_{\text{гет}}^* = 0$. Однако на самом деле работа образования зародыша не может стать равной нулю, так как при выводе (4.27) мы пренебрегли работой образования линейного контура зародыша в силу его малости по сравнению с объемной и поверхностной составляющими изменения свободной энергии при гетерогенном фазовом превращении.¹⁴ Образование зародыша в этих условиях фактически представляет собой рост кристалла путем присоединения двумерных зародышей, что также требует затраты энергии, то есть нужно определенное пересыщение пара, хотя и меньшее, чем в случае образования трехмерных зародышей.

Случай $0^\circ < \vartheta < 180^\circ$ является промежуточным и наиболее распространенным. Он всегда более энергетически выгоден, чем случай гомогенного зарождения. Здесь значения r^* и ΔF^* тем меньше, чем меньше угол ϑ , и подложка, таким образом, является катализатором процесса зародышеобразования.

Вопросы о границах применимости рассмотренной термодинамической теории зародышеобразования, о скорости гетерогенного образования зародышей на единице площади и о влиянии подложки на процессы

¹⁴Эта работа может быть одного порядка с поверхностной составляющей ΔF_s только при малых $r^* \sim 10^{-9}$ м. Выражения для $\Delta F_{\text{гет}}^*$ и r^* с учетом этого вклада можно найти в [37]. В этом случае оказывается, что $\Delta F_{\text{гет}}^* \sim 1/[R_0 T \ln(P_r/P_0) - (\gamma_{cv} + \gamma_{cs} - \gamma_{sv})]$.

Рис. 4.24. Структура поверхности граней разных ориентаций в простой кубической решетке.



зародышеобразования будут подробно рассмотрены в гл. 9, посвященной эпитаксиальному росту пленок.

4.2.2. Рост центров новой фазы

Любое фазовое превращение включает в себя не только образование зародышей новой фазы, но и их рост. С точки зрения термодинамики рост образовавшихся флуктуационным путем кристаллических зародышей должен происходить при сколь угодно малых пересыщениях в исходной фазе (случай полного смачивания). Однако многочисленные экспериментальные исследования показывают, что при заданном пересыщении скорость роста грани кристалла зависит от ряда других, кроме пересыщения, факторов, и прежде всего, от морфологии поверхности растущей грани кристалла. Поверхности граней идеальных кристаллов по своему атомному строению принято подразделять на три типа: *сингулярные*, *вициальные* и *несингулярные*.

Сингулярными называются атомно-гладкие грани, не имеющие никаких ступеней (рис. 4.24). Например, грань (100) у простой кубической решетки, грань (111) у алмазоподобной решетки. Сингулярные грани по сравнению с другими обладают наименьшей свободной поверхностной энергией, наибольшей ретикулярной плотностью (плотностью упаковки атомов) и характеризуются малыми индексами Миллера.

Вициальными называются грани с ориентациями, достаточно близкими к сингулярным граням (рис. 4.24). Вициальные грани, образующие небольшой угол с сингулярными, состоят из больших плоских участков сингулярных граней, отделенных друг от друга ступенями монокристаллической высоты. Из-за существования ступеней вициальные грани характеризуются более высокой поверхностной энергией, чем сингулярные грани.

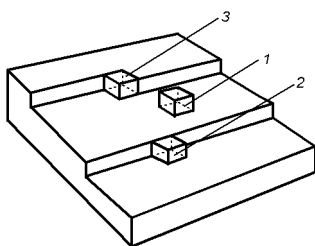


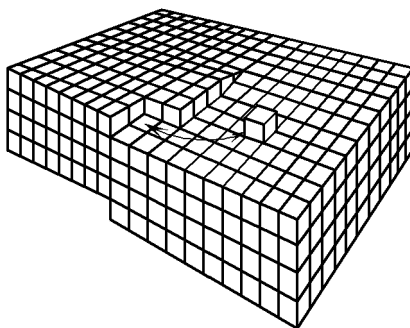
Рис. 4.25. Схема роста атомно-гладкой поверхности кристалла (поз. 1), на ступеньках (поз. 2) и на изломах (поз. 3).

Несингулярные грани составляют достаточно большие углы с сингулярными и имеют высокую концентрацию ступеней (рис. 4.24). Эти грани обладают наибольшей поверхностной энергией.

В зависимости от морфологии растущей поверхности механизм роста кристаллов может быть послойным, спиральным и нормальным. Он связан с характером расположения атомов в решетке кристалла и с характером межатомных взаимодействий. Теория роста кристаллов основана на анализе сил связи, действующих между атомами в кристалле, с помощью теории химической связи. А поскольку теория химических связей носит полуколичественный характер, то для описания механизмов роста кристаллов используют простые модели, например, представляют атомы в виде простых кубиков, плотная упаковка которых позволяет получать как гладкие грани, так и шероховатые.

Послойный рост осуществляется на сингулярных и вицинальных гранях. Основоположниками теории послойного роста являются Коссель, Странский и Фольмер. Согласно их теории атом, попадающий на поверхность растущего идеального кристалла, наиболее прочно связывается в изломе ступени (рис. 4.25, поз. 3), так как в этой позиции атом образует связи с тремя из шести ближайших соседей. На самой ступени (рис. 4.25, поз. 2) связи образуются лишь с двумя, а на гладкой поверхности (рис. 4.25, поз. 1) только с одним из шести соседей. Атом, попадающий из внешней фазы на поверхность кристалла, отдает часть своей энергии решетке и, попадая в поле действия сил связи поверхностных атомов кристалла, переходит в адсорбированное состояние. Как правило, адсорбированные атомы обладают еще достаточным запасом энергии, чтобы передвигаться по поверхности кристаллов. Их средний миграционный путь составляет несколько сотен межатомных расстояний. Следовательно, хотя вероятность попадания атома из внешней фазы в изломы на ступеньках мала, атомы достигают изломов путем поверхностной диффузии сначала к ступенькам, и далее, передвигаясь вдоль них до из-

Рис. 4.26. Схема растущей грани кристалла, содержащей выход винтовой дислокации.



ломов, встраиваются в кристалл. Когда ряд завершается, атом должен присоединиться к ступени, образовав новый излом. Зарождение излома требует более высокого пересыщения, поскольку адсорбированный на ступени атом образует связи только с двумя ближайшими соседями. Однако, как было замечено Френкелем, образование изломов на ступенях не лимитирует скорости их движения, так как вследствие наличия тепловых флуктуаций во внешней среде ступени имеют значительное число изломов, количество которых пропорционально $\exp(-w/kT)$, где w — энергия, необходимая для образования излома на ступени. В реальных кристаллах вероятность образования ступеней и изломов зависит также от ретикулярной плотности кристаллографических плоскостей. У плоскостей с малыми индексами Миллера она меньше, чем у плоскостей с большими индексами.

По мере движения ступени и излома происходит застройка плоскости кристалла. Для дальнейшего роста необходимо образование двумерного зародыша, для чего требуются уже значительные пересыщения. Время ожидания (вероятность) такой флуктуации будет больше (меньше), чем время (вероятность), необходимое для образования изломов на ступенях. Поэтому согласно теории Косселя-Странского-Фольмера сингулярные грани должны расти прерывистым образом, и для их роста необходимо критическое пересыщение, которое обеспечивало бы образование двумерных зародышей.

Однако в реальных условиях прерывистый рост сингулярных граней кристалла наблюдается редко. Тот факт, что кристаллы в большинстве случаев растут с измеримыми скоростями даже при очень малых пересыщениях, привел Франка к заключению, что такие кристаллы имеют на поверхности роста постоянно действующий источник ступеней, в ка-

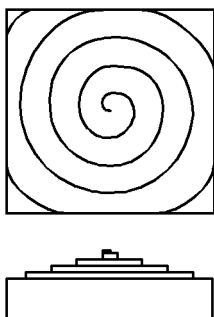


Рис. 4.27. Морфология поверхности роста, обусловленная действием винтовой дислокации.

честве которого может выступать винтовая дислокация. Действительно, при выходе на поверхность винтовая дислокация дает незарастающую ступень (рис. 4.26), и потребность в создании двумерных зародышей уже не возникает. Рост кристалла в этих условиях идет при низких пересыщениях и без прерывания. Процесс роста с помощью винтовой дислокации называется *слоисто-спиральным* механизмом роста кристалла.

Слоисто-спиральный механизм роста аналогичен описанному механизму роста совершенного кристалла со ступенью (только ступенька в нашем случае незарастающая). На ступени, возникающей благодаря винтовой дислокации, имеются изломы вследствие существования тепловых флуктуаций. Адсорбированные атомы диффундируют к ступени, а затем к изломам, где они встраиваются в решетку кристалла, в результате чего ступень движется. Поскольку один конец ступени зафиксирован в точке выхода дислокации, то ступень может двигаться только путем вращения вокруг этой точки. При определенном пересыщении каждый участок на прямой ступеньке движется с одинаковой линейной скоростью. Поэтому участок ступеньки вблизи линии дислокации имеет более высокую угловую скорость и за одинаковое время должен сделать большее число оборотов, чем далеко отстоящие от линии дислокации участки. По мере увеличения кривизны участка ступени в области выхода дислокации равновесное давление пара над этим участком повышается, местное пересыщение понижается и, следовательно, линейная скорость движения этой части ступени замедляется. Спираль закручивается до тех пор, пока радиус кривизны в центре ее не достигнет значения критического радиуса двумерного зародыша. По достижении стационарного состояния спираль вращается как единое целое вокруг линии дислокации, при этом форма ее приближенно может быть описана уравнением архимедовой спирали.

Нормальный рост кристаллов осуществляется на несингулярных гра-

нях. Эти грани, будучи атомно-шероховатыми, равномерно покрыты изломами, и присоединение новых частиц происходит на них практически в любом месте. Такие грани растут перпендикулярно самим себе, то есть осуществляется «нормальный» механизм роста. Процесс роста несингулярных граней идет непрерывно при любых пересыщениях. Поскольку концентрация точек роста на несингулярных гранях значительно выше, чем на сингулярных, то и скорость роста несингулярных граней значительно выше. В процессе роста кристалла несингулярные грани могут исчезнуть или выклинить, вследствие чего кристалл часто принимает огранку и оказывается покрытым медленно растущими, то есть преимущественно сингулярными, гранями с низкими индексами Миллера.

Поверхностная кинетика роста кристаллов. Рост кристаллов в основном осуществляется по слоистому или слоисто-спиральному механизмам. Теория слоисто-спирального роста кристаллов впервые была создана применительно к конденсации из газовой фазы Бартоном, Кабрерой и Франком. Эта теория занимает центральное место для роста кристаллов из газовой фазы, но она существенна также и для роста из разбавленных растворов, и даже в какой-то степени для роста из расплавов. Рассмотрим коротко основные положения этой теории.

Пусть рост кристалла из газовой фазы в основном происходит по слоистому или слоисто-спиральному механизмам. В этом случае источниками ступеней на растущих сингулярных (вицинальных) гранях могут быть винтовые дислокации. Ступень, образованная винтовой дислокацией, при встраивании в нее частиц закручивается в спираль, и образующиеся последовательные витки формируют эшелон ступеней. На растущей поверхности при этом возникают пирамиды (рис. 4.27), причем концентрация ступеней, образующих эти пирамиды, велика и практически не зависит от количества винтовых дислокаций, выходящих на поверхность роста. На поверхности кристалла, контактирующего с питающей средой, присутствуют адсорбированные частицы того же вещества, из которого состоит кристалл. Адсорбированные частицы совершают тепловые колебания в трех направлениях — перпендикулярно плоскости и в двух параллельных плоскости. Флуктуации энергии при колебаниях первого типа приводят к отрыву частиц от поверхности и переходу их в газовую среду (испарение). Колебания второго типа создают условия для диффузионной миграции этих частиц по поверхности. Если над растущей поверхностью создается пересыщение, то начинается диффузия в окружающей среде и адсорбированном слое по направлению к ступени, на которой будет идти конденсация до тех пор, пока это пересыщение не

исчезнет. При этом надо учитывать, что при конденсации из газовой фазы количество частиц, попадающих на торец ступени непосредственно из пара мало ввиду малых плотности пара и площади поверхности торца. Поэтому основной вклад в рост кристалла в этом случае дает плоский диффузионный поток частиц, адсорбированных на различных участках поверхности. При этом нормальная скорость роста поверхности V по слоисто-спиральному механизму будет при низких пересыщениях:

$$V \sim (\Delta P)^2 \exp(-E/kT), \quad (4.28)$$

а при больших пересыщениях

$$V \sim \Delta P \exp(-E/kT), \quad (4.29)$$

где ΔP — пересыщение, а E — энергия испарения частицы.

Морфологическая устойчивость растущего кристалла. Рост кристалла представляет собой процесс, который может протекать только при отклонении системы от равновесия. Термодинамическое изучение процессов роста кристалла в неравновесных условиях показало возможность образование целого ряда устойчивых форм кристалла.

Действительно, морфология растущей грани и, соответственно, форма кристалла определяется как поверхностной кинетикой роста, так и условиями переноса (тепла и массы) в соприкасающихся фазах (питающей среды и кристалла). С увеличением размеров растущих граней кристалла изменяются условия тепло- и массопередачи, что приводит к неоднородности пересыщения (переохлаждения) у растущей грани и способствует нарушению ее плоской формы. В этих условиях флуктуационное появление на грани небольших искажений усиливает неоднородность процессов переноса и еще больше искажает форму растущего кристалла. Однако кинетические явления на растущей поверхности выступают в качестве факторов, стабилизирующих плоскую форму грани. Это связано с тем, что при образовании выступа или впадины на сингулярной грани появляются боковые поверхности, которые растут с большими скоростями, чем сингулярные. Это приводит к сглаживанию формы грани.

В результате конкуренции противоположно действующих факторов растущая плоская грань может потерять свою устойчивость, и изотермическая форма кристалла может в дальнейшем перейти в скелетообразную или дендритную (то есть превратиться в древовидный кристалл, состоящий из центрального ствола, первичных, вторичных и т. д. ветвей).

Дальнейшее, более детальное, обсуждение процессов тепло- и массопереноса и кинетики роста кристаллов из жидкой и газовой фаз будет

продолжено в гл. 6, посвященной методам роста полупроводниковых монокристаллов.

Глава 5

Получение чистых полупроводниковых материалов

5.1. Кристаллизационные методы очистки

Электрофизические свойства кристаллов определяются, как было выяснено в гл. 3, содержащимися в них структурными дефектами и примесями. Требование продолжительности и стабильности работы полупроводниковых приборов делает одной из важнейших задач технологии задачу получения совершенных монокристаллов с заданным значением параметров. Однако получение чистых элементарных веществ, необходимых для производства различных, в том числе легированных и сложных, полупроводников, используемых для создания приборов, является чрезвычайно сложным технологическим процессом.

В основе всех способов очистки полупроводниковых материалов лежит различие в химических и физических свойствах разделяемых веществ. При существенном различии в свойствах разделяемых материалов их легко отделить, и, наоборот, проблема очистки становится сложной, если очищаемый материал и примесь очень близки по своим физико-химическим характеристикам. Как правило, для получения чистых веществ применяют многоступенчатые технологические схемы, включающие различные химические и металлургические методы очистки.

Анализ содержания примесей в чистых веществах осуществляется с применением широкого круга химических, физико-химических и физических методов. Наиболее перспективным для анализа чистых веществ и полупроводников является масс-спектральный анализ, который позволяет в ходе одного сравнительно нетрудоемкого анализа определить большое число элементов с высокой точностью. Другой, не менее перспективный, но более трудоемкий и длительный метод анализа — радиоактивационный — позволяет

исключить загрязнения, вносимые в анализируемый материал в ходе самого анализа. Чувствительность радиоактивационного метода очень велика. Для полупроводниковых материалов наибольшее распространение получили косвенные, физические методы оценки чистоты на основании измерения концентрации носителей заряда, удельного сопротивления, подвижности и др. Физические методы анализа чистых веществ гораздо менее трудоемки и длительны, чем обычные аналитические методы определения содержания примесей. Однако результаты, полученные такими методами, дают представление только о суммарном содержании примесей. Тем не менее они позволяют надежно судить о возможности использования проанализированных чистых веществ в технологии полупроводниковых материалов.

Понятие химической чистоты материалов, а также способы ее выражения различны и зависят от области применения материала. В нашей стране для чистых веществ, используемых в химической и металлургической практике, в зависимости от степени очистки установлены следующие классификации: «чистый» (марка Ч, содержание примесей от $2 \cdot 10^{-5}$ до 1.0%), «чистый для анализа» (марка ЧДА, содержание примесей от $1 \cdot 10^{-5}$ до 0.4%), «химически чистый» (марка ХЧ, содержание примесей от $5 \cdot 10^{-6}$ до 0.5%) и «особо чистый» (марка ОЧ, содержание примесей 0.05%). Особо чистые вещества для полупроводниковой техники разделяют на классы А, В и С. В класс А входят вещества, чистоту которых по *содержанию основного компонента* можно надежно охарактеризовать современными аналитическими методами. Классы чистоты В и С характеризуют чистоту по *содержанию определяемых примесей*. В последнем случае о содержании основного компонента можно говорить только условно, понимая под этим разницу между 100% и суммарным содержанием определяемых примесей.

Неметаллы и химические соединения относят к высоко чистым веществам, если содержание лимитирующих примесей в них не менее чем на порядок ниже по сравнению с соответствующей маркой ХЧ, на два порядка — для марки ЧДА и на три порядка — для марки Ч. К высоко чистым веществам принято относить также металлы и полупроводниковые материалы, если содержание каждой из контролируемых примесей в них не превышает $1 \cdot 10^{-4}\%$ (по массе).

Высоко чистым веществам в зависимости от количества и суммарной концентрации контролируемых примесей присваивают марки двух категорий: ВЭЧ («вещества эталонной чистоты») и ОСЧ («особо чистые вещества»). В веществах марки ВЭЧ лимитируется общее содержание примесей и содержание некоторых нежелательных примесей. В веществах марки ОСЧ лимитируется содержание большого числа нежелательных примесей.

В зарубежной литературе принято квалифицировать вещества по чистоте числом (n) атомов (молекул) примесей, приходящихся на: 1) 1000 атомов (молекул) основного вещества, что обозначают п‰ (p promille); 2) 10^6 атомов (молекул) основного вещества, p ppm (p parts per million); 3) 10^9 атомов (молекул) основного вещества, p ppb (p parts per billion). Кроме такой квалификации для обозначения веществ высокой степени чистоты используют символ N и различают вещества по качеству цифровыми индексами.

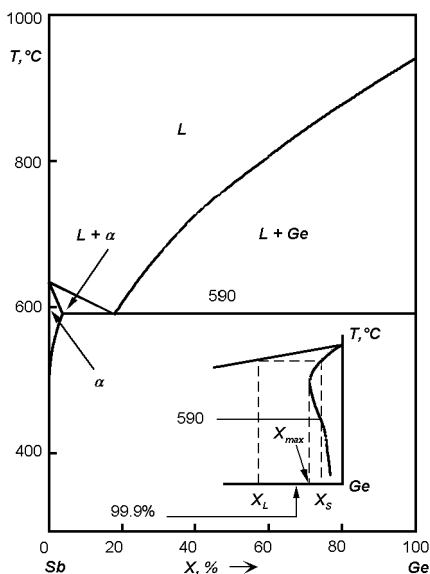
До сих пор речь шла о химической чистоте вещества. Однако понятие чистоты вещества значительно шире. В частности, в электронной технике при использовании полупроводников важна не только степень химической чистоты, но и отсутствие собственных дефектов в кристаллической решетке. Эти дефекты, так же как и химические примеси, сильно влияя на свойства полупроводников. Следует отметить, что концентрации собственных дефектов («физическая» чистота) и примесей (химическая чистота) взаимосвязаны, поскольку структурное совершенство кристаллов во многом зависит от наличия примесей в кристаллической решетке, а предел химической чистоты кристалла часто определяется типом и количеством собственных дефектов, образующихся при выращивании.

Итак, получение чистых полупроводниковых материалов производится химическими и металлургическими методами. Химическая очистка, например, германия и кремния заключается в получении их летучих соединений (как правило, галогенидов, см. гл. 6), которые легко отделяются от примесей и их соединений, а затем в восстановлении предварительно очищенных летучих соединений до элементарных Ge и Si. Металлургические методы заключаются в многократной перекристаллизации слитка основного вещества путем последовательного расплавления его участков. В этом курсе мы не будем рассматривать химические методы очистки веществ, так как они специфичны для каждого вещества, а остановимся на металлургических (кристаллизационных) методах очистки, при которых вещество может очищаться в процессе его выращивания. Кристаллизационные методы очистки основаны на различии содержания примесей в жидкой и твердой фазах, находящихся в равновесии при данной температуре (см. рис. 5.1). Следствием этого свойства является отеснение примеси к концу слитка или ее захват в начальных частях слитка в ходе кристаллизации, то есть очистка вещества от примеси в любом случае.

В полупроводниковой технологии кристаллизация из жидкой фазы, как правило, применяется на последнем этапе технологического процесса очистки вещества. При этом наряду с высокой степенью очистки материала от примесей достигается и необходимое совершенство кристаллической структуры, то есть в конечном счете полупроводники получают в виде высокочистых монокристаллов.

Ge и Si — это два важнейших полупроводниковых элемента. Их способность к взаимодействию друг с другом и с другими химическими элементами имеет большое значение для электронной промышленности, поэтому процессы очистки мы будем рассматривать на их примере. Для разработки методов кристаллизационной очистки вещества необходимо знание фазовой диаграммы состояния вещества с присутствующими в нем примесями. Ge и Si кристаллизуются в кубической структуре типа алмаза и могут образовывать неограниченные твердые растворы только с немногочисленными элементами, которые также имеют структуру типа алмаза и атомные радиусы не сильно отличающиеся от атомных радиусов Ge и Si. Как правило, растворимость большинства примесей в Ge и Si очень невелика (≈ 0.1 ат.%). Кроме того, как уже упоминалось, кристаллизационные методы применяются на конечной стадии процесса очистки, а содержание остаточных примесей в очищаемом веществе настолько мало, что их взаимное влияние или взаимодействие между собой в среде основного материала практически отсутствуют. Все это позволя-

Рис. 5.1. Диаграмма состояния системы Ge–Sb.



ет при описании кристаллизационных методов очистки заменить одну сложную многокомпонентную систему совокупностью фазовых диаграмм бинарных систем: основной компонент — примесь.

Для бинарных систем диаграммы плавления–затвердевания, как уже обсуждалось ранее, весьма разнообразны и сложны. На рис. 5.1 приведена типичная диаграмма полупроводниковой системы: диаграмма состояния системы Ge–Sb. Однако для небольших концентраций примеси в полупроводнике все сложные диаграммы можно свести к двум типам без потери общности рассуждений при описании процессов кристаллизационной очистки. Действительно, в области малых концентраций примеси, то есть в области, примыкающей к точке плавления чистого компонента, различия между разными типами фазовых диаграмм пропадают: линии ликвидуса L и солидуса S в этих областях можно аппроксимировать прямыми линиями, касательными к кривым L и S в точке плавления основного компонента (рис. 5.2). В результате имеем два типа диаграмм: 1) если прямые L и S направлены вверх, то примесь понижает температуру плавления (рис. 5.2,а); 2) если прямые направлены вниз, то примесь повышает температуру плавления (рис. 5.2,б).

Для фиксированной температуры, при которой в системе сосуществуют жидкая и твердая фазы, концентрации примеси в кристалле и в жид-

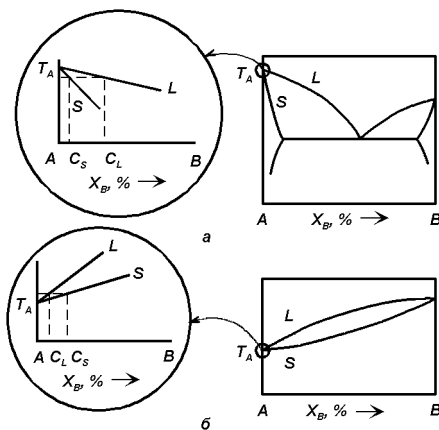


Рис. 5.2. Аппроксимация участка фазовых диаграмм вблизи точки плавления основного компонента отрезками прямых, касательных к кривым солидуса S и ликвидуса L в точке плавления основного компонента T_A : a — примесь, понижающая температуру плавления; b — примесь, повышающая температуру плавления.

кости оказываются различными: концентрация примеси в твердой фазе C_S определяется по абсциссе точки пересечения коноды с линией солидуса, а концентрация примеси в жидкой фазе C_L — по абсциссе точки пересечения коноды с линией ликвидуса (рис. 5.2). В случае, когда примесь понижает температуру плавления, C_S оказывается меньше C_L (рис. 5.2, a), а в случае, когда примесь повышает температуру плавления, $C_S > C_L$ (рис. 5.2, b). В процессе кристаллизации примесь будет либо отесняться в расплав (в первом случае), либо захватываться кристаллом (во втором случае).

Действительно, для любой температуры из интервала, соответствующего области сосуществования жидкой и твердой фаз, концентрации основного вещества и примеси в этих фазах различны. Тем не менее, в равновесных условиях кристаллизации после завершения процесса весь закристаллизовавшийся слиток должен иметь состав, равный составу расплава (см. гл. 4). Соответственно этому не должно возникать никаких закономерных неоднородностей состава в кристаллизующемся материале, то есть в результате перекристаллизации никакой очистки твердой фазы не происходит. Однако в реальных условиях выращивания кристаллов твердая фаза кристаллизуется неоднородной по своему составу. Это происходит из-за медленных диффузионных процессов выравнивания со-

става в твердой фазе (предполагают, что равновесие в жидкой фазе и на границе раздела фаз успевает непрерывно устанавливаться, а диффузия в твердой фазе идет настолько медленно, что изменениями в составе твердой фазы можно пренебречь). Кроме того, при выращивании кристаллов создают такие градиенты температур, при которых нарастающие слои твердой фазы попадают в область с пониженной температурой, что также снижает скорость диффузионного выравнивания состава. Все эти процессы приводят к фиксации неоднородностей состава, возникающих в твердой фазе, и к очистке веществ при кристаллизации.

Если затруднена диффузия в жидкой фазе, то на границе раздела фаз в ходе процесса будет увеличиваться избыток компонента, которым обеднена образующаяся твердая фаза, и недостаток компонента, которым твердая фаза обогащена. Кристаллизация твердой фазы при этом будет происходить не из основной массы жидкой фазы, а из прилегающего к фронту кристаллизации слоя, что приводит к снижению эффекта очистки. Таким образом, затрудненность диффузии компонентов в жидком состоянии снижает эффективность очистки при кристаллизации. Еще одним важнейшим фактором, влияющим на эффективность очистки, являются условия проведения кристаллизации, которые должны обеспечивать монокристалличность получаемого материала.

5.2. Коэффициент разделения примесей

Для оценки эффективности очистки веществ от примесей кристаллизацией из расплава служит *коэффициент разделения K* , представляющий собой отношение концентрации растворенного вещества в твердой фазе к его концентрации в жидкой фазе:

$$K = C_S/C_L. \quad (5.1)$$

Значение K может быть больше или меньше 1. Если примесь понижает температуру плавления вещества, то $K < 1$, что часто встречается на практике. В обратном случае, когда примесь повышает температуру плавления вещества, $K > 1$. Этот случай в технологии полупроводников встречается сравнительно редко. Значение K может изменяться в широких пределах от величины менее 10^{-6} до величины более 10. Величина K определяется типом фазовой диаграммы основного компонента и примеси и зависит от условий кристаллизации (скорости перемещения фронта кристаллизации, интенсивности перемешивания расплава и т. д.). Величина K играет большую роль в кристаллизационных способах очистки.

Различают *равновесный* и *эффективный коэффициенты разделения*.

5.2.1. Равновесный коэффициент разделения

С помощью *равновесного коэффициента разделения* K_0 характеризуют эффект разделения компонентов смеси в *каждый момент равновесия* между кристаллом и расплавом. Его называют равновесным именно потому, что он отражает состав фаз на диаграмме состояния, построенной в близких к равновесию условиях, то есть при достаточно малых скоростях кристаллизации. В этом случае концентрация примеси, растворенной в расплаве C_L , одинакова во всем объеме расплава, а ее концентрация в закристаллизовавшейся части $C_L = K_0 C_S$.

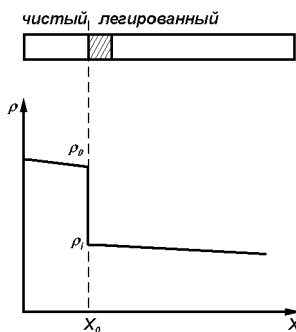
Как можно видеть из рис. 5.2, в пределах прямолинейности L и S линий K_0 не зависит от концентрации и температуры, то есть является постоянной величиной для данной пары веществ (это следует из подобия прямоугольных треугольников, построенных на отрезках C_S и C_L при разных температурах кристаллизации). Однако при достаточно высоких концентрациях примеси, когда линии S и L нельзя аппроксимировать прямыми, K_0 уже не будет постоянной величиной и будет зависеть от концентрации примеси.

Точный расчет K_0 на основе фазовых диаграмм часто затруднен, поскольку во многих случаях имеющиеся фазовые диаграммы недостаточно точны в области малых концентраций примеси. Фазовые диаграммы поэтому часто используются только для примерной оценки K_0 ($K_0 > 1$ или $K_0 < 1$), что позволяет принципиально оценить поведение примеси в условиях очистки вещества кристаллизационными методами.

Наиболее распространенный метод определения коэффициента разделения основан на измерении концентрации свободных носителей заряда в чистом и легированном полупроводнике (предполагается, что примеси электрически активны). Схема определения непосредственно вытекает из самого смысла $K_0 = C_S/C_L$, то есть сводится к определению концентрации примеси в твердой фазе, соответствующей ее заданной концентрации в жидкой фазе. Пусть навеска электрически активной примеси p_i вводится в расплав предварительно очищенного слитка при положении X_0 фронта кристаллизации¹ (рис. 5.3; кристаллизация полупроводникового материала проводится методом зонной плавки (см. ниже)). Если концентрация введенной примеси $N_i = C_S$ заметно превышает разность

¹ Движущаяся граница между твердой и жидкой фазами называется фронтом кристаллизации.

Рис. 5.3. Изменение удельного сопротивления ρ по длине слитка, в который в точке X_0 добавлена примесь.



концентраций остаточных примесей $|N_d^0 - N_a^0|$ в чистом материале, то граница, отделяющая чистый и легированный материал, может быть надежно выявлена по изменению удельного сопротивления материала ρ (рис. 5.3). Вырезав два образца из участков кристалла, прилегающих к границе легирования, и измерив в них коэффициент Холла, можно определить концентрации носителей заряда в нелегированном (n_0 или p_0) и легированном (n или p) образцах. Считая для простоты примеси однозарядными и полностью ионизованными при температуре измерений, можно определить концентрацию N_i введенной примеси из следующих соотношений:

1. Если нелегированный образец n -типа, то при введении донорной или акцепторной примесей имеем соответственно

$$N_d = n - n_0; \quad N_a = p + n_0. \quad (5.2)$$

2. Если нелегированный образец p -типа, то для тех же примесей имеем

$$N_d = n + p_0; \quad N_a = p - p_0. \quad (5.3)$$

Определение C_L сводится к расчету атомных долей примеси, введенной в объем V_L расплава навеской p_i примеси. Объем V_L определяется из измеренного объема твердой легированной части кристалла с учетом теплового расширения и изменения плотности вещества при плавлении (например, для Ge объемный коэффициент теплового расширения равен $18 \cdot 10^{-6}$, а уменьшение объема при плавлении составляет 5.5%; для Si — $12 \cdot 10^{-6}$ и 9% соответственно).

Если M_i — атомный вес примесного элемента, выраженный в граммах, а N_A — число Авогадро, то $N_A p_i / M_i$ — число примесных атомов в

заданной навеске. Тогда

$$\begin{aligned} C_L &= (N_A p_i)/(M_i V_L), \\ K &= C_S/C_L = (N_i M_i V_L)/(N_A p_i). \end{aligned} \quad (5.4)$$

Определенные таким способом коэффициенты разделения точны при следующих условиях: 1) $C_S = N_i$, то есть все атомы примеси электрически активны (обычно соблюдается в пределах растворимости примесей); 2) $K = K_0$, то есть не отличается от оцениваемых из равновесных диаграмм состояния сплавов и K не зависит от V — скорости движения фронта кристаллизации (обычно это имеет место при $V < 1$ мм/ч) и других условий кристаллизации.

Рассмотрим значения K_0 для различных примесей в Ge и Si. На рис. 5.4 эти значения сопоставлены с ковалентными радиусами примесных атомов. Видно, что в Ge лишь две примеси — Si и B — имеют $K_0 > 1$; $T_{пл}$ обоих элементов более высокие, чем у Ge. Оба элемента имеют ковалентные радиусы r_k меньше, чем у Ge, и это объясняет легкость их вхождения в решетку. У остальных примесей третьей, четвертой и пятой групп значения K_0 резко падают с ростом r_k . Общая черта подавляющего числа примесей этих групп — $K_0 < 1$ и, как правило, в этом случае $T_{пл}$ примесей меньше, чем $T_{пл}$ германия.

Для Si характерны те же закономерности, однако сами значения K_0 примесей, естественно, другие. Обращает внимание $K_0 \approx 1$ у примеси бора, имеющее далеко идущие последствия в технологии очистки Si.

Однако следует заметить, что $K_0 \ll 1$ в Ge и Si имеют и примеси, $T_{пл}$ которых выше, чем у основного вещества. Примером может служить Au ($T_{пл} = 1063^\circ\text{C}$) в Ge. Этот случай реализуется, как правило, когда диаграмма состояния системы полупроводник–примесь относится к диаграммам состояния с ограниченной растворимостью компонентов друг в друге. Рассмотрим диаграмму состояния системы Ge–Au, показанную на рис. 5.5. Из рисунка следует, что с ростом содержания Au в интервале 0–73% $T_{пл}$ сплава понижается до 365°C , то есть становится меньше, чем $T_{пл\text{Ge}}$, и лишь затем снова повышается. Линиями солидуса на диаграмме состояния вблизи Ge будут эвтектическая горизонталь (365°C) и линия вблизи чистого германия, которая не изображена на рис. 5.5 из-за малой растворимости Au в твердом Ge (в линейном масштабе она вырождается в вертикальную линию). Известно, что максимальная растворимость Au в твердом Ge не превышает $1.5 \cdot 10^{16} \text{ см}^{-3}$, то есть составляет $4 \cdot 10^{-5}$ ат.% при температуре $T = 921^\circ\text{C}$. Если сопоставить это значение с соответствующим значением концентрации Au в Ge на линии ликвидуса, то K_0 получится тот же, что показан на рис. 5.4, то есть $\approx 10^{-5}$.

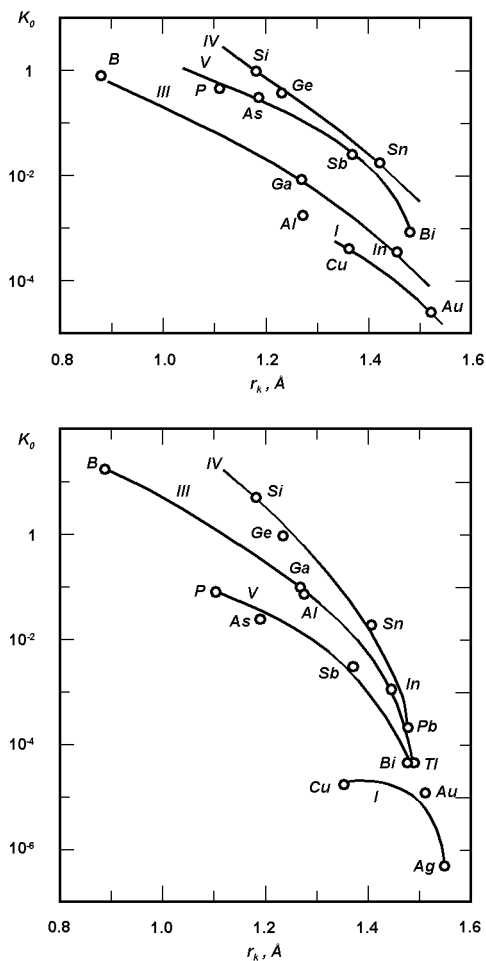


Рис. 5.4. Зависимости коэффициентов разделения примесей в точке плавления Si (сверху) и Ge (снизу) от тетраэдрических радиусов примесных атомов [32].

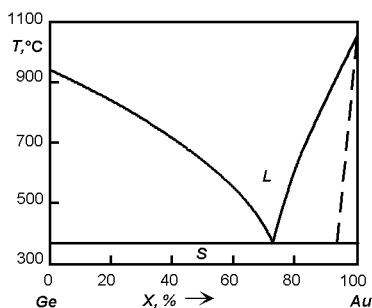


Рис. 5.5. Диаграмма состояния системы Ge–Au.

Таким образом, следует еще раз подчеркнуть, что *равновесный коэффициент разделения* K_0 определяется *видом фазовой диаграммы*, а в конечном счете — свойствами взаимодействующих атомов.

Обращает на себя внимание тот факт, что для всех переходных металлов (Fe, Co, Mn, Ni) в Ge и Si характерны очень малые $K_0 \approx 10^{-5} - 10^{-6}$, точность определения которых, естественно, не может быть высокой. Все эти примеси, включая Au и Ag, относятся к разряду «глубоких» и даже в малых концентрациях оказывают сильное влияние на времена жизни неравновесных носителей заряда. Отделим эту группу примесей от группы «мелких» водородоподобных примесей. В дальнейшем будем особо рассматривать условия их вхождения в решетку, их диффузию и растворимость.

Для полупроводниковых соединений характерен тот же интервал значений K_0 , что и для Ge и Si, но большее количество примесей имеет $K_0 > 1$ и корреляция коэффициентов разделения с ковалентными радиусами примесей более слабая. Очистка полупроводниковых соединений сопряжена с большими трудностями, чем очистка элементарных полупроводников. Поэтому обычно очистку соединений начинают с очистки входящих в соединение компонентов, а затем производят синтез предварительно очищенных компонентов. Суммарное содержание остаточных примесей в исходных материалах обычно составляет величину $\approx 10^{-4} - 10^{-5}\%$. Часто синтез совмещают с последующей дополнительной очисткой соединения путем направленной или зонной кристаллизации расплава (см. ниже). Однако, как правило, даже в чистом с химической точки зрения соединении остаются собственные дефекты, возникающие из-за отклонения состава от стехиометрического (неточности в соотношении числа атомов компонентов в соединении относительно формульного состава). К таким соединениям относятся, например, соединения

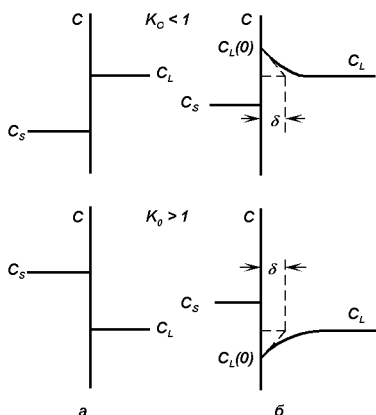


Рис. 5.6. Схематическое изображение распределения примесей на границе раздела твердой и жидкой фаз при различных условиях кристаллизации: а — в равновесных условиях; б — в неравновесных условиях.

$A^{II}B^{VI}$, $A^{IV}B^{VI}$. В этом случае на первый план выходят не методы очистки соединения, а методы, позволяющие приблизить состав кристалла к стехиометрическому (например, отжиг при фиксированном парциальном давлении более летучего компонента, см. гл. 4). Синтез разлагающихся соединений проводят либо в запаянных кварцевых ампулах при контролируемом давлении паров летучего компонента в рабочем объеме, либо под слоем жидкого флюса (например, особо чистый обезвоженный борный ангидрид). Синтез соединений, имеющих большое давление паров летучего компонента над расплавом, осуществляется в камерах высокого давления.

5.2.2. Эффективный коэффициент разделения примесей

До сих пор мы рассматривали равновесные коэффициенты разделения, рассчитываемые из фазовых диаграмм или получаемые методом, основанном на измерении концентрации свободных носителей в легированных полупроводниках. И в том, и в другом методах коэффициент разделения определялся в условиях, близких к равновесным. Эти условия, как правило, выполняются при достаточно малых скоростях кристаллизации расплава. В этом случае концентрация примеси C_L , растворенной в расплаве, одинакова во всем его объеме (рис. 5.6,а), а ее концентрация в закристаллизовавшейся части $C_S = K_0 C_L$.

В реальных условиях фронт кристаллизации движется со скоростью большей, чем скорость диффузии примесей в расплаве. Пусть для определенности $K_0 < 1$. Тогда по мере роста кристалла со скоростью V за счет того, что твердая фаза содержит меньше примеси, чем исходный расплав, из которого она образуется, концентрация примеси в расплаве у границы раздела возрастает. Накопление избытка примеси, выделяющейся у движущегося фронта кристаллизации, приводит к образованию перед ним диффузионного слоя δ , из которого примесь путем диффузии переходит в объем расплава (рис. 5.6,б). Если же равновесный коэффициент разделения примеси $K_0 > 1$, то вблизи поверхности роста ощущается недостаток примеси. В стационарном режиме роста кристалла должен существовать поток примесного компонента от границы роста в глубь расплава, если $K_0 < 1$, или к ней, если $K_0 > 1$. В этих условиях содержание примеси в закристаллизовавшейся части кристалла C_S определяется концентрацией примеси в расплаве у фронта кристаллизации $C_L(0)$. Так как $C_L(0)$, как правило, неизвестна, то в этих неравновесных условиях связь между концентрациями примеси в твердой C_S и жидкой C_L фазах определяется с помощью *эффективного коэффициента разделения* K :

$$K = C_S/C_L, \quad (5.5)$$

где C_L — средняя концентрация примеси в расплаве.

Найдем связь между эффективным K и равновесным K_0 коэффициентами разделения примеси. Для этого необходимо найти зависимость, описывающую распределение концентрации примеси в расплаве перед движущимся фронтом кристаллизации. Математически распределение примеси в системе кристалл–расплав описывается дифференциальным уравнением массообмена в движущейся среде:

$$dC/d\tau = -\varpi \cdot \text{grad } C + D\nabla^2 C, \quad (5.6)$$

где ϖ — скорость движения расплава, D — коэффициент диффузии, τ — время. Без ущерба для общности решение поставленной задачи можно получить, решив это уравнение для одного направления, перпендикулярного границе раздела кристалл–расплав. Пусть начало координат совпадает с границей раздела кристалл–расплав, а положительное направление соответствует движению в глубь расплава. В этом случае уравнение примет вид:

$$dC/d\tau = -\varpi_x dC(x)/dx + Dd^2C(x)/dx^2. \quad (5.7)$$

Пусть $D = \text{const}$ в диффузионном слое, а режим роста кристалла установившийся, то есть $dC/d\tau = 0$. Тогда уравнение (5.7) примет вид:

$$Dd^2C/dx^2 - \varpi_x dC/dx = 0. \quad (5.8)$$

Пусть в пределах слоя δ скорость движения расплава w_x равна скорости роста кристалла V , то есть $w_x = -V$; концентрация примеси в твердой фазе равна C_S , а в глубине жидкой фазы (при $x > \delta$) C_L — средней концентрации примеси в расплаве; $C_L(0)$ — концентрация примеси в расплаве на границе раздела фаз при $x = 0$. Граничные условия определяют из равенства потоков примеси, подводимого к границе раздела фаз за счет оттеснения ее из твердой фазы и отводимого от нее вследствие диффузии. Таким образом,

$$\begin{aligned} V[C_L(0) - C_S] + DdC/dx &= 0 & \text{при } x = 0, \\ C &= C_L & \text{при } x = \delta. \end{aligned} \quad (5.9)$$

Решение уравнения (5.8) с граничными условиями (5.9) при $x = 0$ имеет вид:

$$\frac{C_L(0) - C_S}{C_L - C_S} = \exp\left(\frac{V\delta}{D}\right). \quad (5.10)$$

Вводя равновесный $K_0 = C_S/C_L(0)$ и эффективный $K = C_S/C_L$ коэффициенты разделения и обозначая $V\delta/D = \Delta$, получим уравнение, связывающее их между собой:

$$K(V) = \frac{K_0}{K_0 + (1 - K_0) \exp(-\Delta)}. \quad (5.11)$$

Как видно, значение K зависит от скорости выращивания кристалла V , а также от значения δ и коэффициента диффузии примеси в расплаве D . Величина D для большинства систем полупроводник–примесь (в расплаве) лежит в диапазоне 10^{-5} – 10^{-4} см²/с, а значение δ может меняться от 1 до 10^{-3} см в зависимости от условий выращивания; V обычно выбирают ≈ 1 – 3 мм/ч.

Δ — безразмерная величина, называемая приведенной скоростью кристаллизации. На рис. 5.7 представлен график зависимости K от Δ для разных K_0 . Видно, что увеличение приведенной скорости кристаллизации Δ смещает K в сторону 1 независимо от значения равновесного коэффициента разделения K_0 . И, наоборот, уменьшение Δ приближает K к K_0 .

Ясно, что для повышения эффективности процесса очистки необходимо предельно приблизить значение K к K_0 , что может быть достигнуто уменьшением скорости кристаллизации. Того же эффекта можно достичь, проводя очистку кристаллизацией в условиях интенсивного перемешивания жидкой фазы. В этом случае отвод примеси от фронта кристаллизации (при $K_0 < 1$) в основном осуществляется не за счет диффузии в жидкой фазе, а за счет конвекции, что существенно уменьшает

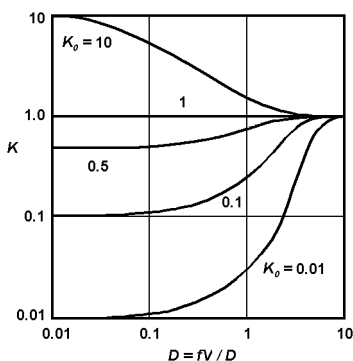


Рис. 5.7. Зависимости эффективного коэффициента разделения примесей K от приведенной скорости роста $\Delta = V\delta/D$ [37].

содержание примеси у фронта кристаллизации по сравнению с ее содержанием, получаемым без перемешивания, и уменьшает K , приближая его к K_0 .

5.3. Методы и принципиальные возможности очистки кристаллизацией

Рассмотрим сущность основных методов очистки кристаллизацией. Для того, чтобы обеспечить получение материала с предельной степенью чистоты кристаллизационными методами необходимо, чтобы кристаллизация начиналась в заданном месте и происходила в определенном направлении, то есть необходимо создать четкую границу между твердой и жидкой фазами и обеспечить ее медленное и равномерное движение вдоль очищаемого слитка. Эти условия достигаются заданием градиента температуры, обеспечивающего *направленный отвод тепла и направленное продвижение фронта кристаллизации*. При этом, в зависимости от значений K , примесь будет или «захватываться» твердой фазой ($K > 1$), освобождая от нее расплав, или оттесняться от границы раздела в расплав ($K < 1$). В первом случае части слитка, затвердевающие позже, будут чище предыдущих; во втором случае наиболее чистой оказывается начальная часть слитка. При рассмотрении процессов очистки кристаллов, выращиваемых из расплава, все многообразие методов выращивания кристаллов направленной кристаллизацией можно свести к двум идеализированным схемам: нормальной направленной кристаллизации и зонной плавке. В частности, первой схемой описывается процесс очистки кристаллов, выращиваемых широко используемым в промыш-

Рис. 5.8. Схема выращивания кристалла методом нормальной направленной кристаллизации.

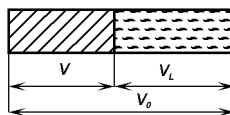
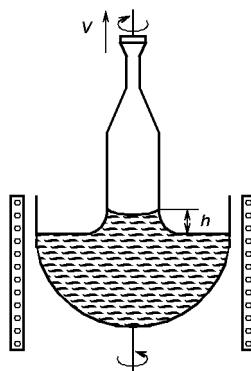


Рис. 5.9. Схема выращивания кристалла методом вытягивания из расплава.



ленной практике методом вытягивания кристаллов из расплава.

В методе *нормальной направленной кристаллизации* вещество расплавляют в тигле заданной формы, который затем медленно охлаждают с одного конца, осуществляя отсюда направленную кристаллизацию (рис. 5.8).

В методе *вытягивания кристаллов из расплава* в расплав опускают затравку в виде небольшого монокристалла, которую затем непрерывно перемещают вверх. Затравка увлекает за собой жидкий столбик, который, попадая в зону более низкой температуры, непрерывно кристаллизуется. Схема этого метода показана на рис. 5.9 (более подробно см. гл. 6).

При *зонной плавке* в слитке вещества расплавляют только небольшую зону, которую перемещают вдоль образца. По мере ее направленного продвижения впереди зоны происходит плавление вещества, а позади — его кристаллизация (рис. 5.10).

Степень чистоты вещества и распределение примеси вдоль кристалла в каждой из идеализированных схем различаются и являются характеристиками схемы. В настоящее время зонная плавка стала одним из наиболее распространенных кристаллизационных методов, применяемых для глубокой очистки многих материалов.

Используя основные законы фазовых превращений в многокомпонент-

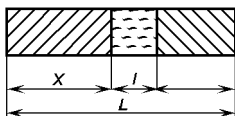


Рис. 5.10. Схема выращивания кристалла методом зонной плавки.

ных системах, выведем аналитические выражения для распределения концентрации остаточной примеси в твердой фазе в методах нормальной направленной кристаллизации и зонной плавки. Рассмотрим наиболее распространенный случай — бинарную систему (полупроводник — примесь) при небольшой концентрации примеси. Достаточно четкое и надежное математическое описание распределения примеси при направленной кристаллизации бинарной системы может быть получено только в некотором приближении.

Пусть:

1. Процессами диффузионного перераспределения компонентов в твердой фазе можно пренебречь, то есть коэффициент диффузии компонентов в твердой фазе

$$D_S = 0. \quad (5.12)$$

2. Перераспределение компонентов и соответственно выравнивание состава в жидкой фазе происходит мгновенно, то есть коэффициент диффузии в жидкой фазе

$$D_L = \infty. \quad (5.13)$$

Это условие обычно называют условием полного перемешивания жидкой фазы.

3. Значение эффективного коэффициента разделения постоянно, то есть

$$K = \text{const}. \quad (5.14)$$

Это допущение справедливо в области малых концентраций примеси, когда K не зависит от концентрации примеси. Кроме того, так как K зависит и от скорости роста кристалла (уравнение (5.11)), то условие $K = \text{const}$ также подразумевает постоянство скорости роста кристалла на протяжении всего процесса:

$$V = \text{const}. \quad (5.15)$$

4. При плавлении и затвердевании объем кристаллизующего материала не изменяется, то есть плотности жидкой и твердой фаз равны:

$$\rho_L = \rho_S = \rho. \quad (5.16)$$

5. В системе нет летучих и диссоциирующих компонентов.

Рассмотрим в рамках сделанных предположений распределение примеси в кристалле, выращиваемом разными методами.

5.3.1. Распределение примесей при нормальной направленной кристаллизации

Этой схемой описывается распределение примесей в кристаллах, выращенных методом нормальной направленной кристаллизации и методом вытягивания из расплава.

Пусть в некоторый момент времени объем расплава составляет V_L (рис. 5.8), а C_L и C_S — концентрации примесного компонента в расплаве и кристалле соответственно. Для определенности пусть $K < 1$. Так как слиток вначале расплавлен весь, то его состав вследствие перемешивания усредняется и равен исходной концентрации примеси в жидкой фазе C_0 . Пусть при понижении температуры закристаллизовался объем расплава, равный dV_L . Тогда количество примесного компонента, перешедшего из расплава в кристалл, составит $C_S dV_L$. В жидкой фазе при этом останется избыточное по отношению к начальному количество примесного компонента $(C_L - C_S)dV_L$. Оно приведет к изменению концентрации примеси в расплаве:

$$dC_L(V_L) = -\frac{C_L(V_L) - C_S(V_L)}{V_L} dV_L. \quad (5.17)$$

Опуская аргумент у функций C_S и C_L и учитывая, что $K = C_S/C_L$, получаем

$$dC_L = -\frac{C_L - KC_L}{V_L} dV_L = -\frac{(1 - K)C_L}{V_L} dV_L. \quad (5.18)$$

Знак минус перед дробью означает, что при уменьшении объема расплава концентрация примесного элемента в нем увеличивается. Интегрирование проводится при начальных условиях:

$$C_S|_{V_L=V_0} = KC_L|_{V_L=V_0} = KC_0, \quad (5.19)$$

где V_0 — общий объем материала.

Разделим в выражении (5.18) переменные и проинтегрируем его:

$$\begin{aligned} -\int_{V_0}^{V_L} \frac{dV_L}{V_L} &= \frac{1}{1-K} \int_{C_0}^{C_S/K} \frac{dC_L}{C_L}, \\ C_S &= KC_0(V_L/V_0)^{(K-1)}. \end{aligned} \quad (5.20)$$

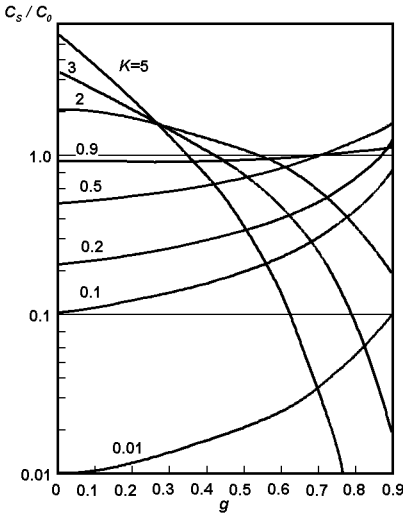


Рис. 5.11. Распределение относительной концентрации примеси с различными значениями K по длине кристаллов, выращенных нормальной направленной кристаллизацией (L — длина слитка) [37].

Переходя теперь от доли расплава V_L/V_0 , остающейся жидкой в момент кристаллизации слоя состава $C_S(V_L)$, к доле вещества g , закристаллизовавшейся к этому моменту,

$$g = 1 - V_L/V_0, \quad (5.21)$$

получим

$$C_S(g)/C_0 = K(1 - g)^{(K-1)}. \quad (5.22)$$

Эта зависимость для различных K представлена на рис. 5.11. Видно, что специфика очистки с помощью метода нормальной направленной кристаллизации заключается в перемещении примеси в тот или другой конец слитка. Максимальная степень очистки имеет место в начале ($K < 1$) или конце ($K > 1$) слитка. Заметим, что при $K < 1$ $C_S(g)/C_0$ по мере приближения g к 1 неограниченно возрастает, то есть как бы ни была мала концентрация примеси в расплаве C_0 в начале, концентрация ее в конечных частях кристалла согласно (5.22) должна быть безгранично большой, что противоречит физическому смыслу. Это противоречие является следствием нарушения допущения, что $K = \text{const}$, в конечных частях кристалла и поэтому выражение (5.22) не справедливо для этих частей кристалла.

Недостаток этого метода — его принципиальная однократность. Чтобы очистить материал более чем в K раз, необходимо удалить загрязнен-

ную часть слитка и только тогда повторить цикл снова. Но любые манипуляции с материалом являются источниками его поверхностного загрязнения. Для повышения эффективности очистки необходимо проведение многократной направленной кристаллизации без какого-либо взаимодействия материала с внешней средой. Эта цель достигается с помощью метода зонной плавки.

5.3.2. Распределение примесей при зонной плавке

При зонной плавке в начале слитка из очищаемого материала с помощью специального нагревателя (например, катушки высокочастотного нагрева) создается узкая расплавленная зона, которая затем движется вдоль слитка. Схема расположения зоны в произвольной части слитка показана на рис. 5.10. Расплавленная зона заключена между двумя твердыми частями кристалла: уже затвердевшей и еще не расплавленной. Проходя по слитку, она «вбирает» в себя примесь (при $K < 1$), перенося ее к концу. Принципиальным отличием зонной плавки от нормальной направленной кристаллизации является то, что доведенная до конца слитка зона может быть «перекинута» в его начало, где при повторном ее проходе она контактирует с уже предварительно очищенными участками, извлекая из них оставшиеся примеси и продолжая таким образом очистку материала.

Пусть при количественном рассмотрении распределения примеси в твердой фазе в случае зонной плавки остаются в силе все допущения, в которых получена формула (5.22), и пусть длина зоны l в течение всего времени проведения процесса остается постоянной: $l = \text{const}$.

Следует отметить, что поскольку при зонной плавке весь образец сразу не плавят, то его состав не усредняется. Поэтому распределение примеси в кристалле после зонной плавки в отличие от нормальной направленной кристаллизации должно зависеть от вида исходного распределения примеси в заготовке.

Рассмотрим распределение примеси в кристалле после первого прохода зоны через слиток с длиной L , постоянной площадью сечения S и первоначальным равномерным распределением примеси C_0 по нему. Пусть зона находится на расстоянии x от начала слитка (рис. 5.10). Положение зоны будем определять по границе ее затвердевания. Концентрация примеси в расплавленной зоне — $C_L(x)$. Рассмотрим баланс примеси в зоне в процессе ее движения.

Если зона сдвинется на расстояние dx , то из расплава у фронта кристаллизации в кристалл перейдет количество примеси, равное

$$dM'_L = KC_L(x)S dx. \quad (5.23)$$

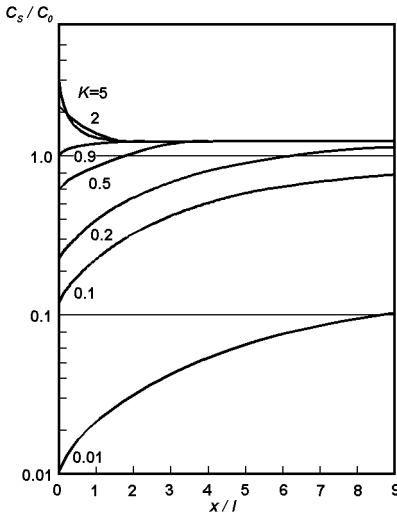


Рис. 5.12. Распределение относительной концентрации примеси с различными значениями K по длине кристаллов, выращенных зонной плавкой, после первого прохода зоны [37].

У другого края зоны, где произошло плавление слоя исходного слитка, в зону поступит количество примеси, равное

$$dM_L'' = C_0 S dx. \quad (5.24)$$

Тогда изменение концентрации примеси в расплавленной зоне при ее перемещении на dx будет

$$dC_L(x) = \frac{dM_L'' - dM_L'}{lS} = \frac{C_0 - KC_L(x)}{l} dx. \quad (5.25)$$

Концентрация примеси в начальном положении зоны

$$C_L(0) = \frac{1}{l} \int_0^l C_0(x) dx = C_0. \quad (5.26)$$

Интегрирование выражения (5.25) с учетом начальных условий (5.26) дает:

$$C_S(x) = C_0[1 - (1 - K) \exp(-Kx/l)]. \quad (5.27)$$

Эта зависимость в приведенных координатах для различных K представлена на рис. 5.12 (в приведенных координатах² вид кривой относительной концентрации $C_S(x)/C_0$ не зависит от длины зоны). Как видно

²В приведенных координатах расстояние от начала образца измеряется в единицах длины расплавленной зоны x/l .

из рис. 5.11 и 5.12, при $x = 0$ отношение $C_S(0)/C_0$ при зонной плавке будет принимать то же значение, что и при нормальной направленной кристаллизации при $V_L = 0$. Однако с ростом x , из-за малого объема зоны, нарастание $C_S(x)$ идет более быстро. При $x \gg l/K$ $C_S(x) = C_0$, то есть очистка прекращается. Это связано с накоплением примеси в зоне до концентрации $C_L(x) = C_0/K$.

Ясно, что при $K < 1$ накопление примеси в зоне будет происходить тем быстрее, чем меньше l . Поэтому длина зоны не должна быть слишком малой. Например, пусть $l = 0.1L$, тогда при $K < 1$ условие «насыщения» зоны примесью не выполняется в пределах всего слитка, и слиток будет эффективно очищаться до $x < 0.9L$, то есть пока правый конец зоны не дойдет до конца слитка. После этого распределение (5.27) перестает действовать, поскольку длина зоны начинает уменьшаться, и дальнейшая кристаллизация расплава (на последнем участке $0.9L < x < L$) происходит по схеме нормальной направленной кристаллизации. Распределение примеси в этом случае описывается зависимостью (5.22), где в качестве C_0 (начальные условия) должна быть взята величина $C_S(L - l)$, определяемая уравнением (5.27).

При $K = 10$ и $l = 0.1L$ условие «насыщения» зоны примесью выполняется практически с самого начала ее движения, поэтому эффекта очистки нет вплоть до $x < 0.9L$, когда начинает действовать распределение (5.22). Зонная плавка неприменима для очистки от примесей с $K > 1$. Однако с учетом распределения примеси в пределах последней зоны (5.22), которая при последующих проходах зоны будет сдвигаться влево, эффект очистки возможен, но он описывается уравнением (5.22).

Рассмотрим результаты последующих проходов зоны по образцу единичного сечения. При продвижении зоны на расстояние dx при n -ом проходе по образцу расплавляется часть слитка с концентрацией примеси $C_S^{(n-1)}(x + l)$, созданной в данном сечении предыдущим проходом зоны, и затвердевает часть зоны с концентрацией примеси $K C_L^{(n)}(x)$. Таким образом,

$$dC_L^{(n)}(x) = \frac{C_S^{(n-1)}(x + l) - K C_L^{(n)}(x)}{l} dx \quad (5.28)$$

или, переходя к концентрации примеси в твердой фазе и опуская для простоты индекс S , получаем

$$\frac{dC^{(n)}(x)}{dx} + \frac{K}{l} C^{(n)}(x) = \frac{K}{l} C^{(n-1)}(x + l). \quad (5.29)$$

Это уравнение применимо для всего слитка, исключая последнюю часть $(L - nl)$ (область последней зоны с каждым проходом расширяется

на длину зоны). Рассмотрим решение этого уравнения для участка от начала слитка до $(L - nl)$. Это тот участок, на котором на распределении примеси не сказывается изменение характера распределения примеси в последней зоне

Общее решение уравнения (5.29) имеет вид:

$$C^{(n)}(x) = \left[\frac{K}{l} \int_0^x C^{(n-1)}(x+l) e^{Kx/l} dx + \text{const} \right] e^{-Kx/l}. \quad (5.30)$$

Постоянная интегрирования определяется из условия:

$$C^{(n)}(0) = KC_L^{(n)}(0) = \frac{K}{l} \int_0^l C^{(n-1)}(x) dx. \quad (5.31)$$

Подстановка ее в предыдущее уравнение дает

$$C^{(n)}(x) = \frac{K}{l} e^{-Kx/l} \left[\int_0^x C^{(n-1)}(x+l) e^{Kx/l} dx + \int_0^l C^{(n-1)}(x) dx \right]. \quad (5.32)$$

Это выражение определяет концентрацию примеси в любом сечении слитка после n -го прохода зоны. Как видно из (5.32), фактором, определяющим $C^{(n)}(x)$, является не только $C^{(n-1)}(x)$, но и $C^{(n-1)}(x+l)$, причем $C^{(n-1)}(x+l)$ с каждым проходом все сильнее отличается от $C^{(n-1)}(x)$ (см. рис. 5.13). С другой стороны, все меньшую роль играет уход примеси в твердую фазу. Поэтому с ростом n отношение $C^{(n)}(x)/C^{(n-1)}(x)$ стремиться в пределе к 1, что означает прекращение очистки. Таким образом, в результате многократных проходов зоны по образцу распределение примеси в нем асимптотически стремится к некоторому предельному распределению, которое не может быть изменено последующим проходом зон.

Для полубесконечного образца предельное распределение примеси в нем может быть найдено следующим образом. Пусть конечное распределение примеси в образце будет $C_k(x)$. Тогда концентрация примеси в расплавленной зоне с учетом того, что при ее перемещении вдоль образца распределение примеси по длине образца не меняется, будет

$$C_L = \frac{1}{l} \int_x^{x+l} C_k(x) dx. \quad (5.33)$$

Так как концентрация примеси в твердой фазе в точке x равна $K C_L$, то

$$C_k(x) = \frac{K}{l} \int_x^{x+l} C_k(x) dx. \quad (5.34)$$

Решение этого уравнения будет:

$$C_k(x) = A \exp(Bx), \quad (5.35)$$

где

$$A = C_0 BL / (e^{BL} - 1), \quad (5.36)$$

а B является решением трансцендентного уравнения

$$K(e^{Bl} - 1) = Bl. \quad (5.37)$$

Например, пусть $l = 0.1, K = 0.1$. Тогда $B = 36$, а $C_k(0)/C_0 = 10^{-14}$. При $l = 0.1$ и $K = 0.01$ $B = 64$, а $C_k(0)/C_0 = 10^{-26}$.

Как видно, принципиальные возможности зонной плавки в очистке весьма велики. Однако формула (5.35) получена без учета эффекта «последней» зоны. Учет этого эффекта накладывает ограничения на длину зоны «сверху», так как он наступает тем раньше, чем больше l . При многократных проходах зоны, как уже отмечалось, последняя зона будет сдвигаться к началу слитка. Результирующие поправки к $C_n(x)$ с учетом этого эффекта рассчитаны в [38], а само предельное распределение примеси показано на рис. 5.13 вместе с распределениями примеси по длине образца после каждого из девяти предыдущих проходов зоны. Видно, что конечное отношение $C_{n=10}(0)/C_0$ на несколько порядков превышает рассчитанное без учета эффекта последней зоны. Таким образом, существуют два ограничения на величину размеров зоны l — сверху и снизу, — поэтому выбор l должен быть оптимальным. Как правило, выбирают $l \approx 0.1L$.

5.3.3. Модификация метода очистки зонной плавкой.

Контроль чистоты материала и оценка содержания в нем примесей

На степень очистки материала методом зонной плавки влияют такие параметры, как ее скорость, чистота контейнера, в котором она производится, и температура плавления очищаемого материала.

В реальных условиях фронт кристаллизации может двигаться со скоростью большей, чем скорость диффузии примеси в расплаве. В этих неравновесных условиях связь между концентрациями в твердой и жидкой фазах описывается с помощью эффективного коэффициента разделения K . Для повышения эффективности процесса очистки необходимо предельно приблизить значение K к K_0 . Это может быть достигнуто за

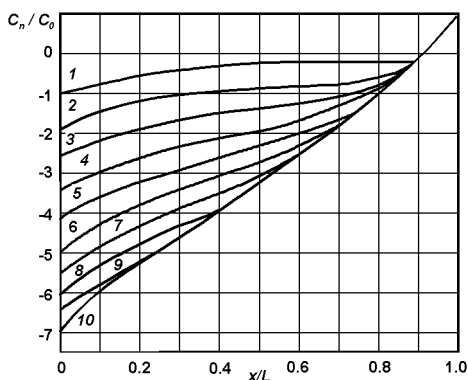


Рис. 5.13. Распределение примеси с $K = 0.1$ при $l = 0.1$ после n проходов зоны ($n = 10$) с учетом эффекта последней зоны [38].

счет уменьшения скорости кристаллизации (уменьшения скорости движения зоны). Как правило, для того чтобы в зоне действительно успевало установиться равновесное распределение отнесенной примеси, берут скорость движения зоны, равную 1–3 мм/ч.

Если учитывать, что для получения предельной очистки необходимо ≈ 10 проходов зоны (см. рис. 5.13), то результирующее время очистки оказывается довольно большим. В процессе всего этого времени расплав контактирует с горячим контейнером и окружающими элементами разогретой установки. Если контейнер недостаточно чист, то он сам служит непрерывным источником примесей, и, таким образом, именно эти примеси, а не примеси в исходном материале, лимитируют степень очистки. В промышленных условиях, где длина очищаемых слитков может превышать 0.5 м, для сокращения времени очистки зонную плавку проводят в многосекционном виде, то есть по длине слитка располагают несколько расплавленных зон одновременно. Однако и в этом случае время контакта очищаемого материала со стенками контейнера достаточно большое. Поэтому проблема контейнера ставит в некоторых случаях непреодолимые трудности перед «глубокой» очисткой материала.

Характерным примером может служить кремний, который в расплавленном состоянии активно реагирует с графитом и кварцем, широко используемыми в качестве материалов для контейнеров. Графит хорошо смачивается расплавленным кремнием, поэтому отделить после плавки материал от контейнера оказывается невозможно. Взаимодействие расплавленного кремния с кварцем приводит к загрязнению расплава при-

месяцами, имеющимися в кварце, в первую очередь кислородом, а также к растрескиванию слитка и кварцевой лодочки при остывании их после прохода зоны. Концентрация кислорода в таком «очищенном» кремнии превышает 10^{18} см^{-3} .

Для устранения этих явлений был разработан специальный вид зонной плавки — *вертикальная бестигельная зонная плавка*, — при которой расплавленная зона контактирует только с газовой фазой. Вертикальная бестигельная зонная плавка обеспечивает возможность металлургической очистки кремния и получения его с широким диапазоном сопротивлений и малым содержанием кислорода. Схема такой установки приведена на рис. 6.4,б.

Важнейшим отличительным признаком монокристаллов кремния, полученных бестигельной зонной плавкой в вакууме, является малое содержание в них кислорода ($\approx 10^{16} \text{ см}^{-3}$). Это связано с тем, что кислород быстро испаряется из расплавленного Si, поэтому для получения кристаллов с малым содержанием кислорода обычно достаточно одного прохода зоны в вакууме независимо от его содержания в исходном Si. При плавке в газовой атмосфере, меняя состав газовой фазы, можно получить Si с заданным содержанием кислорода. Поэтому именно бестигельный метод используется для получения особо чистого, «бескислородного» кремния, необходимого для изготовления работающих при комнатной температуре фотоприемников для ближней ИК- и видимой области спектра. Если кремний удастся очистить до остаточной концентрации носителей заряда $n(p) = |N_d - N_a| < 10^{12} \text{ см}^{-3}$, где N_d и N_a — концентрации доноров и акцепторов в кремнии соответственно ($\rho > 10^3 \text{ Ом} \cdot \text{см}$), и он характеризуется достаточно высокими временами жизни неравновесных носителей заряда $\tau_{n(p)} > 100 \text{ мкс}$, то такой материал имеет марку «детекторный», и из него изготавливают чувствительные фотоприемники, работающие без специального охлаждения. Однако их стабильная по времени работа (фоточувствительность) зависит от такого параметра материала, как степень компенсации примесей N_d/N_a . Чем больше N_d/N_a отличается от 1, тем стабильнее фоточувствительность. Поэтому целью очистки является не только повышение ρ , но и получение слабо компенсированного материала.

Для того, чтобы получить слабо компенсированный материал, годный для изготовления фотодетекторов, надо определить необходимое для этой цели число проходов зоны по кремнию. При этом для оценки степени чистоты материала нужно, во-первых, найти зависимость удельного сопротивления ρ от числа проходов зоны в Si и, во-вторых, определить степень компенсации примесей в уже очищенном материале. Обычно

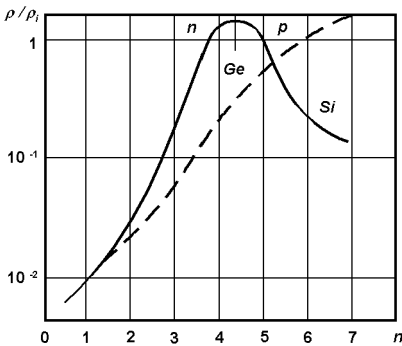


Рис. 5.14. Зависимость отношения удельного сопротивления ρ к собственному ρ_i от числа проходов зоны n в кремнии (сплошная линия) и в германии (пунктирная линия).

при очистке кремния методом вертикальной бестигельной зонной плавки пользуются предварительно очищенным от бора кремнием (например, химическими методами). Это связано с тем, что коэффициент разделения бора в кремнии близок к 1, и избавиться от него металлургическими методами чрезвычайно сложно. Однако было обнаружено, что при промежуточном контроле за степенью очистки материала, требующем нанесения и затем снятия контактов (после каждого прохода зоны), в кристалл при последующем зонном проходе вносится бор в концентрации $2.5 \cdot 10^{12} \text{ см}^{-3}$. Возникает проблема с определением зависимости $\rho(n)$. Для того, чтобы избежать загрязнения кремния бором, но иметь возможность определения ρ после каждого прохода зоны, необходимо вести зонную плавку, непрерывно смещая при каждом последующем проходе начальное положение зоны на контролируемое расстояние $\delta x_n \ll l$ вдоль слитка. Тогда на соответствующих участках после окончания полного цикла очистки можно будет измерить удельное сопротивление, не боясь внести бор в объем материала. Зависимость удельного сопротивления кремния от числа проходов зоны n показана на рис. 5.14. Для сравнения приведена и соответствующая зависимость для германия.

Анализ содержания остаточных примесей в предельно очищенном Si показывает, что основной акцепторной примесью является бор, а неосновной (донорной) — фосфор. При этом в процессе очистки N_a остается примерно неизменным при всех n (так как $K_B \approx 1$), а N_d уменьшается с ростом n . Таким образом (в отличие от Ge), кристалл с удельным сопротивлением ρ , близким к удельному сопротивлению собственного кремния ρ_i , отвечает не наиболее чистому, а компенсированному материалу. Наиболее чистым по суммарной концентрации примесей ($N_d + N_a$) является материал p -типа после 7 проходов, хотя $|N_d - N_a|$ в нем на

порядок выше, чем в том же материале после 4–5 проходов.

После 7–10 проходов зоны Si очищается от всех примесей кроме бора, который определяет величину конечного сопротивления Si при полной очистке или степень компенсации при легировании примесями n -типа. Для увеличения чистоты Si по бору или получения некомпенсированного Si дырочного типа с высоким сопротивлением применяется специальный метод зонной плавки: так называемая плавка в атмосфере влажного водорода. В основу этого метода положено удаление бора из расплава за счет реакции с газовой фазой. Используются обычные установки зонной плавки, но в них над расплавом пропускается водород, насыщенный парами воды. Степень очистки растет с уменьшением скорости перемещения зоны и с увеличением числа проходов зоны. Наряду с летучими соединениями бора происходит улетучивание и части кремния, который осаждается в виде окислов вместе с окислами бора на стенках установки.

Наиболее эффективным методом контроля количественного содержания донорных и акцепторных примесей, степени их компенсации и общего качества материала является исследование температурных зависимостей концентрации основных носителей заряда (постоянной Холла $R_H(T)$) и удельного сопротивления $\rho(T)$ в широкой области температур, включая область частичной и полной ионизации примесей. Параллельные измерения ρ и R_H обеспечивают возможность определения температурной зависимости холловской подвижности носителей заряда $\mu_x = R_H/\rho$, которая характеризует степень компенсации и однородность материала.

Рассмотрим возможности этого метода контроля чистоты материала на примере кремния (для определенности пусть у нас будет Si p -типа). Как известно [6], температурная зависимость концентрации носителей заряда в области полной ионизации примесей определяется соотношением $p(n) = |N_d - N_a|$. В области низких температур, когда справедливо неравенство $p(T) \ll N_d, N_a, |N_d - N_a|$, зависимость $p(T)$ определяется уравнением:

$$p(T) = \frac{(N_a - N_d)N_v}{gN_d} \exp\left(-\frac{E_{av}}{kT}\right), \quad (5.38)$$

где g — фактор вырождения, N_v — эффективная плотность состояний в валентной зоне, E_{av} — энергия ионизации доминирующей (в нашем случае акцепторной: бор) примеси. Соответствующие кривые $p(T)$ после 5 и 7 проходов зоны показаны на рис. 5.15,а; температурный ход подвижности — на рис. 5.15,б. Значительно более короткие кривые $p(T)$ и $\mu_x(T)$ после 5 прохода не случайны, так как они относятся к сильно компенсированному материалу, измерения параметров которого, из-за сильной

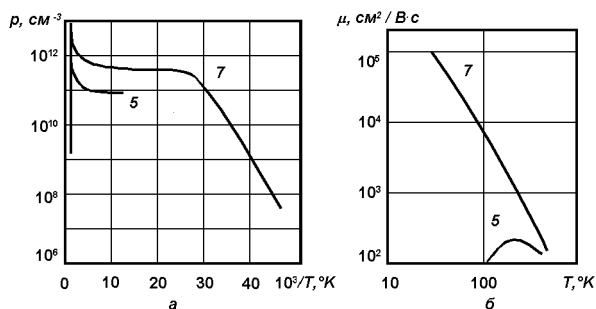


Рис. 5.15. Температурные зависимости концентраций (а) и подвижностей (б) дырок после пяти и семи проходов зоны.

неоднородности и быстрого спада μ_x с понижением температуры, практически невозможны.

Наличие экспоненты на кривой $p(1/T)$ с определенной энергией ионизации (для бора в кремнии $E_{av} \approx 0.05$ эВ) позволяет определить величину $(N_a - N_d)/N_d$, в то время как величина $(N_a - N_d)$ легко находится из высокотемпературного плато $p(T)$. Все вместе это позволяет определить концентрации «основной» и компенсирующей примесей отдельно. Кривая 7 на рис. 5.15 соответствует $(N_a - N_d) \approx 5 \cdot 10^{11} \text{ см}^{-3}$, и расчет дает, что $N_d/N_a = 0.2$. В сочетании с величиной максимальной подвижности дырок $\mu = 10^5 \text{ см}^2/\text{В} \cdot \text{с}$, этот результат очистки кремния можно считать хорошим.

Глава 6

Выращивание объемных полупроводниковых монокристаллов

Раздел 1. Получение кристаллов из жидкой фазы

Все технологические методы выращивания монокристаллов из жидкой фазы можно разделить на две группы: выращивание из собственных расплавов и выращивание из растворов.

Расплав — это жидкая фаза, состав которой соответствует составу кристаллизующегося нелегированного вещества или соединения.

Раствор — это жидкая фаза, состав которой отличается от состава выращиваемого нелегированного вещества или соединения.

Процесс кристаллизации из жидкой фазы состоит из следующих этапов: 1) подвода кристаллизующегося компонента к поверхности роста (фронту кристаллизации); 2) поверхностной диффузии (миграции по поверхности роста и встраивания атомов в кристалл); 3) диффузии в объеме кристалла (миграции в кристалле); 4) отвода скрытой теплоты кристаллизации от поверхности роста.

Одним из основных параметров, характеризующих процесс кристаллизации, является скорость роста кристаллов. Она зависит от способа и скорости подвода питающего компонента к поверхности роста, механизма роста кристаллической грани и интенсивности отвода тепла от фронта кристаллизации. При росте кристалла из газообразной фазы считают, что питание ступеней (изломов) в основном осуществляется по-

средством поверхностной диффузии атомов из адсорбированного слоя на растущей кристаллической грани. Поток вещества из газообразной фазы непосредственно к ступеням полностью пренебрегают (см. гл. 4). Существенной особенностью роста из жидкой фазы по сравнению с ростом из газообразной фазы является то, что подвод питающего материала к растущей грани кристалла обеспечивается в основном объемной диффузией из жидкой фазы, так как даже в условиях интенсивного перемешивания расплава (раствора) у поверхности растущей грани кристалла всегда имеется тонкий слой неподвижной жидкости — *диффузионный слой*. Скорость роста кристалла в таких условиях будет тем меньше, чем больше толщина диффузионного слоя. Однако не следует полностью пренебрегать и питанием ступеней (изломов) посредством поверхностной диффузии адсорбированных частиц на растущей грани.

Механизм роста кристаллической грани определяется главным образом ее строением, как и в случае роста кристаллов из газообразной фазы (см. гл. 4). Атомно-шероховатые (несингулярные) поверхности растут по нормальному механизму. В этом случае плотность центров роста сопоставима с плотностью поверхностных атомов и, как показывают расчеты для случая роста кристаллов из расплава, скорость роста поверхности пропорциональна переохлаждению на фронте кристаллизации ($V \sim \Delta T$). Атомно-гладкие (сингулярные и вицинальные) поверхности растут по слоистому механизму при двухмерном зарождении ступеней роста ($V \sim \Delta T \exp(-1/\Delta T)$) и по слоисто-спиральному механизму с участием винтовых дислокаций ($V \sim (\Delta T)^2$).¹ Анализ процессов роста кристаллов из раствора показывает, что в этом случае, так же как и в случае роста кристаллов из газообразной фазы (см. ниже), при малых пересыщениях зависимость скорости роста поверхности по слоисто-спиральному механизму от пересыщения близка к параболической, а при больших пересыщениях становится линейной.

Величина переохлаждения расплава или пересыщения раствора определяет образующуюся макроформу кристалла. При минимальном переохлаждении (пересыщении) образуются неограниченные монокристаллы (наиболее равновесная форма роста), при больших переохлаждениях (пересыщениях) — ограниченные монокристаллы, при еще больших переохлаждениях (пересыщениях) — дендриты (наименее равновесная форма роста). Внешняя форма неограниченных монокристаллов определяется симметрией теплового поля, в котором происходит рост кристалла (см. ни-

¹Приведенные зависимости скорости роста кристаллов из расплава от переохлаждения для конкретного механизма роста кристаллических граней могут быть получены на основании законов диффузии [17].

же).

Образующиеся центры кристаллизации разрастаются благодаря подводу кристаллизующегося вещества из жидкой фазы к фронту кристаллизации (процесс массопереноса). При этом выделяющаяся при кристаллизации теплота должна отводиться от поверхности раздела фаз, чтобы соблюдались температурные условия кристаллизации. Если внешний теплоотвод отсутствует, то теплота кристаллизации повышает температуру системы, и рост твердой фазы при достижении температуры плавления $T_{пл}$ прекращается. Для непрерывного роста твердой фазы необходимо наличие и постоянство температурного градиента, а скорость роста при этом будет определяться интенсивностью отвода тепла.

Таким образом, процессы тепло- и массопереноса наряду с кинетическими явлениями на поверхности роста играют определяющую роль в кинетике кристаллизации. Они могут быть описаны совокупностью дифференциальных уравнений тепло- и массопереноса. Строгое аналитическое решение такой системы уравнений в большинстве случаев крайне затруднено. Поэтому решение конкретной задачи часто проводят приближенно с привлечением представления о лимитирующей стадии процесса. *Лимитирующей стадией* процесса кристаллизации называется наиболее медленный ее этап, который и будет определять скорость роста монокристалла.

Процессы переноса играют важную, а зачастую и определяющую роль в формировании кристалла, в образовании дефектов в кристалле, в процессах легирования и т. д. Более детально процессы переноса будут рассмотрены ниже при анализе процессов выращивания из жидкой и газовой фаз.

6.1. Выращивание кристаллов из расплава

Выращивание кристаллов из расплава в настоящее время является наиболее распространенным промышленным процессом, так как по сравнению с другими методами методы выращивания из расплава обладают наивысшей производительностью. Это обусловлено тем, что в расплавах диффузионные процессы в жидкой фазе (диффузия к фронту кристаллизации компонентов кристаллизующейся фазы) не являются лимитирующей стадией процесса. С помощью этих методов можно получать достаточно чистые кристаллы Ge и Si с высокими скоростями роста (до 10 мм/ч), более чем в сто раз превышающими скорости роста кристаллов при выращивании другими методами. Применительно к многокомпонентным полупроводникам рост из расплава является сравнительно простым

процессом, если эти материалы плавятся конгруэнтно или если они имеют невысокие давления паров компонентов при температуре плавления. В противном случае методика выращивания и аппаратура процесса значительно усложняются.

Однако методы выращивания кристаллов из расплава обладают и определенными недостатками. Выращивание кристаллов из расплавов требует использования высоких температур роста, что в ряде случаев может создавать проблемы при контроле температурных градиентов, необходимых для выращивания кристаллов высокого структурного совершенства. Высокие температуры требуют также более высоких энергетических затрат при росте и способствуют загрязнению расплава, если он находится в тигле.

В основе всех методов выращивания монокристаллов из расплава лежит направленная кристаллизация расплава, при которой зарождение и рост кристалла при наличии переохлаждения ΔT в расплаве осуществляются на одной фазовой границе, а теплота от фронта кристаллизации отводится преимущественно в одном направлении. Это позволяет кристаллизовать расплав в виде одного монокристалла. Методы направленной кристаллизации подразделяются на три группы: методы *нормальной направленной кристаллизации*; методы *вытягивания из расплава*; методы *зонной плавки*.

6.1.1. Методы нормальной направленной кристаллизации

В методах нормальной направленной кристаллизации заготовка расплавляется целиком, а затем расплав кристаллизуется с одного конца. Общим для этих методов является рост кристалла в контакте со стенками тигля, содержащего расплав. Создание переохлаждения на фронте кристаллизации осуществляют путем перемещения тигля с расплавом относительно нагревателя, создающего тепловое поле с градиентом температуры, или путем перемещения нагревателя относительно тигля. Кристаллизация также возможна и при охлаждении неподвижного тигля с расплавом в тепловом поле с температурным градиентом (получаемым путем неравномерной намотки печи) при программном снижении питающей печь мощности. В зависимости от расположения тигля с материалом различают горизонтальный и вертикальный методы нормальной направленной кристаллизации. Вертикальный метод получил название метода Бриджмена (рис. 6.1,а).

Оборудование, необходимое для проведения процесса нормальной направленной кристаллизации, включает (рис. 6.1):

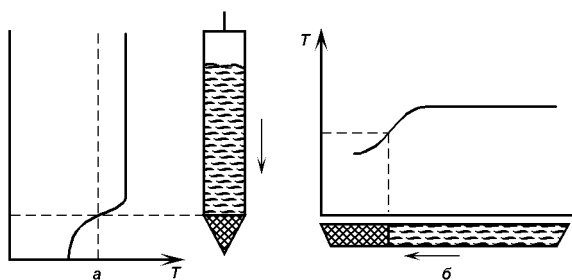


Рис. 6.1. Схема выращивания кристаллов методом нормальной направленной кристаллизации расплавов: а — вертикальная модификация (метод Бриджмена); б — горизонтальная модификация.

- 1) тигель заданной формы, изготовленный из материала, химически стойкого по отношению к расплаву и газообразной среде, в котором проводится процесс кристаллизации;
- 2) печь, обеспечивающую создание заданного теплового поля;
- 3) систему регулирования температуры печи и механического перемещения контейнера или нагревателя.

Суть метода нормальной направленной кристаллизации заключается в следующем. Предварительно тщательно очищенный исходный материал загружают в тигель и расплавляют; процесс проводят в вакууме или в нейтральной атмосфере в герметичной камере. Затем начинается охлаждение расплава, причем наиболее интенсивному охлаждению подвергается оттянутый заостренный участок тигля: здесь зарождаются центры кристаллизации (рис. 6.1). Заостренный конец используется с целью увеличения вероятности образования только одного центра кристаллизации, поскольку объем расплава, находящегося в заостренной части тигля, невелик. Кроме того, в случае образования нескольких центров кристаллизации один из них, имеющий наиболее благоприятную ориентацию для роста, подавляет рост остальных зародышей. С течением времени по мере перемещения тигля с расплавом относительно нагревателя или перемещения нагревателя относительно тигля фронт кристаллизации перемещается в сторону расплава и постепенно весь расплав в тигле закристаллизовывается. Слиток будет монокристаллическим, если будет расти только один зародыш.

Следует заметить, что, если используются методы нормальной направленной кристаллизации для выращивания кристаллов, то процессы зарождения и роста не контролируются с достаточной степенью точно-

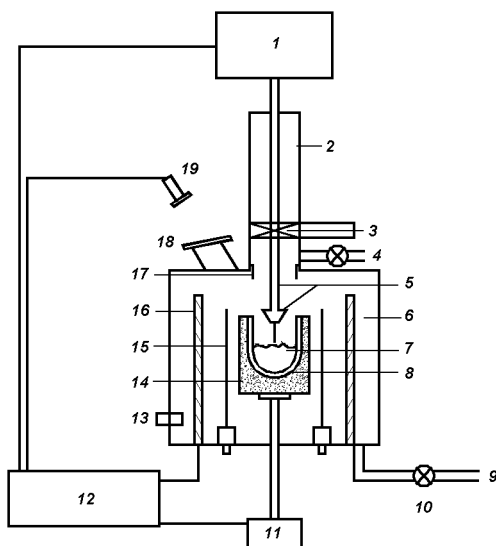
сти. Эти процессы зависят от формы фронта кристаллизации, от материала и качества изготовления тигля и всевозможных изменений условий роста. Особо следует остановиться на сильной зависимости совершенства выращиваемого кристалла от материала тигля. Для получения чистых кристаллов с минимальным количеством собственных дефектов необходимо выполнение следующих жестких требований, предъявляемых к свойствам материала тигля. Расплав и материал тигля не должны вступать в химическую реакцию. Расплав не должен смачивать стенки тигля, а после кристаллизации сцепляться с ним. Теплопроводность и тепловое расширение обоих материалов должны быть близки. Тигель должен обладать достаточной термической и механической прочностью. В качестве материала для изготовления тиглей наиболее часто применяют стекло, плавленый кварц, высокочистый графит, оксид алюминия (алунд), платину, нитрид алюминия и др.

Применение затравок при выращивании кристаллов методами нормальной направленной кристаллизации затруднительно, особенно в случае вертикального расположения тигля. Это связано с недостаточно высокой точностью измерения теплового поля печи (малые температурные градиенты) и часто с невозможностью визуального наблюдения за фронтом кристаллизации и затравкой из-за непрозрачности материала тигля и стенок рабочей камеры. В результате этого в начальный период процесса может происходить либо полное расплавление затравки, либо неполное расплавление исходного материала.

Процессы кристаллизации при горизонтальном расположении тигля отличаются от кристаллизации при вертикальном расположении тем, что, во-первых, механическое воздействие тигля на кристалл значительно меньше, и, во-вторых, тем, что фронт кристаллизации может быть доступен наблюдению, что позволяет использовать затравки. Однако управление формой фронта кристаллизации, особенно по вертикали, при горизонтальном расположении тигля значительно труднее, чем в случае вертикального расположения, ввиду асимметрии теплоотвода.

Выращивание кристаллов проводят в печах при различных градиентах температуры. При механическом перемещении тигля относительно нагревателя или наоборот желательно создание двух изотермических областей с перепадом температуры между ними: одна область с температурой выше, а другая ниже $T_{пл}$ (рис. 6.1). Практически ступенчатое распределение температуры достигается при использовании двухзонных печей с отдельным регулированием температуры и хорошей теплоизоляцией между зонами. Такое распределение температуры в печи позволяет одновременно с ростом проводить отжиг выращиваемого кристалла.

Рис. 6.2. Схема установки для выращивания кристаллов по методу Чохральского. 1 — затравочный шток, устройство подъема и вращения; 2 — верхний кожух; 3 — изолирующий клапан; 4 — газовый вход; 5 — держатель затравки и затравка; 6 — камера высокотемпературной зоны; 7 — расплав; 8 — тигель; 9 — выход; 10 — вакуумный насос; 11 — устройство вращения и подъема тигля; 12 — система контроля и источник энергии; 13 — датчик температуры; 14 — пьедестал; 15 — нагреватель; 16 — изоляция; 17 — труба для продувки; 18 — смотровое окно; 19 — датчик для контроля диаметра растущего слитка.



Метод нормальной направленной кристаллизации имеет важное значение для получения крупных монокристаллов с достаточно высокой скоростью выращивания (1–10 мм/ч). Таким способом растят химические соединения, например PbTe, GaAs. Однако в случае GaAs из-за высокой летучести As приходится проводить выращивание в специальных условиях: при избыточном давлении паров мышьяка. Основным недостатком метода является трудность получения кристаллов достаточно высокого структурного совершенства.

6.1.2. Методы вытягивания кристаллов из расплава

Данная группа методов в настоящее время является наиболее распространенной при промышленном производстве больших монокристаллов полупроводников с контролируруемыми и воспроизводимыми свойствами. Принцип метода вытягивания кристаллов из расплава впервые был предложен Чохральским в 1916 г. Сейчас существует значительное количество различных его модификаций. Суть метода заключается в следующем.

Исходный материал (в виде порошка или кусков поликристаллов), прошедший стадию тщательной очистки, загружают в тигель и нагревают до расплавления. Процесс проводят в вакууме или в атмосфере

инертного газа. Затравочный кристалл размером в несколько миллиметров, установленный в охлаждаемый кристаллодержатель и ориентированный в нужном кристаллографическом направлении, погружают в расплав. После частичного подплавления затравки и достижения определенного температурного режима начинается вытягивание таким образом, чтобы кристаллизация расплава происходила от затравочного кристалла. Диаметр растущего кристалла регулируется подбором скорости вытягивания и нагревом расплава.

Установка для выращивания кристаллов методом Чохральского включает четыре основных узла (рис. 6.2): 1) печь, в которую входят тигель, механизм вращения, нагреватель, источник питания и камера; 2) механизм вытягивания кристалла, содержащий стержень с затравкой, механизм вращения затравки и устройство для зажима затравки; 3) устройство для управления составом атмосферы, состоящее из газовых источников, расходомеров, системы продувки и вакуумной системы; 4) блок управления, в который входят микропроцессор, датчики и устройства вывода.

Рассмотрим более подробно основные стадии технологического процесса выращивания монокристалла методом Чохральского.

Непосредственно перед началом выращивания кристалла проводят выдержку расплава при температуре, заметно превышающей температуру плавления. Такая выдержка необходима для очистки расплава от летучих примесей, которые, испаряясь из расплава, осаждаются на холодных частях камеры. На поверхности расплава не должно быть пленок или посторонних частиц, так как они могут приводить к образованию поликристаллов вместо требуемых монокристаллов. После очистки расплава его температура понижается до температуры, немного превышающей температуру плавления материала. Кроме того, система используемых нагревателя, экранов, подставки для тигля должна обеспечивать такое распределение температуры в расплаве, чтобы кристаллизация началась в точке погружения затравки. Этому случаю соответствует понижение температуры расплава от стенок и дна тигля к центральной части расплава. Температура в центральной части расплава должна немного превышать температуру плавления материала. Температура же стенок тигля в течение всего процесса должна быть выше $T_{пл}$ (во избежание паразитной кристаллизации на стенках тигля).

Затем производят прогрев затравки путем выдержки ее над расплавом при температуре возможно более близкой к температуре кристаллизации. Это необходимо для предотвращения термоудара в момент контакта более холодной затравки с поверхностью расплава, так как термоудар при-

водит к существенному увеличению плотности дислокаций в затравке, которые прорастают в выращиваемый кристалл, ухудшая его структурное совершенство.

Качеству затравки уделяют особое внимание, поскольку ее структурное совершенство во многом определяет совершенство выращиваемого кристалла. Затравки вырезают из монокристаллов, кристаллографически ориентированных относительно направления вытягивания определенным образом, с минимальной плотностью дислокаций. Поверхностные нарушения удаляют химическим травлением и полировкой. Сечение затравки (имеющее форму квадрата или треугольника) делают минимально возможным для того, чтобы число дислокаций, наследуемых выращиваемым монокристаллом и возникающих вследствие термоудара, было минимально. Однако сечение затравки не должно быть слишком малым, иначе произойдет разрыв между затравкой и выращиваемым монокристаллом при вытягивании. Затравка должна быть ориентирована таким образом, чтобы облегчить движение и выход на поверхность кристалла дислокаций, проросших из затравки, и обеспечить максимальную симметризацию формы растущего кристалла. Например, кристаллы со структурой типа алмаза на практике часто выращивают вдоль направления $\langle 111 \rangle$, перпендикулярного плоскости с теми же индексами, так как плоскости $\{111\}$ являются плоскостями скольжения, в которых легко перемещаются дислокации, образованные различными источниками. Стремление к выращиванию симметричных монокристаллов (цилиндры с постоянным по длине диаметром) имеет глубокую основу. Однородность формы означает однородность и симметрию теплового режима роста монокристаллов, постоянство скорости кристаллизации и, в конечном счете, однородность электрофизических параметров по длине и поперечному сечению кристалла (см. ниже). Как правило, для максимальной симметризации теплового режима системы тигель и растущий кристалл вращаются в противоположных направлениях, при этом обеспечиваются и наилучшие условия перемешивания расплава.

После прогрева затравки ее конец погружают в перегретый расплав и частично подплавляют с целью удаления поверхностных дефектов и загрязнений. При этом граница раздела расплав–затравка оказывается расположенной над поверхностью расплава. Высота расположения границы раздела зависит от степени перегрева расплава и условий теплоотвода от затравки.

В первом приближении можно считать фронт кристаллизации при вытягивании плоским, как показано на рис. 6.3. Тогда высоту цилиндрического столба расплава можно оценить, приравняв вес столба жидкости,

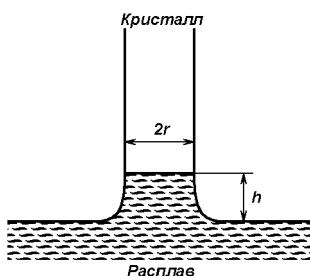


Рис. 6.3. Положение границы раздела кристалл–расплав при выращивании кристаллов методом Чохральского.

висящего на затравке, силам поверхностного натяжения, действующим по окружности фронта кристаллизации, и записать

$$\pi r^2 h \rho_l g = 2\pi r \sigma, \quad (6.1)$$

$$h = 2\sigma / r \rho_l g, \quad (6.2)$$

где σ — коэффициент поверхностного натяжения расплава, ρ_l — плотность расплава, g — ускорение силы тяжести, r — радиус столба расплава, h — высота столба расплава.

При сильном перегреве расплава возможен разрыв столба при вытягивании; при слишком низкой температуре расплава вокруг затравки образуется область переохлаждения, вследствие чего даже при отсутствии перемещения происходит заметное наращивание кристалла на затравку. Поэтому вытягивание кристалла необходимо начинать при промежуточной между двумя этими случаями температуре, то есть когда затравка сцеплена с расплавом, но роста кристалла еще не происходит. При вытягивании кристаллизация происходит у границы раздела.

На начальной стадии вытягивания после оплавления затравки производят формирование так называемой шейки монокристалла. При этом, как правило, диаметр шейки не превышает линейного размера поперечного сечения затравки, а длина составляет несколько ее диаметров. Формирование шейки производят понижая температуру расплава с большой линейной скоростью, что соответствует появлению больших осевых градиентов температуры. Это приводит к пересыщению вакансиями области монокристалла вблизи фронта кристаллизации, что при соответствующей кристаллографической ориентации затравки облегчает движение и выход на поверхность кристалла дислокаций, проросших из затравки.

После операции формирования шейки посредством варьирования температурных условий роста кристалла осуществляется разращивание монокристалла от размеров шейки до номинального диаметра слитка, то

есть «выход на диаметр». Угол разращивания делается достаточно малым для предотвращения возникновения термических напряжений и соответственно увеличения плотности дислокаций.

После выхода на диаметр температурные условия выращивания кристалла стабилизируют с целью получения монокристаллического слитка высокого структурного совершенства. Ведущая роль на данном этапе принадлежит тепловым условиям процесса, так как они определяют градиенты температуры в кристалле и расплаве, от которых, в свою очередь, зависят форма фронта кристаллизации, скорость роста кристалла, диаметр, структурное совершенство и, в конечном счете, электрофизические параметры выращиваемого кристалла.

Рассмотрим процесс теплопередачи на фронте кристаллизации.

Если температура расплава равна T , а температура кристаллизации — T_0 , то теплоподвод от расплава к фронту кристаллизации через столбик жидкости, пренебрегая его теплоизлучением, характеризуется величиной

$$Q_1 = \beta_l(T - T_0)\pi r^2/h, \quad (6.3)$$

где β_l — теплопроводность расплава.

При скорости вытягивания f на фронте кристаллизации будет выделяться количество тепла, равное

$$Q_2 = \pi r^2 L \rho_s f, \quad (6.4)$$

где L — скрытая теплота кристаллизации; ρ_s — плотность кристаллизующего материала.

Теплоотвод через твердый образец равен

$$Q_3 = \pi r^2 \beta_s dT_s/dx, \quad (6.5)$$

где β_s — теплопроводность твердого вещества вблизи температуры плавления; dT_s/dx — температурный градиент в кристалле. В этом соотношении неявно учтен и теплоотвод посредством излучения, так как последний учитывается в результирующем градиенте температуры вдоль растущего кристалла [22].

Условие теплового баланса на границе раздела будет выполнено, если тепловой поток Q_3 , отводимый от фронта кристаллизации растущего кристалла, равен сумме теплового потока Q_1 , поступающего к фронту кристаллизации из расплава, и количества теплоты Q_2 , выделяющегося на фронте кристаллизации вследствие самой кристаллизации: $Q_1 + Q_2 = Q_3$. Подстановка соответствующих значений Q и h приводит к выражению:

$$\beta_l(T - T_0)/h + L\rho_s f = \beta_s dT_s/dx, \quad (6.6)$$

откуда

$$r = \frac{(\beta_s dT_s/dx - L\rho_s f)2\sigma}{\rho_l g \beta_l (T - T_0)}, \quad (6.7)$$

Из этого выражения видно, что при заданном градиенте температур вдоль растущего кристалла и постоянной скорости вытягивания диаметр выращиваемого кристалла обратно пропорционален перегреву расплава.

Максимально возможная скорость вытягивания кристалла достигается тогда, когда поток тепла от расплава к границе раздела фаз полностью прекращен, то есть $Q_1 = 0$. Это условие реализуется при сведении к минимуму градиента температуры в расплаве, то есть при $T - T_0 = 0$. В этом случае от фронта кристаллизации отводится только скрытая теплота кристаллизации, а максимально допустимая скорость вытягивания кристалла определяется выражением

$$f_{max} = (\beta_s \text{grad} T_s)/(L\rho_s) \quad (6.8)$$

и может достигать десятков мм/мин.

Хотя в это выражение радиус кристалла явно не входит, можно показать, что он косвенно влияет на максимальную скорость вытягивания кристалла через результирующий градиент температур посредством учета теплоизлучения с поверхности растущего кристалла [38].

Существенное влияние на процессы теплопередачи оказывает характер атмосферы в ростовой камере. При выращивании кристаллов в вакууме теплопередача осуществляется только путем излучения. В газообразной же среде основную роль в процессе теплопередачи играют конвекционные процессы. В этом случае интенсивность теплопередачи возрастает с увеличением давления газа и его теплоемкости.

Любые изменения теплового баланса на границе раздела кристалл–расплав нарушают установившиеся условия роста и приводят к изменениям диаметра выращиваемого слитка, что, как правило, сопровождается возникновением различных структурных дефектов в кристалле. Обязательным условием для выращивания совершенных монокристаллов является высокая стабильность скорости вытягивания, скоростей вращения растущего кристалла и тигля с расплавом, постоянство мощности, подводимой к источнику нагрева тигля. Для стабилизации условий роста современные установки для выращивания кристаллов оснащены автоматизированными системами поддержания температуры нагревателя, непрерывного контроля диаметра выращиваемого слитка, подъема и вращения тигля и кристалла. Наиболее перспективными управляющими системами являются цифровые микропроцессорные системы. Они поз-

воляют уменьшить непосредственное участие оператора в процессе выращивания и дают возможность организовать программное управление многими этапами технологического процесса.

Процесс выращивания кристалла завершается отрывом его от расплава. Перед отрывом диаметр кристалла плавно уменьшают, создавая обратный конус, чтобы ослабить тепловой удар, приводящий к размножению дислокаций в конечной части слитка. Затем производят достаточно медленное охлаждение выращенного кристалла для предотвращения образования в материале термических напряжений и дислокаций. Для этого после отрыва монокристалл поднимают на небольшое расстояние над расплавом и производят медленное снижение температуры нагревателя.

Основными преимуществами методов вытягивания кристаллов из расплава по сравнению с методами нормальной направленной кристаллизации являются следующие.

1. Кристалл растет в свободном пространстве, не испытывая никаких механических воздействий со стороны тигля; при этом размеры растущего кристалла можно достаточно произвольно изменять в пределах, допускаемых конструкцией установки.

2. Имеется возможность визуального наблюдения за процессом роста вытягиваемого кристалла. Это позволяет сопоставлять свойства получаемых кристаллов с условиями их выращивания и производить оптимизацию технологического процесса.

3. Имеется возможность использования затравки минимального размера.

Масштабы применения метода вытягивания кристаллов из расплава можно проиллюстрировать на примере кремния. В настоящее время приборы на основе кремния составляют ~98% всех производимых в мире полупроводниковых приборов, а основную массу монокристаллического кремния (мировой выпуск которого превышает 2 тыс. тонн в год) выращивают методом вытягивания из расплава. Непрерывно совершенствуется технологическая аппаратура, используемая при производстве Si. В результате выращенные кристаллы достигают в диаметре 150 мм и по длине 1 м. Кристаллы указанных размеров выращиваются практически бездислокационными. Этот же метод применяют и для выращивания большинства полупроводниковых соединений, имеющих не слишком высокие давления паров своих компонентов. Например, таким способом выращивают кристаллы InSb.

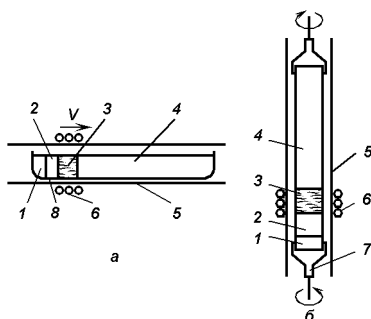


Рис. 6.4. Схема выращивания монокристаллов методом зонной плавки: *а* — горизонтальная зонная плавка; *б* — вертикальная бестигельная зонная плавка (1 — затравка; 2 — выращиваемый кристалл; 3 — расплавленная зона; 4 — исходный материал; 5 — стенки герметичной камеры; 6 — индуктор; 7 — кристаллодержатель; 8 — тигель).

6.1.3. Выращивание кристаллов методами зонной плавки

Зонная плавка может производиться в тигле и без тигля. В «тигельном» варианте, так же как и в случае нормальной направленной кристаллизации, предварительно очищенный материал загружается в тигель, в одном конце которого располагается монокристаллическая затравка. Тигель помещается в кварцевую трубку, которая, как правило, заполняется инертным газом. Узкая жидкая зона создается с помощью специального нагревателя, который, перемещаясь вдоль кварцевой трубки, обеспечивает передвижение зоны расплава вдоль обрабатываемого слитка и, следовательно, рост кристалла (рис. 6.4, *а*).

С точки зрения процесса кристаллизации метод «тигельной» зонной плавки мало чем отличается от метода нормальной направленной кристаллизации со всеми его недостатками (наличие стенок тигля и свободной поверхности, что не позволяет обеспечить полную симметризацию теплового режима). Однако есть и преимущества. Преимуществом метода «тигельной» зонной плавки по сравнению с методом нормальной направленной кристаллизации является то, что время, в течение которого расплав находится в контакте с материалом тигля, в этом случае меньше, а поэтому и загрязнение материала менее значительно. Кроме того, используя метод зонной плавки, можно регулировать ширину расплавленной зоны, создавать вдоль слитка несколько отдельных расплавленных зон, а также применять монокристаллическую затравку.

Схема установки бестигельной вертикальной зонной плавки приведена на рис. 6.4,б. В зажимах (цангах) укрепляется заготовка — цилиндрический или плоский (вначале) стержень перекристаллизуемого материала — и монокристаллическая затравка. Расплавление зоны, как и в горизонтальной плавке, осуществляется с помощью нагревателя. В зависимости от значения удельного электрического сопротивления исходного материала формирование расплавленной зоны осуществляется либо с помощью высокочастотного нагрева (индукционный нагрев), либо с помощью электронно-лучевого нагрева, либо сфокусированным излучением источника света. Такие способы нагрева не вносят загрязнений в обрабатываемый материал. Индукционный нагрев более предпочтителен, поскольку он обеспечивает эффективное перемешивание расплава и, следовательно, выравнивание его состава. Он наиболее часто используется при зонной плавке Si и некоторых других полупроводниковых материалов. Электронно-лучевой нагрев используется для тугоплавких неразлагающихся материалов, а радиационный — для обработки непроводящих и диссоциирующих материалов в атмосфере паров и газов. Специальные механизмы обеспечивают вращение верхней и нижней частей стержня относительно друг друга (с целью перемешивания расплава и симметризации теплового режима). Движение зоны вдоль образца осуществляется либо его перемещением относительно источника нагрева, либо перемещением нагревателя относительно образца. Расплав в пределах зоны удерживается силами поверхностного натяжения.

К числу важных преимуществ кристаллизации методом вертикальной зонной плавки относится возможность выращивания кристаллов без использования тиглей. В этом случае не происходит загрязнения расплава за счет растворения в нем материала тигля, а в выращиваемом кристалле не возникают дефекты вследствие различия коэффициентов линейного расширения кристалла и материала тигля. Метод вертикальной зонной плавки широко применяется при выращивании особо чистых монокристаллов полупроводников, а также материалов с высокой температурой плавления, обладающих в расплавленном состоянии высокой реакционной способностью, а также однородно легированных полупроводниковых материалов.

Принципы получения бездислокационных монокристаллов на затравке при зонной плавке такие же, как и при вытягивании монокристаллов из расплава, но при этом площадь сечения кристалла обычно имеет размеры, близкие к диаметру затравки.

Существенно повысить диаметр кристалла, выращиваемого бестигельной зонной плавкой, позволяет модификация этого метода, а именно:

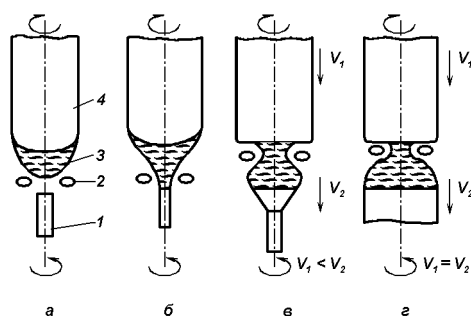


Рис. 6.5. Этапы процесса выращивания кристалла бестигельной зонной плавкой, когда диаметр индуктора меньше диаметра проплавляемого стержня (V_1 и V_2 — скорости движения стержня и затравки соответственно): *а* — создание капли расплава; *б* — смачивание затравки; *в* — выход на диаметр; *г* — выращивание кристалла постоянного диаметра.

использование индуктора (нагревателя), диаметр которого меньше диаметра проплавляемого стержня (рис. 6.5). Суть метода заключается в следующем. Заготовку 4, которая подвергается переплавке, закрепляют в верхнем держателе. Соосно с верхним держателем снизу располагают монокристаллическую затравку 1. На нижнем конце заготовки с помощью индуктора 2 создается капля расплава 3, которая удерживается на ней за счет сил поверхностного натяжения. К образовавшейся капле подводят затравку до соприкосновения с расплавом и полного ее смачивания. Далее индуктор перемещается относительно заготовки вверх, в результате чего заготовка над индуктором расплавляется, а под индуктором расплав кристаллизуется на затравку, образуя монокристалл. Верхний и нижний держатели вращаются в противоположных направлениях, что способствует перемешиванию расплава, симметризации теплового поля и выращиванию кристалла цилиндрической формы. Применение этого метода в технологии Si позволило увеличить диаметр выращиваемых монокристаллов до 100 мм. Существуют и другие разновидности метода зонной плавки: выращивание монокристаллов с пьедестала, метод Вернейля.

Основными недостатками бестигельной зонной плавки являются трудность управления величинами температурных градиентов вблизи фронта кристаллизации и трудность управления распределением температур вдоль слитка, что ведет к достаточно высокой концентрации собственных дефектов в выращиваемых кристаллах.

6.2. Выращивание монокристаллов из раствора

В технологии полупроводников наряду с кристаллизацией веществ из собственных расплавов широко используются процессы, основанные на кристаллизации из растворов. Основным преимуществом этого метода выращивания кристаллов является то, что процесс проводят при значительно более низких температурах, чем кристаллизация из расплавов. Выращивание из растворов зачастую является единственным методом, позволяющим получать объемные полупроводниковые монокристаллы с очень высокими температурами плавления ($T_{пл} > 2000^\circ\text{C}$, например, алмазы), кристаллы соединений, имеющие при температуре плавления высокое давление паров компонентов (GaP, GaAs, InP и др.), а также incongruently плавящиеся полупроводниковые соединения (GaTe₂ и др.). Основой такой возможности являются особенности фазовых диаграмм системы растворимое вещество–растворитель: зависимость температуры начала кристаллизации от процентного содержания компонентов в сплаве.

Кристаллизация из раствора часто способствует получению более чистого и совершенного по структуре полупроводникового монокристалла. Это связано с тем, что в этом случае процесс кристаллизации проводится при низких температурах, что ведет к меньшему загрязнению выращиваемого кристалла. Кроме того, происходит дополнительная очистка кристалла от примесей, если коэффициент разделения примеси в нем $K < 1$. Например, соединение GaP в процессе кристаллизации из раствора может очищаться от таких примесей, как Ag, Cu, Fe, коэффициент разделения которых в GaP меньше 1.

При кристаллизации из раствора растворитель должен удовлетворять следующим требованиям: существенно снижать температуру процесса выращивания кристалла; иметь достаточно малое давление собственных паров; не загрязнять выращиваемый кристалл, то есть иметь малый коэффициент разделения; желательно, чтобы атомы растворителя в кристалле (растворяемое вещество) являлись нейтральной примесью. Подобрать растворитель, удовлетворяющий всем сформулированным требованиям, очень трудно. Различают следующие случаи: 1) растворителем служат вещества, не входящие в состав выращиваемого кристалла (например, H₂O для NaCl); 2) растворителем служит один из компонентов выращиваемого кристалла (Ga для GaP).

В первом случае выращиваемые кристаллы содержат в качестве примесей все компоненты растворителя, включая остаточные примеси в нем, поэтому полученные кристаллы имеют невысокую степень чистоты. Во

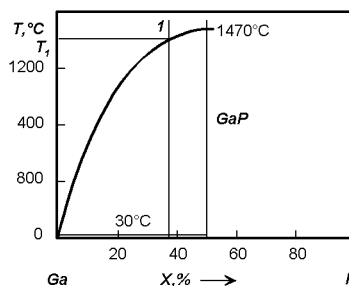
втором случае, поскольку в растворе отсутствуют посторонние вещества, чистота выращиваемых кристаллов определяется только чистотой компонентов, входящих в состав раствора. Однако возможность применения метода выращивания кристаллов из растворов для второго случая определяется типом диаграммы состояния выращиваемого соединения, поэтому сам метод оказывается менее универсальным, чем рост из посторонних растворителей.

При кристаллизации из раствора процесс роста кристалла складывается в общем случае из следующих стадий: 1) растворение исходных компонентов; 2) подвод питающих компонентов к поверхности роста за счет их диффузии через жидкую фазу к фронту кристаллизации; 3) осаждение на фронте кристаллизации и поверхностная диффузия; 4) диффузия в кристалле; 5) отвод теплоты кристаллизации от поверхности роста.

Если при выращивании монокристаллов из расплава лимитирующей стадией является отвод теплоты кристаллизации, то при выращивании из раствора самый медленный этап, как правило, — диффузия растворенного вещества к фронту кристаллизации. Вследствие этого линейная скорость роста кристаллов из раствора ($10^{-2} - 10^{-1}$ мм/ч) на 2–3 порядка меньше скорости роста из расплавов.

Монокристаллы полупроводниковых соединений АВ, изготовленные методом выращивания из раствора, обычно получают из раствора $A_{1-x}B_x$ с x , отличающимся от стехиометрического значения, характерного для соединения. В качестве примера рассмотрим кристаллизацию соединения GaP из раствора $Ga_{1-x}P_x$. Диаграмма состояния системы Ga–P представлена на рис. 6.6. Температура плавления GaP — 1470°C ; Ga и P при одной и той же температуре имеют резко различающиеся давления паров. Все составы с $x < 0.5$ будут плавиться при температурах ниже, чем само соединение. Таким образом, Ga можно использовать как растворитель, а из раствора $Ga_{1-x}P_x$ при температурах, соответствующих заданным значениям x , выращивать соединение GaP. Действительно, охлаждая раствор по линии 1 на рис. 6.6, мы пересечем линию ликвидуса при температуре T_1 , то есть при T_1 раствор окажется в пересыщенном состоянии и из него начнет выделяться твердая фаза. Первые кристаллики будут иметь состав, соответствующий пересечению коноды с линией солидуса, а состав жидкой фазы будет определяться по пересечению коноды с линией ликвидуса. При дальнейшем охлаждении (равновесном) составы жидкой и твердой фаз будут изменяться по линиям ликвидуса и солидуса соответственно. При этом жидкая фаза будет обогащаться Ga, а состав твердой фазы будет неизменным — GaP, так как линия солидуса вертикальная. Количество твердой и жидкой фазы при каждом значении

Рис. 6.6. Диаграмма состояния системы Ga–P.



температуры определяется правилом рычага.

Выращивание кристаллов из растворов можно производить как без специальных затравок, путем спонтанного образования и роста центров кристаллизации, так и контролируемым ростом на затравке. Практически выращивание крупных монокристаллов производят на затравках методами, аналогичными кристаллизации из собственных расплавов: методами нормальной направленной кристаллизации, методами вытягивания из раствора и методами зонной плавки. Однако технологическая аппаратура при выращивании кристаллов из растворов усложняется устройствами для обеспечения равномерной подачи исходных материалов в зону кристаллизации, то есть устройствами для обеспечения поддержания жидкой фазы в состоянии пересыщенного раствора.

Методы выращивания полупроводников из растворов принято различать по способу создания пересыщенного раствора, из которого будет выращиваться кристалл: 1) испарение растворителя; 2) подпитка раствора кристаллизующимся веществом (путем создания перепада температур между источником кристаллизующегося вещества и затравкой); 3) направленная кристаллизация насыщенного раствора; 4) зонное плавление при наличии температурного градиента, приложенного ко всему образцу. Первые три группы методов являются изотермическими, то есть фронт кристаллизации в течение всего процесса роста кристалла находится при постоянной температуре. Кристаллизация в изотермических условиях позволяет обеспечить лучший контроль параметров выращиваемого кристалла, зависящих от температуры процесса.

Наиболее широко используемым методом выращивания кристаллов из раствора в настоящее время является метод зонной плавки с градиентом температуры (более известный под названием «метод движущегося растворителя»), впервые предложенный в 1955 г. Пфанном. Сущность метода можно понять рассматривая диаграмму состояния системы

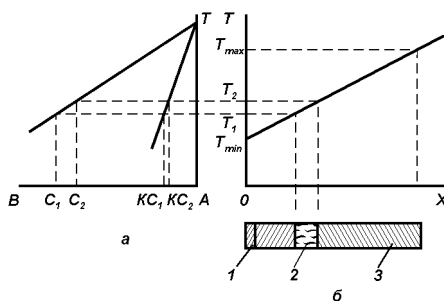


Рис. 6.7. Схема выращивания кристалла методом движущегося растворителя: *a* — участок диаграммы состояния системы А (кристаллизующий материал) — В (растворитель); *б* — распределение температуры по длине образца (1 — затравка материала А; 2 — зона жидкого раствора А–В; 3 — исходный материал А).

А–В (рис. 6.7). Здесь А — растворяемое вещество, В — растворитель, T_{max} — максимальная температура слитка, T_{min} — минимальная температура слитка, T_1 — минимальная температура зоны, T_2 — максимальная температура зоны.

Между затравкой и исходным материалом А помещают тонкий слой растворителя В, а вся система находится в таких условиях, что температура слоя В ниже температуры плавления вещества А, но выше температуры плавления вещества В. Находясь в соприкосновении с твердым веществом А, слой В после расплавления растворит в себе некоторое количество вещества А и займет больший объем. Растворение вещества А будет происходить до тех пор, пока границы зоны не достигнут температур T_1 и T_2 , соответствующих состоянию равновесия системы А–В. Действительно, по мере того, как у обеих поверхностей раздела фаз будет происходить растворение вещества А, концентрация В в промежуточном слое будет убывать до тех пор, пока у менее горячей поверхности зоны с температурой T_1 не достигнет концентрации ликвидуса C_1 . Тогда растворение вещества А здесь прекратится. У более горячей поверхности зоны растворение будет продолжаться до тех пор, пока при температуре T_2 не будет достигнута концентрация ликвидуса C_2 . В расплавленной зоне создается градиент концентрации вещества А и происходит его диффузионный перенос от более нагретой к более холодной границе зоны. Таким образом, у «холодной» границы зоны образуется пересыщенный раствор и происходит кристаллизация вещества А, содержащего примесь В с концентрацией KC_1 , а у «горячей» границы зоны образуется ненасыщенный

раствор и происходит дальнейшее растворение вещества А.

Непрерывно протекающие процессы растворения, диффузии и кристаллизации приводят к тому, что расплавленная зона под действием градиента температуры перемещается к более нагретому концу образца, то есть «ползет» вверх по линии ликвидуса. Перемещение расплавленной зоны в поле градиента температуры приводит к изменению в ней концентрации растворителя в соответствии с линией ликвидуса. Количество растворителя В уменьшается, кроме того, и за счет его растворения в твердой закристаллизовавшейся части А. От этих двух факторов и зависит длина зоны при перемещении ее вдоль слитка. Скорость перемещения расплавленной зоны тем выше, чем больше градиент температуры, коэффициент диффузии компонента А в растворителе В и чем меньше наклон линии ликвидуса.

Изменение концентрации растворителя в расплавленной зоне по мере ее перемещения вдоль слитка приводит к изменению состава кристаллизующегося вещества. Выращивание более однородных по составу кристаллов возможно при уменьшении градиента температуры и выборе растворителя В, обладающего малой растворимостью в твердой фазе А.

Метод движущегося растворителя успешно применяется при выращивании кристаллов InSb и $\text{Cd}_{1-x}\text{Hg}_x\text{Te}$. Кроме того, методы выращивания кристаллов из растворов широко используются в технологии полупроводниковых приборов для получения эпитаксиальных слоев (см. гл. 9).

6.3. Условия выращивания и дефектность монокристаллов

Важнейшей задачей технологии полупроводниковых материалов, как уже неоднократно отмечалось, является получение их в форме совершенных монокристаллов с определенной кристаллографической ориентацией и с минимальным или контролируемым и нужным образом распределенным количеством примесей. Поэтому так важно изучение влияния условий выращивания на дефектность получаемых монокристаллов.

Во время роста кристаллов могут образовываться равновесные дефекты и неравновесные дефекты (см. гл. 3).

Точечные дефекты (вакансии, междоузельные атомы, комплексы вакансий, антиструктурные дефекты) образуются в кристаллах в равновесных концентрациях в тех случаях, когда условия роста не сильно отличаются от равновесных. В то же время при выращивании кристаллов в неравновесных условиях кристалл может захватывать точечные дефекты в неравновесных концентрациях. Поэтому для получения при

комнатной температуре кристаллов с равновесной концентрацией точечных дефектов должны быть обеспечены условия аннигиляции неравновесных точечных дефектов при охлаждении кристаллов от температуры выращивания до комнатной.

Наиболее распространенным точечным дефектом в кристаллах являются вакансии. Уменьшение концентрации вакансий происходит либо путем заполнения вакансий диффундирующими к ним междоузельными атомами, которые переходят при этом в состояния с минимальной энергией; либо путем диффузии вакансий к поверхности кристалла и аннигиляции на ней; либо путем диффузии вакансий к дислокациям и аннигиляции на них. В первом приближении время, необходимое для уменьшения концентрации неравновесных вакансий в кристалле, можно оценить из соотношения $\tau \sim L^2/D_v$, где L — длина пути вакансии до места аннигиляции, а D_v — коэффициент диффузии вакансий. Соответственно время охлаждения выращенного монокристалла от температуры роста до комнатной должно быть не менее τ для получения полупроводника с равновесной концентрацией вакансий.

Помимо точечных дефектов, при кристаллизации могут возникать и такие неравновесные дефекты, как дислокации, малоугловые границы, двойники, включения второй фазы, поры и т.д. Энергия образования этих дефектов слишком велика, чтобы в кристаллах они могли образовываться в заметной равновесной концентрации (см. гл. 3).

Рост в виде монокристалла с минимальным количеством неравновесных дефектов будет происходить в неравновесных условиях только при выполнении определенных требований. Основные из них: обеспечение преимущественного роста одного центра новой фазы и обеспечение отвода выделяющейся на фронте кристаллизации скрытой теплоты кристаллизации только через твердую фазу в направлении нормальном фронту кристаллизации. Рассмотрим более детально влияние условий роста кристаллов на их дефектность.

При беззатравочном росте основным источником *дислокаций* может быть образование и рост нескольких центров новой фазы. Для уменьшения возможности их образования необходимо выполнение следующих условий.

1) Создавать выпуклый в расплав фронт кристаллизации с помощью дополнительных нагревателей (рис. 6.8). Это связано с тем, что на плоском или вогнутом в кристалл фронте кристаллизации случайные кристаллические зародыши будут расти совместно с основным, а на выпуклом в расплав фронте кристаллизации они будут оттесняться к периферии и вырождаться.

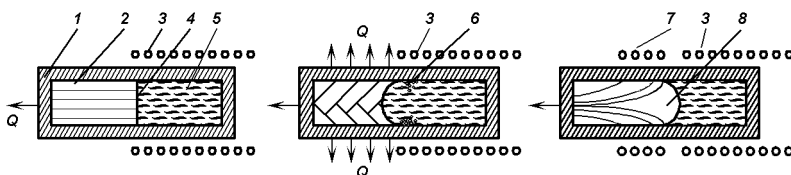


Рис. 6.8. Влияние формы фронта кристаллизации на структуру твердой фазы, образующейся при кристаллизации расплава без затравки: *а* — плоский; *б* — вогнутый в кристалл фронт; *в* — выпуклый в расплав фронт кристаллизации (*1* — контейнер; *2* — твердая фаза; *3* — нагреватель; *4* — фронт кристаллизации; *5* — расплав; *6* — случайные кристаллические зародыши; *7* — дополнительный нагреватель; *8* — благоприятно ориентированный кристалл; *Q* — тепло, выделяющееся на фронте кристаллизации).

2) Иметь в виду, что вероятность образования случайных зародышей зависит от градиента температуры в расплаве $W \sim \exp(-B/\Delta T)$ и уменьшается с уменьшением скорости кристаллизации, то есть, например, при слоисто-спиральном механизме роста растущей грани $V \sim (\Delta T)^2$, поэтому уменьшение скорости кристаллизации способствует уменьшению вероятности образования центров новой фазы.

3) Использовать тигли специальной формы, сделанные из несмачиваемого расплавом материала с коэффициентом термического расширения близким к коэффициенту термического расширения кристалла. Это позволит уменьшить вероятность образования случайных зародышей на стенках тигля.

Поскольку при беззатравочном росте, даже при соблюдении всех перечисленных выше требований, обычно получают крупноблочные поликристаллы, то в основном стараются выращивать монокристаллы с применением затравок. В этом последнем случае источниками дислокаций могут быть дислокации затравочного кристалла, термические напряжения в кристалле, примеси.

Влияние качества затравки на структурное совершенство выращиваемого кристалла и требования, предъявляемые к затравке, рассматривались при описании метода Чохральского (см. выше). Было выяснено, что этой проблеме необходимо уделять большое внимание.

Теперь рассмотрим вопрос о влиянии термических напряжений на появление дислокаций при росте кристалла. Известно, что термические напряжения могут вызвать пластическую деформацию в кристалле. Действительно, большие разности температур на малых расстояниях вызывают значительные механические напряжения благодаря различному

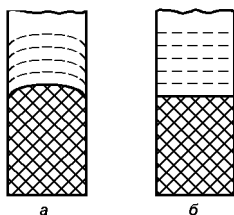


Рис. 6.9. Изотермические поверхности в растущем кристалле при вогнутом в кристалл (а) и плоском (б) фронте кристаллизации.

температурному сжатию отдельных участков кристалла. Если при этом напряжения превысят предел текучести кристалла, то возникнет его пластическая деформация с рождением большого числа дислокаций и других дефектов. Условия становятся особенно жесткими для материалов с низкой теплопроводностью, в которых большие перепады температур образуются особенно легко.

При росте монокристаллов из жидкой фазы термические напряжения вызываются наличием осевых (продольных) и радиальных (поперечных) температурных градиентов в кристаллах при их выращивании. Величины осевых температурных градиентов зависят от интенсивности охлаждения выращиваемого кристалла (главным образом, от интенсивности теплоотвода вдоль оси роста). На величины радиальных температурных градиентов влияет степень асимметрии теплового поля, в котором выращивают монокристалл, и соотношение величин теплооттоков вдоль оси растущего кристалла и с поверхности.

Градиенты температуры (в основном осевой), кроме того, зависят от скорости кристаллизации вследствие выделения скрытой теплоты кристаллизации. С увеличением скорости выращивания осевой градиент температуры уменьшается. Следует иметь в виду, что на процессы роста вытягиваемого из расплава кристалла значительное влияние оказывает также характер распределения температуры в расплаве. Большие градиенты температуры в расплаве приводят к конвективным потокам в нем, что, в свою очередь, приводит к хаотическим колебаниям температуры на фронте кристаллизации и скорости роста кристалла, а следовательно, к увеличению количества собственных дефектов в кристалле.

Величины радиального и осевого градиентов температуры определяют форму фронта кристаллизации, который может быть выпуклым в жидкую фазу, плоским или вогнутым. Наиболее благоприятным для выращивания монокристаллов с низкой плотностью дефектов является плоский фронт кристаллизации. Поскольку на практике обеспечить его трудно, то выращивание проводят, как правило, при слегка выпуклом в жидкую фа-

зу фронте кристаллизации. При наличии только продольных градиентов мы имеем плоскую форму фронта кристаллизации, лишь дополнительные поперечные разности температур вызывают его искривление. Например, рассмотрим рис. 6.9. Из него видно, что более холодная (из-за теплоизлучения с поверхности) поверхность кристалла при его выращивании действует подобно трубке, стремящейся сжать более теплую внутреннюю область. Плоской формы фронта кристаллизации можно достичь уменьшением потерь тепла с поверхности выращиваемого кристалла, направляя тепло однородным потоком вдоль его длины. Практически это осуществляется либо с помощью экранирования растущего кристалла, либо с помощью «посленагревателей» — печей, поддерживающих температуру растущего кристалла близкой к температуре плавления. Таким способом получают кристаллы Ge и Si с плотностью дислокаций $N_D < 100 \text{ см}^{-2}$.

При выращивании легированных монокристаллов к возникновению дислокаций могут приводить и примеси, неоднородно захватываемые растущим монокристаллом в концентрациях, близких к пределу растворимости или превышающих его (включения второй фазы). Однако примеси не всегда способствуют повышению плотности дислокаций во всем объеме растущего кристалла. При выращивании из расплава некоторых монокристаллов с алмазоподобной структурой, легированных простыми донорами, плотность дислокаций в них оказывается ниже по сравнению с нелегированными монокристаллами, выращенными в тех же условиях. Например, такая закономерность наблюдается в кристаллах Si, легированных отдельно фосфором, мышьяком, сурьмой; в кристаллах InSb, легированных отдельно селеном, теллуром. Однако пока механизм влияния легирующих примесей на плотность дислокаций в этих монокристаллах не выяснен до конца.

При значительных и заметно анизотропных термических напряжениях дислокации располагаются в плоскостях скольжения и группируются в *малоугловые границы*. Большое влияние на возникновение МУГ оказывает форма фронта кристаллизации, которая определяется осевыми и радиальными температурными градиентами при выращивании. При выпуклом в расплав фронте кристаллизации рост грани кристалла происходит чаще всего в результате разрастания одного двумерного зародыша, возникающего в наиболее холодной центральной части грани. При вогнутом в кристалл или плоском фронте кристаллизации рост грани может происходить от нескольких одновременно разрастающихся двумерных зародышей. Таким образом, источниками малоугловых границ являются все те же термические напряжения в кристалле и одновременный рост нескольких центров новой фазы. В монокристаллах со структурой типа

алмаза, выращенных в кристаллографических направлениях $\langle 111 \rangle$, термические напряжения снимаются путем скольжения дислокаций вдоль плоскостей $\{111\}$. Поэтому в монокристаллах, выращенных в данном направлении, МУГ встречаются реже, чем в кристаллах, выращенных в других кристаллографических направлениях.

Двойники и дефекты упаковки в монокристаллах также являются довольно распространенными дефектами структуры, возникающими при росте полупроводников. Основными причинами образования двойников в монокристаллах при их выращивании из жидкой фазы являются большие термические и механические напряжения на периферии фронта кристаллизации, а также включения второй фазы вблизи фронта кристаллизации. Двойники, причины появления которых перечислены выше, получили название двойников прорастания. Приведем примеры ситуаций, в которых появляются двойники прорастания. Значительные напряжения в выращиваемом кристалле, возникающие при кристаллизации полупроводников в тиглях методами Бриджмена и горизонтальной зонной плавки, особенно часто приводят к появлению двойников прорастания. Эти напряжения возникают в результате увеличения объема монокристалла при кристаллизации. Резкие изменения диаметра свободно растущего в методе Чохральского монокристалла также могут вызвать напряжения и появление двойников прорастания на конусной части кристалла. Попадание посторонних включений на поверхность растущего монокристалла в области фронта кристаллизации способствует появлению двойников.

Наряду с двойниками прорастания в монокристаллах, особенно при их выращивании из растворов, могут образовываться двойники срастания. Причиной образования двойников срастания является отклонение направления роста монокристалла от некоторого определяемого кристаллической структурой полупроводника направления. Например, рост монокристаллов с алмазоподобной структурой происходит преимущественно в результате развития наиболее плотноупакованных атомами плоскостей $\{111\}$. Выращивание таких монокристаллов в направлениях, отличных от направления $\langle 111 \rangle$, приводит к возникновению двойников срастания вследствие стремления одной из систем плоскостей $\{111\}$ в растущем монокристалле стать в положение, определяемое основным градиентом температуры. Поэтому наиболее склонным к двойникованию направлением роста будет направление $[100]$, в котором четыре плоскости (111) расположены под углом $36^\circ 16'$ к направлению роста кристалла. В этом случае плоскость двойникования оказывается параллельной одной из плоскостей $\{111\}$, а осью двойникования является одно из направлений $\langle 111 \rangle$.

Очевидно, что для получения монокристаллов с минимальным количеством неравновесных собственных дефектов необходимо исключить при их выращивании и охлаждении причины, приводящие к возникновению этих дефектов. Методы борьбы с дислокациями, малоугловыми границами, двойниками и дефектами упаковки, как видно из предыдущих рассуждений, — это использование тиглей специальной формы, сделанных из подходящего материала; создание слегка выпуклого в расплав фронта кристаллизации; максимальная симметризация теплового поля, в котором проводится рост кристалла; уменьшение градиентов температуры в растущем кристалле и уменьшение скорости кристаллизации; уменьшение градиента температуры в расплаве; использование совершенных затравочных кристаллов с минимально возможным сечением. Кроме того, для уменьшения вероятности возникновения МУГ и двойникования кристаллы необходимо выращивать в определенном кристаллографическом направлении.

Помимо перечисленных собственных дефектов при выращивании полупроводниковых материалов могут возникать и неоднородности состава кристаллов, обусловленные технологическими причинами. Рассмотрим основные факторы, приводящие к таким неоднородностям.

Одним из важнейших факторов, влияющих на локальную однородность состава кристалла, является неравномерное распределение потоков жидкости в расплаве, омывающем фронт кристаллизации. В результате этого толщина диффузионного слоя δ в расплаве у поверхности фронта кристаллизации в разных его точках различна, что приводит к неоднородному распределению примеси по сечению растущего кристалла, так как $K = K_0/[K_0 + (1 - K_0) \exp(-V\delta/D)]$. Эти процессы наиболее ярко проявляются в тех методах выращивания кристаллов, где объем расплава велик, например, в методе Чохральского. На практике для того, чтобы избежать подобного явления, как правило, производят одновременное вращение тигля и кристалла в противоположных направлениях (рис. 6.10). Вращение тигля необходимо для того, чтобы обеспечить осевую симметрию теплового поля в области расплава, так как практически трудно выставить на одну ось нагреватель, тигель и затравку кристалла. При одновременном вращении тигля и кристалла характер движения потоков в расплаве очень сложен и зависит от таких факторов, как направление вращения тигля и кристалла, число оборотов, конфигурация тигля, соотношение диаметров тигля и кристалла и т. д. Тем не менее, как показывает опыт, для повышения однородности кристалл необходимо вращать с максимально возможной скоростью (частотой), а тигель — с

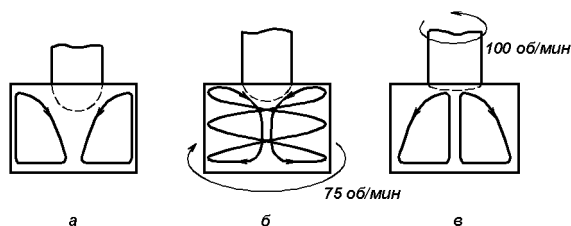
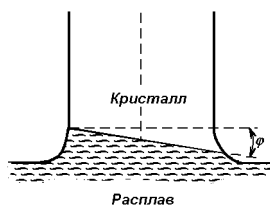


Рис. 6.10. Схема распределения гидродинамических потоков в расплаве и их влияние на форму диффузионного слоя (отмечено пунктиром) на фронте кристаллизации при выращивании кристаллов вытягиванием из расплава: а — в отсутствие вращения; б — при вращении тигля; в — при вращении кристалла.

минимальной. Помимо влияния на толщину диффузионного слоя δ тепловая конвекция в расплаве приводит к случайным колебаниям температуры расплава у фронта кристаллизации (об этом уже говорилось ранее), что ведет к хаотическим колебаниям скорости роста кристалла, и, следовательно, опять к неравномерному распределению примесей в растущем кристалле.

Искривление фронта кристаллизации также оказывает значительное влияние на распределение примеси в кристалле. Изменение формы фронта кристаллизации может вызвать колебания скорости роста кристалла V , что приведет к возникновению в слитке примесных полос, воспроизводящих в каждый момент роста форму фронта кристаллизации. В этом случае распределение примесей в продольном сечении кристалла представляет собой полосы, воспроизводящие последовательные положения фронта кристаллизации, а в поперечном сечении монокристалла — полосы в виде колец, спиралей или фигур кольцеобразной формы. На поверхности растущего кристалла слоистая неоднородность проявляется в виде рельефных углублений (типа винтовой нарезки у кристаллов, выращенных методами Чохральского или вертикальной зонной плавки, или пилообразных выступов у кристаллов, полученных методами горизонтальной зонной плавки или Бриджмена). Например, в методе Чохральского нарушение симметрии теплового поля вокруг растущего кристалла приводит к наклону фронта кристаллизации относительно поверхности расплава (рис. 6.11). В этом случае различные участки фронта кристаллизации вращающегося кристалла периодически проходят через области расплава в тигле с более высокой и более низкой температурами. В первом случае скорость роста замедляется (иногда части кристалла даже частично сплавляются), а во втором — ускоряется, то есть кристалл имеет флук-

Рис. 6.11. Отклонение фронта кристаллизации от горизонтального положения, обусловленное осевой асимметрией теплового поля.



туации скорости роста с периодом, равным периоду вращения. Следовательно, и микрообъемы с одинаковым содержанием примеси в кристалле располагаются внутри него по винтовой поверхности с шагом V/ω , где V — скорость роста кристалла, ω — скорость вращения кристалла. Внешним проявлением такого характера роста кристалла будет возникновение на его поверхности винтовой нарезки. Рассмотренный пример приводит к слоистой неоднородности в распределении примеси, имеющей правильный характер. Другие технологические неперіодические факторы могут искажать правильный характер этой периодичности. К ним относятся неперіодические нарушения механических и температурных условий роста кристалла. Нарушения механических условий роста, приводящие к искривлению фронта кристаллизации, следующие: плохая центровка кристалла относительно тигля с расплавом, нестабильность скоростей перемещения кристалла и вращения его и тигля, вибрация кристалла или расплава и др. К нарушениям температурных условий роста кристалла относятся колебания температуры нагревателей, колебания тепловых потерь от расплава и кристалла и др. Уменьшить этот вид неоднородностей можно спрямляя фронт кристаллизации путем устранения механических нарушений условий роста и путем тщательного подбора тепловых условий роста монокристалла.

При выращивании сильно легированных кристаллов к рассмотренным видам неоднородностей добавляется еще *канальная неоднородность*, связанная с кристаллографической ориентацией растущей поверхности, и неоднородность, связанная с *концентрационным переохлаждением* расплава.

Канальная неоднородность возникает вследствие существования зависимости эффективного коэффициента разделения K от кристаллографической ориентации растущей поверхности. В ряде случаев отношение K в разных кристаллографических направлениях может достигать 10–15. Так, например, в Si и Ge для некоторых примесей наибольший коэффициент разделения наблюдается при росте кристалла в направлениях $\langle 111 \rangle$.

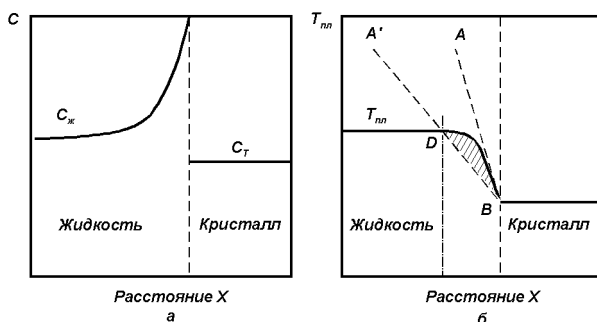


Рис. 6.12. *а* — Распределение примеси перед фронтом растущего кристалла, *б* — схема возникновения перед ним переохлажденной области.

Поэтому при выращивании кристалла в данном направлении при выпуклом или вогнутом фронте кристаллизации по оси выращенного слитка образуется «канал», имеющий форму трубки, в которой концентрация примеси в несколько раз выше, чем в остальной части кристалла. В частности, при выпуклом в расплав фронте кристаллизации анизотропия коэффициента разделения K примеси в зависимости от ориентации влечет за собой появление такого эффекта, как *эффект грани*. Суть этого эффекта состоит в следующем. Рассмотрим монокристалл со структурой типа алмаза, растущий в направлении $\langle 111 \rangle$. В центральной части фронта кристаллизации возникает хорошо развитая грань $\{111\}$, размеры которой зависят от величины теплового переохлаждения. Скорость роста этой грани, находящейся в более переохлажденной области, будет намного больше, чем других участков фронта кристаллизации, где переохлаждение существенно меньше. Различие условий роста на грани и вне ее способствует разному распределению примеси между жидкой и твердыми фазами в этих областях и образованию «канала» по оси растущего кристалла, в котором концентрация примеси оказывается выше, чем в других частях кристалла. Основные пути предотвращения канальной неоднородности — выращивание кристалла в кристаллографических направлениях, отличающихся от направлений, где коэффициент разделения максимален; спрямление фронта кристаллизации растущего кристалла, небольшие скорости роста кристалла, хорошие условия перемешивания расплава.

Образование примесных структур и включений второй фазы в растущем кристалле при *концентрационном переохлаждении* происходит следующим образом. Известно, что перед фронтом кристаллизации есть

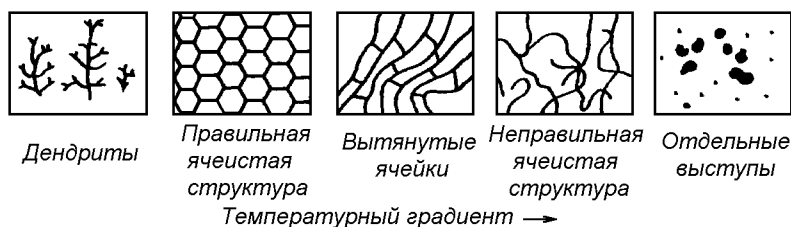


Рис. 6.13. Изменение морфологии поверхности кристалла, растущего в условиях концентрационного переохлаждения, в зависимости от градиента температуры у фронта кристаллизации.

область расплава, обогащенная примесью при $K < 1$. Так как примесь с $K < 1$ уменьшает температуру плавления $T_{\text{пл}}$ системы кристалл–примесь, то с приближением к поверхности растущего кристалла температура плавления будет уменьшаться. (рис. 6.12). На рисунке 6.12 пунктирными линиями AB и $A'B$ показаны два возможных случая распределения температуры вблизи фронта кристаллизации при выращивании кристалла. При большом градиенте температуры (прямая AB) концентрационного переохлаждения перед фронтом кристаллизации нет (температура в нем выше температуры ликвидуса). При небольшом градиенте температуры (прямая $A'B$) перед фронтом кристаллизации появляется слой расплава DB , обогащенный примесью и находящийся в переохлажденном состоянии, то есть температура в нем ниже температуры ликвидуса. Повышение переохлаждения перед фронтом роста приводит к возрастанию скорости роста кристалла. Поэтому любой случайно образовавшийся на фронте кристаллизации выступ, попадая в область концентрационного переохлаждения, будет быстро расти до тех пор, пока не достигнет точки, где температура расплава будет равна $T_{\text{пл}}$. Выступ, разрастаясь, будет оттеснять примесь в стороны, в результате чего на фронте кристаллизации может образоваться ячейка со скоплением примеси на ее границе. Совокупность таких ячеек образует ячеистый фронт кристаллизации. При очень высоких переохлаждениях ячеистый рост сменяется дендритным. Сравнительно невысокие концентрационные переохлаждения могут приводить к образованию в кристалле беспорядочно распределенных примесных включений второй фазы (рис. 6.13).

Избежать концентрационного переохлаждения можно путем создания в процессе выращивания кристалла больших градиентов температуры у фронта кристаллизации, а также путем интенсивного перемешивания расплава для разрушения диффузионного слоя перед фронтом кристал-

лизации.

Раздел 2. Выращивание кристаллов из газообразной фазы

Широко распространено мнение, что выращивание монокристаллов из газообразной фазы не имеет большого практического значения ввиду малых скоростей роста, присущих этому методу. Действительно, скорость роста монокристаллов из газообразной фазы обычно равна сотым долям мм/ч, что на несколько порядков ниже, чем при вытягивании кристаллов из расплава. Рост из газообразной фазы применяется в основном для выращивания тонких эпитаксиальных пленок, используемых в технологии полупроводниковых приборов, и для получения небольших монокристаллов тугоплавких материалов, а также полупроводниковых соединений, которые плавятся с разложением. Кроме того, поскольку высокопроизводительные методы выращивания монокристаллов из расплавов не всегда обеспечивают высокую однородность их свойств, то для получения особо качественных небольших кристаллов полупроводников используются методы выращивания из газообразной фазы. Эти методы, естественно, не устраняют все причины, приводящие к дефектности кристаллов. Процессы выращивания монокристаллов из газообразной фазы тоже весьма чувствительны к колебаниям внешних условий и составу питающей фазы. Однако влияние этих колебаний значительно сглажено благодаря малым скоростям роста, что способствуют приближению к более равновесным условиям роста.

Сам процесс роста кристаллов из газообразной фазы состоит из тех же этапов, что и процесс кристаллизации из жидкой фазы: 1) подвод кристаллизующегося компонента к поверхности роста (процессы массопереноса); 2) поверхностная диффузия (миграция по поверхности роста и встраивание атомов в кристалл); 3) диффузия в объеме кристалла (миграция в кристалле); 4) отвод теплоты кристаллизации от поверхности роста (процессы теплопереноса). Различие процессов конденсации из газообразной фазы и кристаллизации из жидкой фазы заключается в том, что при росте из газообразной фазы, как правило, не реализуется сплошность контакта этой фазы с фронтом кристаллизации.

С технологической точки зрения методы выращивания кристаллов из газообразной фазы делятся на три большие группы, отличающиеся способом доставки атомов от источника к растущему кристаллу:

- 1) метод сублимации-конденсации;
- 2) метод химических реакций разложения–восстановления;

3) метод химического транспорта.

Рассмотрим сначала особенности первой группы методов. В современной технологии полупроводников широкое распространение получили технологические процессы получения монокристаллов из паровой фазы в системах с пониженным давлением (в вакууме). К преимуществам выращивания кристаллов в вакууме относятся чистота условий получения материалов и изученность закономерностей массопереноса в вакууме, последнее позволяет научно обосновать и автоматизировать технологические процессы.

Массоперенос в вакууме включает следующие три основные стадии: 1) переход вещества из конденсированной (твердой) фазы в газообразную; 2) перенос вещества в пространстве от источника до подложки при пониженном давлении газа; 3) конденсация паров вещества на подложке.

Первая стадия — перевод вещества из конденсированной фазы в газообразную в основном осуществляется методами термического испарения или ионно-плазменного распыления.

Сущность процесса термического испарения заключается в нагреве вещества (методами резистивного, электронно-лучевого, высокочастотного, светового и др.) до температуры, когда энергия поверхностных атомов вещества становится выше их энергии связи с соседними атомами, в результате чего они приобретают способность к переходу в паровую фазу.

Ионное распыление (часто называемое катодным распылением) производят путем импульсной передачи энергии от ионов плазмы атомам вещества с последующим радиационным разрушением материала. Ион плазмы при соударении с атомами вещества передает им часть своей энергии. Если переданная энергия превышает пороговую энергию смещения атомов, то последние покидают свои места. При достаточном запасе энергии эти первичные смещенные атомы смещают другие атомы и т. д., в результате чего вдоль пути иона образуются каскады смещенных атомов. Часть из этих атомов достигает поверхности облучаемого кристалла и покидает ее.

В зависимости от соотношения между длиной свободного пробега атома или молекулы в паре l и расстоянием от источника паров до подложки L различают два основных режима второй стадии массопереноса вещества в вакууме.

При $l \ll L$ пар от поверхности испаряемого вещества движется в виде сплошного потока, и массоперенос в этом случае описывается уравнениями газодинамики [37]. Перенос вещества происходит путем молекуляр-

ной диффузии² или конвекции и молекулярной диффузии одновременно.³ В неподвижной среде вещество перемещается, строго говоря, лишь посредством одной молекулярной диффузии. В движущейся среде перенос вещества осуществляется как молекулярной диффузией, так и самой средой в направлении ее движения (ламинарный поток) или ее частицами в разных направлениях (турбулентный поток). В турбулентном потоке конвективный перенос вещества, осуществляемый под действием турбулентных пульсаций (нерегулярных пульсаций скорости), часто называют *турбулентной диффузией*.

При $l \gg L$ движение пара от поверхности испарения происходит в виде молекулярных пучков, в пределах которых атомы или молекулы распространяются по прямолинейным траекториям. Основные закономерности для этого режима массопереноса при термическом испарении были установлены Герцем, Кнудсеном и Ленгмюром. Плотность потока испаряющихся частиц (скорость испарения) в режиме молекулярных пучков описывается уравнением Герца-Кнудсена:

$$j_0 = (a_i P) / (2\pi RT / M)^{1/2}, \quad (6.9)$$

где j_0 — плотность потока испаряющихся частиц, P — равновесное давление пара испаряемого вещества при температуре T его поверхности, M — молекулярная масса вещества, R — универсальная газовая постоянная, a_i — коэффициент испарения, учитывающий факторы, препятствующие уходу атомов или молекул с поверхности конденсированной фазы в паровую.

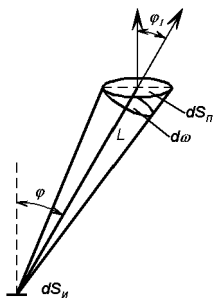
Если поверхность испарения и паровая фаза находятся в состоянии, близком к термодинамическому равновесию, то распределение по энергиям молекул, падающих на поверхность или покидающих ее, описывается распределением Максвелла. Из него непосредственно вытекает закон косинуса для пространственного распределения молекул, испускаемых нагретой поверхностью:

$$j(r, \varphi) = j_0 ds_i \cos \varphi / \pi L^2, \quad (6.10)$$

²Молекулярной диффузией называется перенос вещества, обусловленный беспорядочным движением самих молекул в неподвижном веществе. В однофазной системе при постоянной температуре и при отсутствии внешних сил диффузия выравнивает концентрацию каждого компонента фазы во всей системе. Если на систему действуют внешние силы или поддерживается градиент температуры, то в результате диффузии устанавливаются градиенты концентрации отдельных компонентов (электродиффузия, термодиффузия; см. 8).

³Суммарный перенос вещества молекулярной диффузией и движущимся потоком вещества называется *конвективной диффузией* или *конвективным переносом*. Очевидно, что скорость молекулярной диффузии существенно ниже скорости конвективного переноса вещества.

Рис. 6.14. Схема пространственного распределения молекул, испускаемых нагретой поверхностью, в зависимости от взаимного положения элементов поверхности испарения ds_{π} и подложки ds_{Π} .



где $j(r, \varphi)$ — плотность потока испаряемых молекул на расстоянии L от источника под углом φ к нормали элемента поверхности испарения ds_{π} (см. рис. 6.14).⁴

Количество вещества, которое осаждается на противоположной источнику паров растущей поверхности кристалла (подложке), также зависит от положения этой поверхности относительно испаряемой. Как видно из рисунка 6.14, количество вещества, испаряемого в пределах пространственного угла $d\omega$, осаждается на площади, которая возрастает с увеличением как расстояния до испаряемой поверхности L , так и угла падения φ_1 . Площадь элемента подложки ds_{Π} для данных $d\omega$, L и φ_1 такова:

$$ds_{\Pi} = L^2 d\omega / \cos \varphi_1. \quad (6.11)$$

Таким образом, поток осаждающегося вещества на единицу площади поверхности растущего кристалла (скорость осаждения) будет:

$$I = a_c I_0 \cos \varphi \cos \varphi_1 / \pi L^2, \quad (6.12)$$

где I_0 — поток испаряемых молекул, φ_1 — угол падения вещества на поверхность растущего кристалла, а a_c — коэффициент конденсации. Коэффициент конденсации a_c изменяется от 1 до 10^{-10} для разных веществ и условий. Было выяснено, что коэффициент конденсации зависит в первую очередь от механизма роста и чистоты растущей поверхности.

Для количественной характеристики эффективности ионного испарения используется коэффициент распыления R , равный числу атомов

⁴В соответствии с законом косинуса испарение вещества происходит неравномерно во всех направлениях. Оно идет преимущественно в направлениях, близких к нормали к испаряемой поверхности.

материала, выбитых одним падающим ионом [37],

$$R = K_{\alpha} \frac{M_1 M_2}{l(E)(M_1 + M_2)} E, \quad (6.13)$$

где K_{α} — коэффициент, учитывающий агрегатное состояние мишени; M_1 — масса бомбардирующего иона; M_2 — масса выбитого атома мишени; $l(E)$ — средняя длина свободного пробега иона в материале мишени; E — энергия иона. Значение R в сильной степени зависит от энергии иона, природы распыляемого материала, массы распыляющихся ионов, кристаллографической ориентации поверхности мишени и угла падения ионов. С практической точки зрения можно считать, что угловое распределение атомов, выбиваемых из поликристаллической мишени, близко к закону косинуса, а выбиваемых из монокристаллов, характеризуется наличием в кристалле наиболее предпочтительных направлений испускания атомов. Анизотропный характер эмиссии атомов из монокристаллической мишени при ионном испарении приводит к получению неоднородных по толщине пленок, а поэтому на практике применяется редко.

Рост кристалла из паровой фазы, в основном, происходит по слоистому или слоисто-спиральному механизмам с соответствующими зависимостями скорости роста V кристаллизующейся поверхности от пересыщения (см. гл. 4).

Для выращивания кристаллов из паровой фазы, так же как и для процесса выращивания из жидкой фазы, можно ввести понятие о лимитирующей стадии процесса. При выращивании из паровой фазы принято выделять диффузионную и кинетическую области роста. Если лимитирующей стадией процесса является подвод кристаллизующего материала к поверхности роста (осуществляемый посредством молекулярной или конвективной диффузии, а в вакууме — посредством молекулярного потока), то о таком процессе роста говорят как о идущем в «диффузионной области». Если же лимитирующей стадией являются процессы миграции и встраивания в решетку атомов на растущей поверхности, то процесс роста идет в «кинетической области». Если скорости подвода кристаллизующего материала и миграционных процессов соизмеримы, то говорят, что процесс протекает в «переходной области».

В случае выращивания кристаллов методами химических реакций и химического транспорта химические процессы, связанные с изменением химического состава и свойств исходного вещества, являются важнейшей стадией технологии получения полупроводниковых материалов наряду с тепло- и массопередачей. Скорость протекания химических про-

цессов определяется законами химической кинетики. В большинстве случаев при анализе химико-технологических процессов выделяют отдельные, наиболее важные стороны процесса и рассматривают их взаимное влияние. В первую очередь выделяют основную химическую реакцию, определяющую выход целевого продукта, и устанавливают особенности ее протекания.

Кратко рассмотрим основные закономерности протекания химических реакций с точки зрения химической технологии.

Химические реакции классифицируют по ряду признаков. Реакции подразделяются на *простые* и *сложные*. Для описания простых реакций достаточно одного уравнения. К сложным относятся реакции, состоящие из двух и более взаимосвязанных между собой (параллельных или последовательных) простых реакций. Для описания сложных реакций необходимо несколько уравнений. Разделяют *необратимые* и *обратимые* химические реакции. Необратимые реакции протекают лишь в одном направлении. Большинство реакций при соответствующих условиях обратимы, то есть могут протекать в прямом и обратном направлении. На производстве обычно стремятся создать такие условия, чтобы основные химические реакции протекали необратимо. Различают *гомогенные реакции*, когда реагирующие вещества находятся в одной какой-либо фазе, и *гетерогенные реакции*, когда реагирующие вещества находятся в разных фазах. Большинство химико-технологических процессов при получении полупроводниковых материалов протекает с участием гетерогенных химических реакций.

Разработка технологического процесса получения целевого продукта требует установления всей совокупности протекающих химических реакций в процессе переработки исходного сырья. Далее определяется вероятность протекания основной реакции, обеспечивающей получение заданного продукта, и рассчитывается максимальный выход целевого продукта (см. ниже) в технологическом процессе.

Вероятность протекания той или иной обратимой реакции определяется по изменению полной энергии Гиббса ΔG , которое рассчитывается с помощью стандартных таблиц или определяется на основе экспериментальных данных. Если $\Delta G > 0$, то наиболее вероятно протекание реакции слева направо; если $\Delta G < 0$, то наиболее вероятно протекание реакции справа налево; если $\Delta G = 0$, то существует состояние равновесия. На основании данных о равновесии основной реакции можно получить представление о максимальном выходе целевого продукта. В состоянии равновесия, к которому стремятся все химические реакции, скорости прямой и обратной реакции становятся одинаковыми, а соотношение концентраций компонентов в рассматриваемой системе (исходные реагенты и продукты реакции) остается неизменным. Концентрации реагирующих компонентов в состоянии равновесия связаны между собой определенным соотношением (*константой равновесия K*), которое является выражением *закона действующих масс*. Величина K дает количественную оценку равновесия. Если ее численное значение велико, то равновесие реакции смещено в сторону образования продуктов реакции, если мало, то в смеси преобладают исходные вещества. Знание констант равновесия для химических реакций позволяет рассчитать такие важнейшие характеристики химической технологии, как *равновесную степень превращения реагента* и *равновесный выход продукта*. *Степень превращения* — это отношение количества израсходованного основного реагента к общему его количеству в начале процесса. *Выход продукта* — это отношение количества полученного целевого продукта к его количеству, которое должно быть получено согласно химическому уравнению.

Основными параметрами, влияющими на равновесие химических реакций, являются температура, давление и концентрация реагирующих веществ. Эти параметра на практике обычно используются для сдвига равновесия в желаемую сторону, то есть для регулирова-

ния равновесной степени превращения в соответствии с законом действующих масс. Так, при повышении температуры равновесие смещается в сторону получения целевых продуктов, если химическая реакция протекает с поглощением тепла, а для реакции, протекающей с выделением тепла, повышение температуры смещает равновесие в сторону исходных веществ. При повышении давления равновесие смещается в сторону получения целевых продуктов, если реакция идет с уменьшением числа молекул в газовой фазе. Если же реакция идет с увеличением числа молекул в газовой фазе, то увеличение давления смещает равновесие в сторону исходных продуктов. При повышении концентрации исходных реагентов равновесие реакции смещается в сторону образования целевого продукта.

Рассмотренные основные закономерности равновесия химических реакций позволяют определить лишь максимально возможный (теоретический) выход целевого продукта. Однако в технологических процессах, ограниченных временем, теоретический выход, как правило, не достигается вследствие уменьшения с течением времени суммарной скорости прямой и обратной реакций. В необратимых процессах скорость реакции стремиться к нулю из-за уменьшения концентрации одного из исходных реагентов. В обратимых процессах уменьшение скорости химической реакции до нуля происходит по мере достижения системой равновесия, когда скорости прямой и обратной реакций становятся равными.

Кроме того, рассмотренные термодинамические соображения, приводящие к увеличению выхода продуктов реакции, часто находятся в противоречии с кинетикой процесса. Например, с точки зрения термодинамики экзотермический процесс желательно вести при низкой температуре, так как равновесная степень превращения увеличивается при понижении температуры. Однако скорость реакции при уменьшении температуры, наоборот, снижается. Таким образом, для установления оптимальных технологических условий проведения процесса необходимо одновременно учитывать как термодинамические, так и кинетические факторы.

Скорость химической реакции зависит от многих факторов, основными из которых являются температура, концентрация компонентов, давление и катализаторы. Сильное влияние на скорость реакции оказывает температура. Однако влияние температуры оказывается неодинаковым для различных типов реакций. Скорость простой необратимой и обратимой эндотермических реакций с увеличением температуры в первом приближении экспоненциально возрастает, а обратимой экзотермической реакции сначала увеличивается, достигает максимума, а затем уменьшается. При увеличении концентрации исходных реагентов скорость химической реакции обычно возрастает. Давление также оказывает большое влияние на скорость химических процессов. Особенно сильно это влияние проявляется в тех случаях, когда процессы протекают в газовой фазе или при взаимодействии газов с жидкими или твердыми веществами. Это объясняется тем, что повышение давления равносильно увеличению концентрации реагирующих веществ. Механизм влияния катализаторов на скорость химической реакции связан со снижением энергии активации процесса. Скорость реакции в присутствии катализатора иногда увеличивается в миллионы раз.

В общем случае скорость роста кристалла из газовой фазы меньше или сравнима со скоростью гетерогенной химической реакции, так как взаимодействию реагентов, находящихся в разных фазах, предшествует их доставка к поверхности раздела фаз и массообмен между фазами. В реальных условиях протекания большинства гетерогенных химико-технологических процессов (процесс роста кристалла из газовой фазы) наряду с химической реакцией необходимо учитывать сопутствующие физические процессы, связанные с макросостоянием системы и накладывающиеся на нее. Гетерогенный химико-технологический процесс представ-

ляет собой совокупность физических (тепло- и массопередача) и химических процессов. Для количественной характеристики такого сложного процесса роста допустимо разделение его на отдельные стадии и анализ каждой из них. Принято выделять по меньшей мере три одновременно протекающих стадии: 1) перенос реагирующих веществ к поверхности раздела фаз — зоне реакции; 2) химическое взаимодействие в зоне реакции; 3) отвод продуктов реакции из зоны реакции. Характер процесса роста может значительно усложняться условиями теплопередачи в зоне реакции (особенно при высоких скоростях выращивания кристаллов, когда скорость процесса главным образом зависит от скорости отвода скрытой теплоты кристаллизации).

Общая скорость процесса будет определяться скоростью самой медленной (лимитирующей) стадии. Если самым медленным звеном процесса в методах химических реакций и химического транспорта является подвод реагирующих веществ к зоне реакции или отвод от нее продуктов реакции, то кинетика суммарного процесса будет лимитироваться диффузионным массопереносом (процесс роста идет в диффузионной области). Если самой медленной стадией процесса является химическое взаимодействие, то скорость всего процесса будет определяться скоростью химических реакций, и тогда считается, что процесс протекает в кинетической области. При соизмеримых скоростях отдельных стадий считается, что процесс протекает в переходной области. Лимитирующая стадия гетерогенного химико-технологического процесса определяется опытным путем при изучении влияния различных факторов технологического режима на общую скорость процесса. Как правило, скорость роста кристалла, выращиваемого из газовой фазы методами химических реакций или химического транспорта, определяется скоростью протекания химических реакций.

Остановимся на методах выращивания кристаллов из газообразной фазы более подробно.

6.4. Метод сублимации–конденсации

Метод сублимации–конденсации можно применять для выращивания конгруэнтно испаряющихся полупроводниковых соединений и твердых растворов на их основе, компоненты которых обладают достаточно высокими давлениями паров (≈ 1 мм рт. ст.), а также в случаях, когда кристаллы материалов сложно вырастить другими методами. Выращивание кристаллов этим методом проводят в замкнутых или проточных систе-

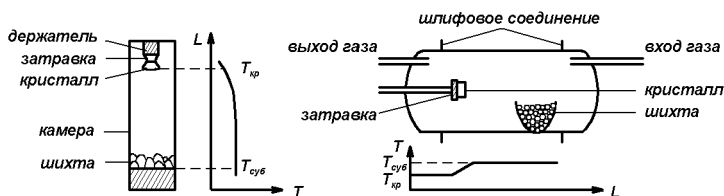


Рис. 6.15. Схемы выращивания кристаллов методом сублимации-конденсации и распределение температуры в реакционной камере: *а* — замкнутая система; *б* — проточная система.

мах, в вертикальной или горизонтальной модификации, с затравкой или без затравки. Схемы различных модификаций этого метода показаны на рис. 6.15.

Сущность метода сублимации-конденсации заключается в возгонке (или испарении) исходного вещества и последующей конденсации паров на растущей поверхности. Метод характеризуется тем, что состав растущего кристалла практически идентичен составу источника, а паровая фаза состоит только из атомов или молекул, образующих источник и кристалл. В зависимости от паровой среды массоперенос в замкнутых системах осуществляется либо молекулярными пучками (в вакууме), либо молекулярной или конвективной диффузией. Лимитирующей стадией роста обычно является массоперенос: подвод кристаллизующегося материала к поверхности роста. Рост кристалла происходит в диффузионной области.

Замкнутая система реализуется в вакуумированной и запаянной кварцевой ампуле. Для выращивания монокристаллов таким способом необходимо следовать следующим правилам.

1. Использовать докритические пересыщения с целью обеспечения селективного зарождения на небольшом числе активных мест на стенках ампулы. Для локализации зарождения и образования единственного зародыша целесообразно вести кристаллизацию в сужающемся конце специально обработанной кварцевой ампулы, также можно использовать создание относительно холодного пятна на стенке ампулы или проводить рост на затравке.

2. Скорость протягивания ампулы должна быть согласована с линейной скоростью роста в направлении оси ампулы. Благодаря выполнению этого условия фронт критического пересыщения будет совпадать с фронтом кристаллизации, поэтому зарождение будет происходить на растущей поверхности кристалла, а не на стенках ампулы.

3. Для обеспечения конвективного переноса следует пользоваться ампулами большого диаметра, выполненными в виде комбинации цилиндра и конуса.

Таким способом, как правило, растят совершенные кристаллы соединений $A^{IV}B^{VI}$ (PbTe, PbS и др.), $A^{II}B^{VI}$ (CdS, CdSe и др.) и твердые растворы на их основе.

При выращивании кристаллов из газообразной фазы в проточных системах методом сублимации-конденсации пары исходного вещества вводятся в зону кристаллизации потоком инертного газа (водород, аргон и др.).

6.5. Метод химических реакций (диссоциации или восстановления газообразных химических соединений)

Метод диссоциации или восстановления газообразных химических соединений оказывается весьма эффективным для выращивания из газообразной фазы монокристаллических слитков тугоплавких соединений, компоненты которых при приемлемых технологических температурах обладают незначительными давлениями паров. В этом методе для получения монокристаллов из газовой фазы используются химические реакции. Источник состоит из газообразных молекул сложного состава, содержащих атомы кристаллизующегося вещества. Кристалл заданного состава образуется в результате химической реакции, протекающей на поверхности затравки или подложки (или вблизи нее) и приводящей к выделению атомов кристаллизующегося вещества.

Суть метода заключается в следующем. Источником материала для роста кристалла служат предварительно очищенные легколетучие химические соединения, которые подаются в реакционную камеру. Эти соединения подвергаются термической диссоциации или восстановлению соответствующим газообразным восстановителем на поверхности роста, что приводит к выделению на ней атомов кристаллизующегося вещества. Примером такого процесса может служить восстановление кремния из четыреххлористого кремния



или получение кремния при термической диссоциации SiH_4



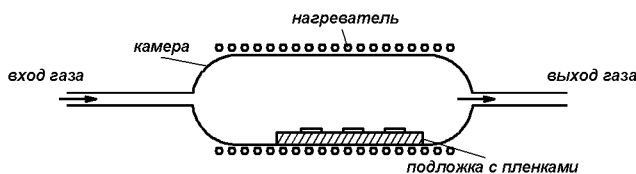


Рис. 6.16. Схема выращивания кристаллов методом химических реакций в проточной системе.

Процессы кристаллизации на растущей поверхности происходят в этом случае в две последовательные стадии: 1) выделение вещества на поверхности роста в результате химической реакции восстановления (разложения); 2) встраивание атомов в решетку кристалла. Поскольку при реакции восстановления (разложения) выделяются, кроме основного, газообразные продукты, то для достижения стационарного, равномерного процесса их необходимо непрерывно удалять, для чего наиболее подходят проточные системы (рис. 6.16).

Поверхность раздела фаз (поверхность растущего кристалла или поверхность подложки, на которой растет монокристаллическая пленка) играет роль катализатора реакции. Каталитическая активность поверхности зависит от природы вещества и его агрегатного состояния, а также определяется морфологией этой поверхности, присутствием на ней активных мест для зарождения центров новой фазы.

Обычное понятие пересыщения, которое определяет термодинамическую скорость роста кристалла, для метода химических реакций оказывается неприменимым. Основным параметром, характеризующим скорость роста, является изменение полной энергии ΔG , равное сумме изменений, соответствующих химическому процессу и встраиванию в решетку выделившихся на поверхности атомов. Первое слагаемое этой суммы зависит от значений констант равновесия основных химических реакций, а второе — от условий закрепления атомов в решетке (механизмов роста). Как показывает опыт, в большинстве случаев лимитирующей стадией роста кристалла методом химических реакций оказываются поверхностные химические реакции.

Крайне важным вопросом технологии выращивания кристаллов при помощи метода химических реакций диссоциации или восстановления является выбор, синтез и подача смеси реагентов. Смесь должна быть такова, чтобы во время химических реакций при температуре роста все продукты реакции (за исключением компонентов кристалла) были газообразными и не встраивались в решетку кристалла. Смесь должна

обладать высокой летучестью при довольно низких температурах. Реакция синтеза реагентов должна обеспечивать получение продукта высокой чистоты. Выход химической реакции должен быть регулируемым в достаточно широком интервале температур путем изменения состава газовой смеси реагентов. Крайне желательны подробные сведения о механизме и кинетике химических процессов, происходящих у поверхности роста. Таким образом, количество кристаллизующегося вещества определяется выходом реакции восстановления (разложения) газообразного соединения при данных условиях (температуре, давлении, концентрации компонентов) и скоростью подачи газовой смеси.

Определение оптимальных условий для выращивания монокристаллов с заданными свойствами методом химических реакций требует в каждом отдельном случае тщательных исследований. Общие же правила таковы: необходима тщательная подготовка подложки, на которой предполагается выращивать кристалл; температура на поверхности подложки, состав и скорость потока газовой смеси должны быть неизменными в течение всего процесса.

6.6. Метод химического транспорта (реакций переноса)

В основе получения монокристаллов из газовой фазы методом химического транспорта лежат химические реакции также, как и в основе метода диссоциации и восстановления газообразных соединений. Рост кристалла происходит в результате реакций разложения газообразных молекул. Примером использования этого метода может служить получение монокристаллического германия:



Сущность метода реакций переноса заключается в следующем (рис. 6.17). При взаимодействии газообразного реагента X (I_2 в нашем примере) с твердым, нелетучим веществом А (источник Ge) при определенных значениях температуры и давлениях паров X могут образовываться разные по составу летучие соединения. При неизменных условиях между ними устанавливается некоторое состояние равновесия. Если изменить температуру системы, то состояние равновесия нарушится и состав компонентов смеси изменится. Выбирая условия, при которых в области источника идет преимущественно реакция образования летучего соединения XА (GeI_4), а в области кристаллизации — его разложение с выделением компонента А (Ge), можно обеспечить условия, при которых

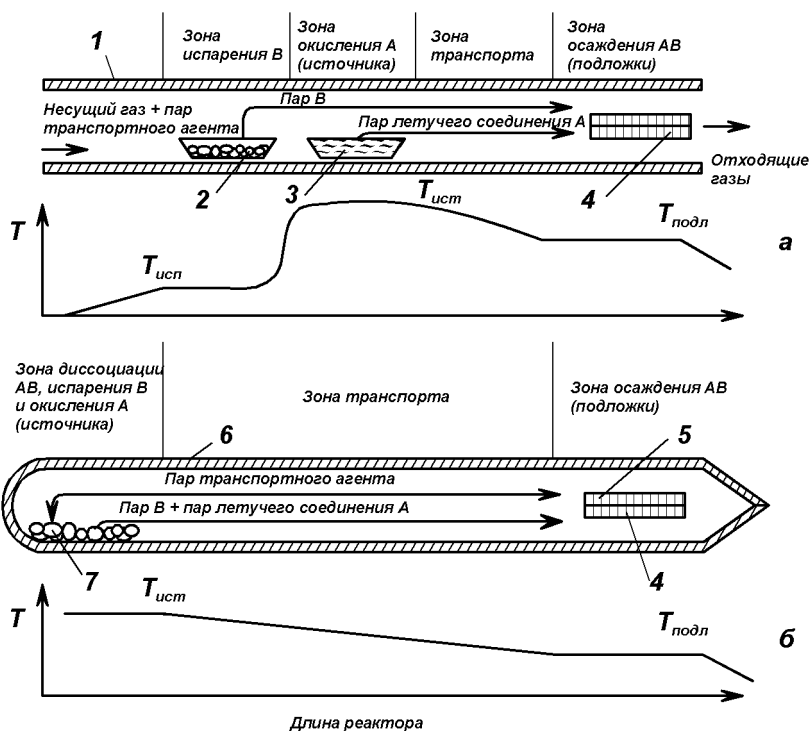


Рис. 6.17. Схема выращивания кристаллов с использованием газотранспортных реакций в открытой (а) и закрытой (б) системах: 1 — реактор; 2 — лодочка с летучим компонентом; 3 — лодочка с нелетучим компонентом; 4 — подложка с кристаллом 5; 6 — кварцевая ампула; 7 — исходное соединение.

будет происходить перенос А (Ge) от источника к растущей поверхности и рост монокристалла вещества А (Ge).

Для того чтобы осуществить перекристаллизацию вещества А, его помещают в один конец откачанной кварцевой ампулы, в которую вводится определенное количество газообразного реагента Х, после чего ампула запаивается. Затем ампула помещается в двухзонную печь таким образом, что источник находится при температуре T_1 , а зона кристаллизации, расположенная на другом конце ампулы, находится при температуре T_2 . Температуры T_1 и T_2 подбираются так, чтобы при T_1 в результате химических реакций реагента Х с материалом источника образовывалось летучее соединение, которое, попадая в зону с температурой T_2 , разла-

галось, образуя вещество А и газообразный реагент Х. Реагент Х, диффундируя по ампуле, снова попадает в зону источника, снова образует летучее соединение, молекулы которого, попадая в зону кристаллизации, распадаются, и т. д. Таким образом, небольшое количество реагента Х может обеспечить перенос и перекристаллизацию произвольно большого количества вещества А.

Необходимыми условиями для осуществления переноса являются обратимость химической реакции и наличие градиента концентрации, которое является следствием наличия перепада температур или давлений. Процесс переноса вещества состоит из трех этапов: 1) химической реакции газообразного переносчика с веществом источника; 2) перемещения газообразного соединения от источника к зоне роста; 3) химической реакции, в результате которой выделяется кристаллизующее вещество.

Выбор химических реакций переноса осуществляется на основе следующих общих правил:

1. Перенос вещества посредством обратимой реакции может происходить лишь в том случае, если переносимый компонент конденсированной фазы присутствует только в одной части уравнения.

2. Реакции, которые при равновесных условиях сильно смещены в одну или другую сторону, не приводят к заметному переносу вещества.

3. Эффективность и направление переноса определяются суммарным изменением полной энергии $\Delta G = \Delta H - T\Delta S$.

4. Выход реакции тем больше, чем больше суммарное изменение свободной энергии.

Скорость роста монокристаллов в этом методе определяется либо диффузионными и конвективными процессами при малых давлениях в системе — тогда скорость роста невелика; либо скоростями реакций, протекающих в зоне источника и в зоне кристаллизации при высоких давлениях — при этом наблюдается резкое повышение скорости роста за счет интенсификации конвекционных потоков.

Таким способом могут выращиваться монокристаллы Ge, GaAs; этот метод является весьма удобным способом получения тугоплавких полупроводников: халькогенидов переходных и редкоземельных элементов (например MnTe, GdS, EuS), сложных полумагнитных полупроводников (HgCr_2Te_4).

Глава 7

Легирование полупроводниковых материалов

Для изготовления многих полупроводниковых приборов необходим легированный материал. Возможны следующие способы легирования: 1) легирование уже выращенных кристаллов; 2) легирование кристаллов в процессе выращивания из жидкой фазы; 3) легирование кристаллов в процессе выращивания из газовой фазы.

1. Легирование уже выращенных кристаллов.

Оно осуществляется методом диффузии примеси из внешней газовой, жидкой или твердой фаз, методом радиационного легирования и методом ионной имплантации. Метод диффузии в технологии производства объемных легированных материалов не получил распространения из-за малых скоростей диффузии в кристаллах. Тем не менее сами процессы диффузии играют большую роль в технологии получения и обработки полупроводниковых материалов и создании приборов на их основе. Рассмотрению этих процессов посвящена гл. 8.

Суть же радиационного (или трансмутационного) метода легирования выращенных кристаллов сводится к следующему. При облучении кристаллов полупроводников и диэлектриков ядерными частицами (нейтронами, протонами, γ -квантами и др.) в результате протекания ядерных реакций может происходить превращение части атомов основного вещества в атомы других химических элементов, которые отсутствовали ранее в веществе. Например, при облучении Si медленными нейтронами образуется нестабильный изотоп ^{31}Si , который, распадаясь с периодом полураспада 2.6 часа, превращается в стабильный изотоп ^{31}P . Эти явления находят все более широкое применение для однородного легирова-

ния выращенных кристаллов. Неоднородность удельного сопротивления при радиационном легировании не превышает 1% на глубине 50 мм. Это значительно превосходит степень однородности легирования кристаллов другими методами. В настоящее время наибольшее практическое применение при радиационном легировании получило использование тепловых нейтронов.

Ионной имплантацией называют процесс внедрения в кристалл ионизированных атомов с энергией, достаточной для проникновения в его приповерхностные области. В настоящее время в электронной промышленности ионная имплантация наиболее широко применяется для ионного легирования кремния при изготовлении полупроводниковых приборов. Энергия легирующих ионов (бора, фосфора или мышьяка) обычно составляет 3–500 кэВ, что достаточно для их имплантации в приповерхностную область кремниевой подложки на глубину 10–1000 нм. Глубина залегания имплантированной примеси, которая пропорциональна энергии ионов, может быть выбрана исходя из требований конкретного применения имплантированной структуры. Основным преимуществом технологии ионной имплантации является возможность точного управления количеством внедренных атомов примеси. Нужную концентрацию легирующей примеси (в кремнии в диапазоне 10^{14} – 10^{21} см⁻³) получают после отжига мишени. Кроме того, можно легко управлять профилем распределения внедренных ионов по глубине подложки. Процесс ионной имплантации, проводимый в вакууме, относится к категории «чистых и сухих» процессов.

Одним из недостатков ионной имплантации и метода радиационного легирования является одновременное с легированием образование в облучаемых кристаллах радиационных нарушений кристаллической решетки, что существенно изменяет электрофизические свойства материала. Поэтому необходимой стадией процесса при получении ионно-легированных и радиационно-легированных кристаллов является термообработка (отжиг) материала после облучения. Отжиг ионно-имплантированных слоев проводится для активирования имплантированных атомов, уменьшения дефектов кристаллической структуры, образующихся при ионной имплантации и радиационном легировании, и в конечном счете, для создания области с заданным законом распределения легирующей примеси и определенной геометрией.¹ Другими недостатками данного метода легирования являются стоимость облучения и необходимость соблюдения

¹Нагрев при термической обработке можно осуществлять кроме обычных методов и методами поверхностного воздействия с помощью лазерного, фотонного, электронного и других видов лучевого воздействия.

мер радиационной безопасности.

Тем не менее ионная имплантация и радиационное легирование кристаллов сейчас — важные и быстро развивающиеся области технологии полупроводников. Так как ионная имплантация обеспечивает более точный контроль общей дозы легирующей примеси в диапазоне 10^{11} – 10^{16} см⁻², там, где это возможно, ею заменяют процессы диффузионного легирования. Очень интенсивно ионная имплантация используется для формирования сверхбольших интегральных схем. Метод радиационного легирования используется для получения кремния, необходимого для производства силовых приборов, где в качестве главного требования выступает высокая однородность распределения примесей в кристалле. Метод радиационного легирования также находит все большее применение и для легирования других полупроводниковых материалов. Так, им осуществляют легирование Ge галлием и мышьяком, InSb оловом, GaAs германием и селеном и т. д.

В данном курсе мы не будем касаться физических процессов, лежащих в основе ионной имплантации и метода радиационного легирования, и вопросов, связанных с применением этих методов для легирования полупроводников и создания полупроводниковых приборов. Эти вопросы подробно изложены в специальной литературе, например [28, 39].

7.1. Легирование объемных кристаллов в процессе выращивания из жидкой фазы (общие принципы)

Рассмотрим особенности легирования кристаллов в процессе их выращивания из жидкой фазы. Широко применяемым методом получения легированных монокристаллов полупроводников является выращивание их из расплава, к которому добавлена нужная примесь.

Общие принципы такого легирования заключаются в следующем. Навеска примеси p_i , подлежащая введению в расплав или жидкую зону для получения в твердом кристалле концентрации N_i , рассчитывается по формуле, определяющей коэффициент разделения примеси (см. гл. 5).

$$\begin{aligned} K_0 &= C_S/C_L = N_i M_i V_L / N_A p_i, \\ p_i &= N_i M_i V_L / K_0 N_A, \end{aligned} \quad (7.1)$$

где M_i — атомный вес примесного элемента, N_A — число Авогадро, V_L — объем расплава.

При выборе примесей для легирования необходимо учитывать величину коэффициента разделения и ее изменение при изменении условий

выращивания. Эти факторы являются чрезвычайно важными для получения монокристаллов с равномерным распределением примесей из-за принципиальной однократности процесса легирования и невозможности исправления ошибок в дозировании примеси.

Пусть, для определенности, $K_0 < 1$. Тогда если скорость роста кристалла V больше, чем скорость выравнивания состава в жидкой фазе, то из-за оттеснения примеси из твердой фазы в жидкую и замедленности диффузионных процессов установления равновесия в жидкой фазе концентрация примеси в расплаве у границы раздела будет возрастать. Накопление избытка примеси приведет к образованию перед движущимся фронтом кристаллизации диффузионного слоя δ , из которого примесь путем диффузии переходит в объем расплава. Если $K_0 > 1$, то вблизи поверхности роста ощущается недостаток примеси. Таким образом, от равновесного коэффициента разделения K_0 мы переходим к эффективному K и учитываем влияние условий выращивания на процессы легирования:

$$K(f, D, \delta) = K_0 / [K_0 + (1 - K_0) \exp(-V\delta/D)], \quad (7.2)$$

где D — коэффициент диффузии примеси в расплаве (см. гл. 5).

В некоторых случаях, если характер перемешивания расплава задан, определение значения δ возможно аналитически. Так, если при получении легированных кристаллов методом Чохральского перемешивание расплава осуществляется вращением кристалла и тигля вокруг своих осей во встречных направлениях с угловыми скоростями ω_k и ω_t , то значение δ при невысоких скоростях роста V может быть определено по формуле

$$\delta = AD^{1/3} \nu^{1/6} (\omega_k + \omega_t)^{-1/2}, \quad (7.3)$$

где A — постоянная, принимающая значение от 1.3 до 1.6, а ν — кинематическая вязкость расплава [37].

Зависимость K от V показана на рис. 7.1 для двух скоростей вращения и трех значений равновесного коэффициента разделения K_0 . Видно, что только при скоростях $V < 4 \cdot 10^{-3}$ см/с можно говорить о совпадении K и K_0 . Из рисунка также видно, сколь важно вращение расплава и (или) кристалла для выравнивания концентрации примеси в расплаве. Особенно это важно в условиях зонной плавки: без перемешивания расплава δ может достигать размеров зоны. При $\delta = 1$ см K и K_0 становятся сопоставимы по величине только при $V < 10^{-4}$ см/с. Именно поэтому скорости выращивания кристаллов в методе зонной плавки значительно меньше, чем при вытягивании кристалла из расплава.

При выборе примеси чрезвычайно важным является учет ее чистоты, так как попадание в растущий кристалл вместе с легирующей приме-

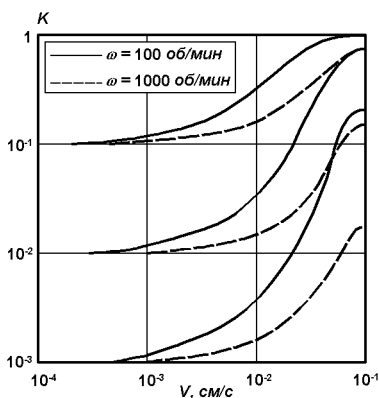


Рис. 7.1. Зависимость K от скорости кристаллизации V при разных скоростях вращения расплава и кристалла.

стью неконтролируемых сопутствующих примесей даже в очень малых количествах может приводить к существенному ухудшению параметров выращиваемых кристаллов (например, существенно снижать время жизни неосновных носителей заряда). Поэтому перед легированием оценивают требуемую чистоту легирующего элемента с учетом особенностей используемого метода легирования и допустимого содержания в легируемом кристалле посторонних примесей.

7.2. Методы выравнивания состава вдоль кристалла

Вопрос об однородности распределения примесей по длине и поперечному сечению выращиваемого монокристалла уже обсуждался в гл. 5 и гл. 6. Из рассмотренного материала можно сделать вывод, что существует несколько факторов, вызывающих появление неоднородностей состава в растущем кристалле. Неоднородности по причинам их возникновения можно разделить на две группы: сегрегационные и технологические.

Сегрегационные (или как их часто называют фундаментальные) связаны с закономерными изменениями состава растущего кристалла, обусловленными основными законами фазовых превращений в многокомпонентных системах (см. гл. 5). Эти закономерные неоднородности охватывают весь объем выращенного кристалла.

Технологические неоднородности имеют незакономерный характер. Они связаны с нарушениями стабильности условий роста кристаллов и охватывают небольшие объемы кристалла (см. гл. 6). Ясно, что техно-

логические неоднородности могут быть устранены усовершенствованиями технологической аппаратуры для выращивания монокристаллов полупроводников и подбором оптимальных условий роста. Подобные способы уже рассматривались в предыдущей главе. В тоже время сегрегационные неоднородности таким путем устранены быть не могут. Для борьбы с ними необходимо разрабатывать специальные методы.

Методы выравнивания сегрегационных неоднородностей состава кристалла делятся на две группы: пассивные и активные методы. В первом случае монокристаллы с заданной однородностью распределения примеси получают без внесения каких-либо изменений в кристаллизационный процесс, то есть используются части кристалла с приблизительно равномерным распределением примеси. Под активными методами подразумеваются такие, которые позволяют активно влиять на ход процесса легирования во время роста, то есть по существу, позволяют программировать процесс изменения состава.

Эффективность метода выравнивания состава оценивают величиной, называемой *выходом процесса* или *выходом годного материала*. Выходом принято называть отношение части количества вещества с необходимыми для дальнейшего использования свойствами к общему его количеству, подвергнутому технологической обработке. В литературе также часто используется такое понятие, как *теоретический выход* годного материала. Имея аналитическое выражение для распределения состава в кристалле, выращиваемом каким-либо методом, можно рассчитать значение теоретического выхода для этого процесса. Поскольку теоретический расчет учитывает только сегрегационные неоднородности состава, то значение теоретического выхода будет характеризовать максимальный предел, выше которого нельзя увеличить реальный выход годного материала.

7.2.1. Пассивные методы выравнивания состава

Однородные кристаллы полупроводников проще всего получить, используя без всяких изменений обычные кристаллизационные процессы: нормальную направленную кристаллизацию и зонную плавку. В этом случае используют приблизительно однородно легированную часть кристалла. Анализ кривых распределения примесей (см. гл. 5) в этих процессах показывает, что наиболее равномерно легированная часть кристалла прижимается к одному из его концов, поэтому целесообразно для дальнейшей работы использовать именно эти части.

При выращивании легированных кристаллов методом нормальной на-

правленной кристаллизации состав исходного расплава обычно задают так, чтобы требуемая концентрация примеси оказалась в начале слитка. Оценка значения расчетного выхода с приемлемым ($\pm 10\%$) разбросом состава по длине слитка при выращивании легированных кристаллов методом нормальной направленной кристаллизации показала низкую эффективность этого метода для получения однородного материала (при $K = 0.1$ теоретический выход равен 10.5%). Однако, благодаря своей простоте этот метод иногда применяется для получения легированных монокристаллов методом Чохральского.

Рассмотрим теперь возможности использования приблизительно равномерно легированной части кристаллов, выращиваемых методом зонной плавки. Теоретический выход процесса в этом случае, как показывают оценки, зависит от распределения примеси в исходном материале. Наиболее распространенным в данной технологии является в среднем равномерное исходное распределение примеси.

Анализ теоретического выхода процесса зонной плавки при этих условиях показал [37], что выход тем больше, чем коэффициент распределения ближе к 1; он возрастает с увеличением приведенной длины кристалла $A = L/l$, где L — длина кристалла, l — длина зоны, и в пределе стремиться к $(A - 1)/A$. Выход в этом случае может быть сделан достаточно близким к 100% для однократного прохода расплавленной зоны через равномерный в среднем по составу образец. Таким образом, метод зонной плавки может удовлетворить большинству практических требований при получении однородных кристаллов, если только приведенная длина слитка достаточно велика.

Эффективным способом повышения теоретического выхода при выращивании кристаллов методом зонной плавки является выравнивание существенно неоднородного распределения примесей в начальной части слитка. Для этого в начальной части кристалла длиной в одну зону создается средняя концентрация примеси в K раз меньшая, чем в остальной части кристалла. Такой метод создания исходного распределения примеси получил название *целевой загрузки*. Для примесей с $K < 1$ и с равномерным в среднем начальным распределением по объему кристалла необходимое распределение проще всего создается легированием зоны в начальной части образца. После расплавления зоны в нее вводится примесь в таком количестве, что при движении зоны вдоль образца она с самого начала имеет постоянный состав: в нее через границу плавления входит ровно столько примеси, сколько уходит через границу кристаллизации. Вследствие этого состав выращиваемого кристалла постоянен по всей длине за исключением его конца, где процесс идет по закону

нормальной направленной кристаллизации.

Также на практике часто используется метод *прохождения легирующей зоны через чистый исходный образец* для получения равномерно легированных кристаллов. Суть метода сводится к следующему. В расплавленную в начале чистого кристалла зону вводят легирующую примесь. Для примеси с $K \ll 1$ при кристаллизации расплава из зоны уходит настолько мало примеси, что состав жидкой фазы практически не меняется и, таким образом, получается однородно легированный материал. Поэтому этот метод наиболее эффективен для примесей с $K \ll 1$.

7.2.2. Активные методы выравнивания состава кристаллов

Эти методы служат для повышения выхода материала с равномерным распределением примеси. Их отличительной особенностью является то, что в течение всего кристаллизационного процесса в него вводятся определенные изменения. Активные методы выравнивания состава подразделяются на две основные группы.

Первая — включает методы, в которых с целью поддержания концентрации примеси в расплаве в течение всего процесса выращивания монокристалла постоянно проводят подпитку расплава либо нелегированной твердой, жидкой или паровой фазой (если $K < 1$), либо фазой, содержащей легирующую примесь (если $K > 1$). В этих методах в процессе кристаллизации должны соблюдаться следующие условия: $C_L = \text{const}$, $K = \text{const}$, $V = \text{const}$, $C_S = KC_L = \text{const}$. Система процесса выращивания однородного кристалла в наиболее общем виде включает в себя следующие элементы: растущий кристалл, расплав и поступающую в него подпитывающую массу.

Вторая — методы, в которых изменяются сами условия роста монокристаллов. Процессы выращивания однородных кристаллов этими методами протекают при следующих условиях: $C_L \neq \text{const}$; $K \neq \text{const}$; $f \neq \text{const}$; $C_S = K_n C_L^n = K_k C_L^k = \text{const}$, где n — начало, а k — конец. При этом программа изменения кристаллизационного процесса обеспечивает постоянство скорости захвата примеси в течение всего процесса выращивания монокристалла.

Помимо этих двух групп методов возможны и их комбинации. Рассмотрим сначала первую группу методов более подробно.

Первая группа методов выравнивания состава кристаллов

Механическая подпитка расплава твердой фазой

Возможны два способа:

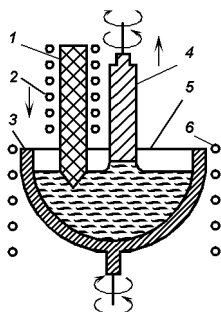
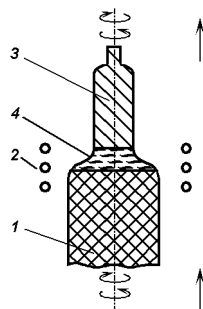


Рис. 7.2. Схема метода механической подпитки расплава твердой фазой: 1 — питающий кристалл; 2 — нагреватель для подогрева питающего кристалла; 3 — тигель; 4 — выращиваемый кристалл; 5 — расплав; 6 — основной нагреватель.

1. Идея первого способа состоит в опускании в расплав подпитывающего стержня (рис. 7.2). Процессом подпитки можно управлять, меняя площадь поперечного сечения подпитывающего стержня, его состав и механическую скорость его подачи. При необходимости в расплав может одновременно вводиться несколько стержней. Для того, чтобы получить математическое выражение, описывающее процесс выравнивания состава в данном методе, необходимо составить уравнение баланса примеси в расплаве и приравнять изменение концентрации примеси нулю. Из этого уравнения для любого варианта механической подпитки расплава опускающимся стержнем легко найти условия, обеспечивающие получение однородного кристалла. Так, для наиболее интересного с практической точки зрения режима получение легированных кристаллов решение уравнения сводится к отысканию либо нужной концентрации подпитывающего стержня при заданных остальных параметрах, либо к отысканию его площади поперечного сечения. В частности, если сечения вытягиваемого кристалла и стержня подпитки равны, и равны их плотности, то состав подпитывающего стержня должен быть равен составу растущего кристалла. Этот способ выравнивания состава позволяет получать однородные монокристаллы с высоким выходом и большим диапазоном уровней легирования. Он используется и для выращивания монокристаллов твердых растворов (см. гл. 6), например, в таких системах как Ge–Si, Bi–Sb, InAs–GaAs и т. д.

2. Второй способ подпитки — *метод расплавленного слоя* (рис. 7.3). В этом случае слиток подпитывающего материала помещают в нижней части кристалла, выращивание которого ведут с вершины подпитывающего слитка, подплавляемого специальным нагревателем. Рост кристалла в этом случае сопровождается синхронным перемещением подпитывающего слитка вверх.

Рис. 7.3. Схема метода подпитки расплава твердой фазой (метод расплавленного слоя): 1 — питающий кристалл; 2 — нагреватель; 3 — выращиваемый кристалл, 4 — расплав.



Анализ условий выращивания однородного по составу кристалла и выхода годного материала в *методе расплавленного слоя* проводится аналогично предыдущему случаю. Однако этот метод имеет определенные преимущества: а) возникающие в расплаве концентрационные и тепловые потоки симметричны; б) процесс выращивания проводится в бестигельных условиях. Эти преимущества послужили стимулом для развития этого метода и изготовления легированных бездислокационных кристаллов полупроводников.

Механическая подпитка расплава жидкой фазой

Механическую подпитку кристаллизующего расплава жидкой фазой чаще всего осуществляют при выращивании кристаллов методом Чохральского. Наибольшее распространение получили две модификации этого метода: первый — вытягивание монокристалла из расплава в плавающем тигле или в тигле, механически перемещающемся относительно внешнего контейнера, с которым они связаны капиллярным каналом (рис. 7.4); второй метод — вытягивание кристалла из тигля, разделенного перегородкой, через которую рабочая и подпитывающая части тигля соединены капиллярным каналом (рис. 7.5). В обеих модификациях в рабочем режиме в соединительном канале идет непрерывный поток расплава по направлению к рабочему объему. При этом перенос примесного компонента в канале состоит из двух частей: потока, вызванного потоком жидкости, и потока, обусловленного молекулярной диффузией. Для управления процессом выравнивания состава вытягиваемого кристалла необходимо, чтобы перенос примеси в канале осуществлялся только механическим перетоком расплава, а выравнивающее действие молекулярной диффузии было подавлено. Это условие легче всего выполняется при использовании длинных и узких соединительных каналов-капилляров.

В процессах с плавающим или механически перемещаемым рабо-

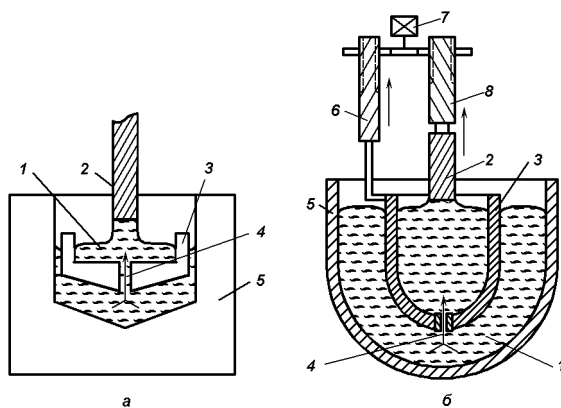


Рис. 7.4. Основные варианты капиллярной подпитки: *а* — плавающий тигель; *б* — механически опускаемый тигель (*1* — расплав; *2* — кристалл; *3* — внутренний тигель; *4* — капилляр трубка; *5* — внешний тигель; *б* — держатель внутреннего тигля; *7* — двигатель; *8* — держатель кристалла).

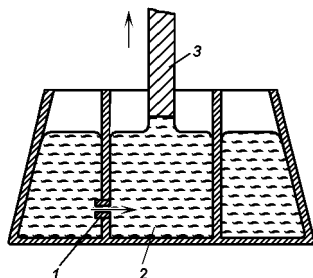


Рис. 7.5. Схема капиллярного двойного тигля: *1* — капиллярный канал; *2* — расплав; *3* — кристалл.

чим тиглем наибольшее распространение получили режимы выращивания, при которых обеспечивается постоянство рабочего объема расплава. Поэтому распределение примесного состава в кристалле, выращиваемом этим методом, аналогично распределению состава в кристалле, получаемом с помощью метода зонной плавки. Однако рассмотренный метод получения однородно легированных кристаллов имеет ряд важных преимуществ по сравнению с методом зонной плавки: 1) отсутствие необходимости приготовления исходных образцов нужного состава; 2) возможность непосредственного наблюдения за процессом вытягивания; 3) постоянство рабочего расплава (в случае с плавающим тиглем) или возможность его варьирования в широких пределах в случае с механическим перемещением тигля.

Процесс выращивания легированных кристаллов из капиллярного двойного тигля является своеобразной комбинацией процессов зонной плавки и нормальной направленной кристаллизации. Для обеспечения постоянства состава кристаллов, выращиваемых этим методом, достаточно, чтобы на протяжении всего процесса роста при изменении высоты расплава отношение площади рабочей части тигля ко всей его площади сохранялось постоянным. Это требование легко реализовать, например, в тигле внутри которого коаксиально устанавливается цилиндрическая перегородка с капиллярным каналом (рис. 7.5). Данный метод успешно применяется для легирования кристаллов примесями с $K > 1$ и $K < 1$. Однако анализ теоретического выхода годного материала показал, что для примесей с $K \ll 1$ применение капиллярных тиглей с целью получения однородно легированных кристаллов нецелесообразно.

Механическая подпитка расплава газовой фазой

Механическую подпитку кристаллизующего расплава газовой фазой осуществляют как при выращивании кристаллов методом Чохральского, так и при выращивании методом зонной плавки. Подпитка может проводиться как нелетучей, так и летучей примесью.

В случае нелетучей примеси обычно эта примесь подводится к расплаву в виде летучего химического соединения. Последнее, взаимодействуя с расплавом, разлагается, вводимая примесь выделяется и легирует расплав. Такой процесс выделения примеси в результате химических реакций используется лишь для примесей с $K > 1$.

При легировании кристаллов летучими примесями, обладающими при температуре расплава высоким давлением пара, необходимо учитывать взаимодействие расплава с паровой фазой. Процесс подпитки кристаллизующего расплава паровой фазой может осуществляться как путем поглощения примеси ($K > 1$) из газовой фазы, так и путем ее испарения

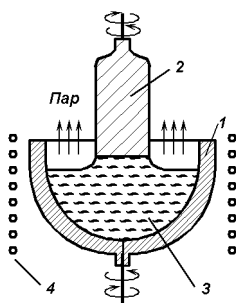


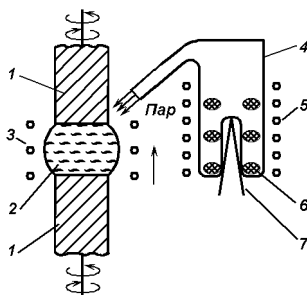
Рис. 7.6. Схема подпитки кристалла с помощью испарения летучей примеси с $K < 1$ из расплава: 1 — тигель; 2 — выращиваемый кристалл; 3 — расплав; 4 — нагреватель.

($K < 1$) из расплава. Переход летучего компонента через поверхность раздела расплав–газ будет отсутствовать только в том случае, когда концентрация растворенной в расплаве летучей примеси находится в равновесии с ее концентрацией в газовой фазе. Изменение состава расплава за счет сегрегации вызывает обмен летучим компонентом между расплавом и газом. Процесс межфазного обмена, который в конце концов приводит концентрации примеси в расплаве и газе к равновесным значениям, прекратится только тогда, когда прекратится изменение состава расплава. Если в системе имеется летучий компонент, то при наличии свободных поверхностей расплава обязательно происходит обмен летучим компонентом между расплавом и газом и равновесная подпитка расплава этим компонентом.

При работе с летучими компонентами использование равновесной газовой подпитки можно вести по двум схемам. В первой из них стенки установки для выращивания кристалла специально не подогреваются, и их температура, как правило, близка к комнатной. В стационарных условиях вся примесь должна испариться из расплава и осесть на холодных стенках системы. При выращивании же легированных кристаллов все происходит по-другому. Равновесное испарение примеси компенсируется механической подачей в расплав летучего компонента. Например, простейший вариант такого процесса реализуется при равенстве оттесняемого на фронте кристаллизации в расплав в единицу времени количества примеси и количества примеси, испаряющейся за это же время со свободной поверхности расплава (рис. 7.6). Необходимая для этого процесса скорость роста кристалла находится из уравнения баланса примеси в расплаве.

По второй схеме выращивание кристаллов ведут в установке со специально подогреваемыми стенками рабочей камеры. В этом случае си-

Рис. 7.7. Схема механической подпитки расплава паровой фазой: 1 — выращиваемый кристалл; 2 — расплав; 3 — основной нагреватель; 4 — сосуд с летучей примесью; 5 — вспомогательный нагреватель, регулирующий давление пара, скорость его подачи к поверхности расплава; 6 — конденсат летучей примеси; 7 — термопара.



стема представляет собой замкнутый, предварительно вакуумированный объем с определенным температурным полем (рис. 7.7). В зоне пониженных температур помещают избыточное количество летучего компонента (источник). Температуру основной части системы, в которой происходит рост легированного кристалла, выбирают промежуточной между температурой источника и температурой расплава. Процесс кристаллизации ведут так, чтобы изменение состава расплава, вызванное сегрегационными явлениями, компенсировалось поглощением примеси из газовой фазы расплавом (должно установиться равновесное распределение примеси в системе расплав–подпитывающая газовая фаза). Для получения однородного по составу кристалла необходимо, чтобы равновесие в системе устанавливалось со скоростью большей, чем скорость сегрегационных изменений состава жидкой фазы. Это возможно только при очень малых скоростях роста кристалла, значительно меньших, чем обычно используют на практике. Поэтому этот метод при выращивании однородных объемных кристаллов применяют сравнительно редко (обычно для выращивания концентрированных твердых растворов тугоплавких материалов и диссоциирующих соединений).

Вторая группа методов выравнивания состава кристаллов (изменение условий выращивания)

Выравнивание состава выращиваемого кристалла с помощью программного изменения условий роста можно построить исходя из двух принципов.

Процессы, основанные на первом принципе, сводятся к программному изменению скорости вытягивания и вращения кристалла. Эта методика получила достаточно широкое распространение и имеет много модификаций. Суть методов, основанных на этом принципе, сводится к следующе-

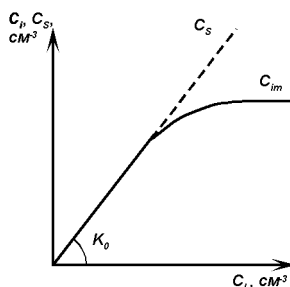
му. Если легирование кристалла проводится нелетучей примесью с $K < 1$, то по мере вытягивания монокристалла концентрация примеси в расплаве непрерывно увеличивается и для получения равномерно легированного кристалла режим выращивания должен быть построен так, чтобы по мере роста кристалла коэффициент разделения примеси непрерывно уменьшался. Управлять изменением K от K_{max} до K_{min} можно изменением скорости выращивания кристалла V , частоты его вращения ω , а также выбором кристаллографической ориентации затравки и, соответственно, направлением выращивания.

Согласно второму принципу параметры системы меняются так, что состав расплава в течение процесса выращивания кристалла остается постоянным. Этого можно добиться в случае, когда подпитка не производится и легирование кристалла осуществляется нелетучей примесью, изменением объема расплава в ходе процесса, то есть либо изменением длины расплавленной зоны, либо применяя слитки переменного сечения. Однако вследствие сложности практического осуществления эти методы выравнивания состава на практике не применяются.

В заключении следует отметить, что с точки зрения получения высоко совершенных кристаллов, подпиточные методы вследствие стационарности условий роста кристаллов обладают существенными преимуществами перед методами программного изменения условий роста, при использовании которых, в принципе, трудно ожидать получения совершенных кристаллов. Это, в свою очередь, и обуславливает более широкое применение подпиточных методов для получения однородных кристаллов.

7.3. Растворимость примесей

Для ряда практических применений (создание туннельных диодов, светодиодов и других полупроводниковых приборов) необходимо получать сильно легированные полупроводники. Поэтому представляется важным знание *предельной растворимости* C_{Smax} примесей в материале (в твердой фазе). Под этим термином подразумевается концентрация примеси в насыщенном твердом растворе, образованном основным веществом и данной примесью. Если концентрация примеси в полупроводнике меньше C_{Smax} , то примесь распределяется в кристаллической решетке монотонно; если превышает C_{Smax} , то, как показывают исследования, в выращиваемом кристалле появляются структурные нарушения, например, макроскопические частицы инородной фазы, что сопровождается резким ростом, в первую очередь, плотности дислокаций. При легировании кристаллов большими концентрациями примесей важно иметь «запас» в рас-

Рис. 7.8. Зависимость C_i и C_S от C_L .

творимости, чтобы обезопасить кристалл от появления подобных структурных нарушений.

Кроме того, когда речь идет о сильном легировании полупроводников электрически активной примесью (легирующие примеси), то надо иметь в виду, что, например, в элементарных полупроводниках довольно часто наблюдается несоответствие между концентрациями носителей заряда и электрически активной примеси $n(p) < C_{Smax}$. Поэтому было введено также понятие предельной растворимости электрически активной примеси. *Предельной растворимостью электрически активной примеси C_{im}* называют максимальную концентрацию электрически активной примеси в твердой фазе, которую можно создать введением данной примеси. Экспериментально C_{im} можно определить измеряя зависимость концентрации электрически активной примеси в твердой фазе C_i от концентрации примеси в расплаве C_L . Если C_i измерить с помощью эффекта Холла (см.гл. 5), то есть электрическим методом, а полную концентрацию примеси в твердой фазе C_S измерить каким-либо другим методом, например, радиоактивным или масс-спектроскопическим, то общий вид зависимостей C_i и C_S от C_L может быть представлен графиком, изображенным на рис. 7.8. В области достаточно малых C_L значения C_S и C_i совпадают; однако, начиная с некоторых C_L , кривая $C_i(C_L)$ обнаруживает тенденцию к насыщению, в то время как $C_S(C_L)$ продолжает расти. Значение C_{im} и есть предельная растворимость электрически активной примеси.

При $C_S < C_{im}$ примесь распределяется в решетке моноатомно, создавая в запрещенной зоне локальные энергетические уровни. При $C_S > C_{im}$ распределение примеси по кристаллу перестает быть однородным. Причем в начале наблюдается небольшое расхождение в значениях C_i и C_S , но вскоре C_i достигает своего максимального значения C_{im} , а C_S продолжает расти с ростом C_L . Этот эффект связан с тем, что при больших концентрациях примеси примесные атомы, по-видимому, могут образовывать

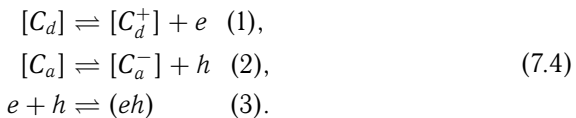
ассоциации (комплексы, кластеры), состоящие из нескольких атомов, а в пределе даже частицы инородной фазы. Атомы, образующие ассоциацию, ионизируются лишь частично (отсюда наблюдаемое небольшое расхождение в значениях C_i и C_S в начале). Возможны и другие причины. В частности, при значительных концентрациях донорной добавки в полупроводнике могут возникать компенсирующие дефекты акцепторного типа (в силу условия электронейтральности кристалла).

Таким образом, в сильно легированном полупроводнике электрически активная примесь может существовать в нескольких формах: одиночные примесные атомы (ионы); ассоциации примесных атомов различного типа; ассоциации, состоящие из примесных атомов и собственных дефектов.

При одновременном введении в полупроводник двух электроактивных добавок они при определенных условиях взаимно влияют на свою предельную растворимость в полупроводнике. Возможны различные случаи взаимодействия примесей при совместном легировании.

1. Взаимодействие примесей, связанное с электронно-дырочным равновесием.

В качестве примера рассмотрим полупроводник, который легируется донорной примесью в присутствии акцепторной. Если в результате электронно-дырочного взаимодействия при таком легировании возникают недиссоциированные или слабодиссоциированные электронно-дырочные пары, то растворимость донора в присутствии акцептора растет. Действительно,

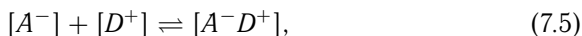


Из закона действующих масс следует, что дырки h , образованные вследствие реакции (7.4.2), могут взаимодействовать с электронами e (согласно реакции (7.4.3)), и это изменит равновесие реакции (7.4.1) в сторону больших значений $[C_d^+]$, а следовательно, и $[C_d]$. Этот процесс эффективен при низких температурах.

Очевидно, что введение в кристалл двух добавок одного знака приводит к взаимному понижению растворимости. Рассмотрение процесса легирования полупроводника акцепторной примесью в присутствии донорной приводит к результатам, аналогичным полученным для случая легирования полупроводника донорной примесью в присутствии акцепторной.

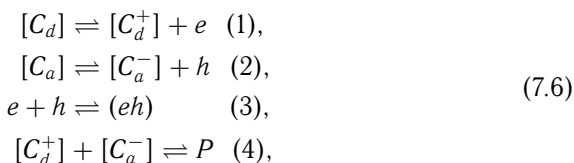
2. Взаимодействие между примесными ионами, приводящее к образованию нейтральных пар, устойчивых при низких температурах.

Обычно такое взаимодействие реализуется, если один из ионов подвижен. Подвижными ионами в полупроводнике могут являться, например, ионы примесей внедрения с малыми ионными радиусами; таким образом, нейтральные пары могут образовываться при взаимодействии иона примеси внедрения с ионом примеси замещения. Реакция взаимодействия может быть записана следующим образом:



где A^- — ион акцептора; D^+ — ион донора; $[A^-D^+]$ — нейтральная ионная пара.

Уравнения равновесия реакции (7.5)



где P — концентрация нейтральных пар.

Образование нейтральных пар приводит к смещению равновесий (7.6) вправо, а следовательно, к взаимному увеличению растворимости донора и акцептора.

3. Взаимодействие между примесными ионами, приводящее к образованию комплексов, устойчивых в широком интервале температур.

Устойчивыми могут являться комплексы, имеющие структуру, подобную структуре матрицы. Такие комплексы возникают при взаимодействии примесей, образующих двойную фазу, подобную по структуре матрице. В тройной системе элементарный полупроводник — первая примесь — вторая примесь эта фаза находится в равновесии с элементарным полупроводником. Протяженность области твердых растворов на основе элементарного полупроводника, как правило, наибольшая в направлении изоструктурной фазы.

Присутствие дефектов в кристалле обычно повышает предельную растворимость примеси в нем. Так, если растворимость выбранной примеси увеличивается вблизи дислокаций, то в кристалле, содержащем дислокации, равновесная растворимость примеси также будет больше:

$$C_d = C_0 \exp(W_d/kT), \quad (7.7)$$

где C_d — равновесная концентрация примеси вблизи дислокации, C_0 — равновесная концентрация примеси вдали от дислокации, а W_d — энергия взаимодействия примесного атома с дислокацией.

С предельной растворимостью электрически активной примеси в кристаллах тесно связан и ее коэффициент разделения K_0 . Анализ корреляции между предельной растворимостью и коэффициентом разделения K_0 электрически активной примеси в германии и кремнии показывает связь типа:

$$X_m \equiv C_{im}/N_0 \approx 0.1K_0, \quad (7.8)$$

где X_m — относительная предельная растворимость электрически активной примеси в атомных долях, а N_0 — концентрация атомов основного вещества.

Таким образом, предельные растворимости электрически активных примесей и их коэффициенты разделения связаны прямо пропорциональной зависимостью и отражают один и тот же процесс вхождения и устойчивости положения примесных атомов в узлах или междоузлиях кристаллической решетки основного вещества (см. гл. 8). Из соотношения (7.8) и значений K_0 , приведенных на рис. 5.4, можно определить значения C_{im} , учитывая, что N_0 для германия и кремния равны соответственно $4.5 \cdot 10^{22} \text{ см}^{-3}$ и $5 \cdot 10^{22} \text{ см}^{-3}$.

Широкий диапазон K_0 и C_{im} для разных примесей указывает на одну весьма важную проблему легирования полупроводниковых материалов примесями с малой растворимостью и малыми K_0 : высокое требование к чистоте самих примесей. Для примера сравним легирование германия галлием, содержащим примесь индия в количестве 10^{-4} атомных долей; и легирование золотом, содержащим примесь галлия в том же количестве. В первом случае соотношение концентраций основной и неосновной примесей в легированном кристалле (с учетом соответствующих K_0) будет $C_{Ga}/C_{In} \cong 10^6$, то есть загрязнением галлия индием можно полностью пренебречь. Во втором случае $C_{Au}/C_{Ga} \cong 1$ и, таким образом, само понятие «основная примесь» теряет смысл. Особые требования к чистоте золота диктуются еще и тем, что галлий (сопутствующая примесь в золоте) относится к группе «мелких» примесей, поэтому многозарядная система глубоких уровней, создаваемых золотом, может быть полностью или частично искажена.

Глава 8

Диффузия примесей в кристаллах

Процессы диффузии очень распространены и играют огромную роль во многих важнейших технологических процессах получения и обработки полупроводников, а также при фазовых и структурных превращениях. Диффузия примесей лежит в основе процесса гомогенизации свойств материала при термообработках, в ряде случаев лимитирует очистку, играет огромную роль при распаде пересыщенных твердых растворов, их упорядочении и разупорядочении. Процессы диффузии используются для получения $p-n$ -переходов, для формирования базовых и эмиттерных областей и резисторов в биполярной технологии изготовления полупроводниковых приборов, для создания областей истока и стока в МОП-технологии и т. д. Поэтому знание основных законов диффузии, диффузионных параметров примесей необходимо для выбора оптимальных режимов технологических операций, а также для понимания некоторых эффектов, проявляющихся в процессе изготовления полупроводниковых схем и приборов.

Исторически изучение процессов диффузии велось в направлении создания на основе экспериментальных результатов моделей, которые давали бы возможность предсказывать протекание процесса диффузии путем теоретического анализа. Для технологов конечной целью исследования процесса диффузии являлась возможность расчетным путем определять электрические характеристики полупроводниковых приборов на основе технологических параметров процесса. Диффузионные модели развивались с позиции двух основных приближений: 1) теории сплошных сред с использованием основных уравнений диффузии и 2) атомистической тео-

рии, которая принимает во внимание взаимодействие между точечными дефектами и примесными атомами.

При низких значениях концентрации примеси измеренные диффузионные профили хорошо согласуются с решениями уравнения диффузии с постоянными значениями коэффициентов диффузии. В этих случаях конкретные пути перемещения атомов примеси, как правило, не известны. При высоких значениях концентрации примеси форма диффузионных профилей отклоняется от формы, предсказанной простой диффузионной теорией. Это отклонение обусловлено влиянием на процесс диффузии примесей факторов, не принятых во внимание в диффузионных уравнениях, и в первую очередь, концентрационной зависимостью коэффициентов диффузии, входящих в уравнение диффузии. Для объяснения экспериментальных результатов концентрационной зависимости коэффициентов диффузии и для анализа других аномальных результатов процесса диффузии были предложены различные атомные модели, основанные на взаимодействии дефектов с примесными атомами. Атомистическая теория диффузии находится сейчас еще в процессе активного развития, поэтому мы начнем с обсуждения диффузионных процессов в рамках первого приближения.

Под диффузией понимают процесс установления внутри кристалла равновесного распределения концентраций, обусловленный тепловым движением атомов и реализуемый обычно их перемещением. Элементарный акт диффузии состоит в скачке атома на расстояние порядка межатомного, то есть переходе из одного равновесного положения в решетке в другое.

Диффузию атомов основного вещества называют *самодиффузией*, а атомов примеси — *гетеродиффузией*. Движущей силой диффузии является градиент химического потенциала μ , наличие которого может быть вызвано различными причинами. Согласно термодинамике необратимых процессов для диффузии в одном направлении (простейший случай) количество вещества, продиффундировавшего в единицу времени через единицу площади, перпендикулярной направлению диффузии, то есть плотность диффузионного потока атомов M , пропорционально этому градиенту химического потенциала

$$M = -P \frac{\partial \mu}{\partial x}, \quad (8.1)$$

где P — коэффициент пропорциональности, определяющий скорость выравнивания химического потенциала. Знак минус в уравнении означает, что поток направлен в сторону, противоположную градиенту.

Наиболее распространена диффузия для случая, когда градиент химического потенциала вызван градиентом концентрации dC/dx какого-либо компонента в многокомпонентной системе. В случае $dC/dx \neq 0$ тепловое движение перестает быть хаотическим, оно становится направленным в сторону выравнивания концентрации. Уравнение (8.1) в этом случае примет вид

$$M = -D \frac{\partial C(x, t)}{\partial x} \quad (8.2)$$

где C — концентрация диффундирующего вещества, которая в нашем случае зависит только от x и t ; x — ось координат, совпадающая с направлением потока диффундирующего вещества; t — время диффузии; D — коэффициент диффузии, который определяет скорость выравнивания концентрации диффундирующего вещества. Коэффициент диффузии имеет размерность $L^2 t^{-1}$ и обычно выражается в $\text{см}^2 \text{с}^{-1}$. Формула (8.2) представляет собой первый закон диффузии (первый закон Фика).

Для практических расчетов часто удобнее иметь формулу, выражающую концентрацию диффундирующего вещества в разных точках образца в разные моменты времени (в зависимости от времени диффузионного отжига).

Если взять две параллельные единичные площадки с координатами x и $x + \Delta x$, то внутри объема, ограниченного указанными площадками, за время dt войдет диффундирующее вещество в количестве M_1 , а выйдет из этого объема в количестве M_2 . В итоге внутри объема количество диффундирующего вещества изменится на

$$\Delta M = M_1 - M_2 = \left(D \frac{dC}{dx} \Big|_{x+\Delta x} - D \frac{dC}{dx} \Big|_x \right) dS dt = d \left(D \frac{dC}{dx} \right) dS dt. \quad (8.3)$$

Изменение концентрации диффундирующего компонента в единице объема запишется так:

$$dC = \frac{\Delta M}{dS dx} = d \left(D \frac{dC}{dx} \right) \frac{dt}{dx}, \quad (8.4)$$

откуда скорость изменения концентрации:

$$\frac{dC}{dt} = \frac{d}{dx} \left(D \frac{dC}{dx} \right), \quad (8.5)$$

или в частных производных для случая, когда коэффициент диффузии D не зависит от концентрации (при низких концентрациях диффундирующего вещества),

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial C}{\partial x} \right) = D \frac{\partial^2 C}{\partial x^2}. \quad (8.6)$$

Полученное выражение представляет собой второй закон диффузии (второй закон Фика). Оно описывает процесс изменения концентрации диффундирующего вещества в разных точках среды как функцию времени.

Направленный поток атомов может возникнуть не только из-за наличия градиента концентрации, но он может возникнуть и в его отсутствии под воздействием направленных внешних сил. Такая диффузия называется *вынужденной*. Она может приводить не только к выравниванию, но и к увеличению градиента концентраций. Основными видами вынужденной диффузии являются следующие.

Термодиффузия — диффузия в условиях наличия градиента температур dT/dx . Характерный пример — диффузия носителей тока от горячего к холодному в термоэлектрических преобразователях энергии.

Электродиффузия — диффузия в условиях наличия градиента электрического потенциала $d\varphi/dx$. Характерные примеры электродиффузии — электролитические процессы.

Восходящая диффузия — диффузия в условиях наличия градиента упругих напряжений dU/dx . Схематически это можно представить себе следующим образом. Пусть некоторый объем кристалла содержит атомы двух элементов А и В, имеющих существенно различные размеры: $R_A > R_B$. В равновесном состоянии распределение атомов А и В в разных макрообъемах будет однородным. Если теперь изогнуть кристалл и в изогнутом состоянии нагреть до такой температуры, при которой станет возможной интенсивная диффузия атомов, то под влиянием стремления системы к уменьшению упругих напряжений атомы А будут диффундировать в растянутую область кристалла, а атомы В — в сжатую. В результате возникнет градиент концентраций. Характерным примером такой диффузии является образование «примесных атмосфер» вокруг дислокаций.

8.1. Возможные механизмы диффузии в твердых телах

Для совершения элементарного акта диффузии атом должен преодолеть энергетический барьер, величина которого определяется энергией активации диффузии Q . Последняя зависит от прочности связи атомов в кристаллической решетке вещества. В зависимости от конкретного механизма диффузии Q по порядку величины может составить от долей до нескольких эВ.

Средняя тепловая энергия колеблющихся атомов kT в зависимости

от температуры составляет $\sim 0.025\text{--}0.2$ эВ. Для того, чтобы диффузия стала возможной, нужна флуктуация энергии, превышающая Q . Вероятность элементарного акта диффузии, то есть частота перескоков, следовательно, зависит от вероятности таких флуктуаций.

При обычных тепловых колебаниях относительно положения равновесия атом ν раз в секунду «сталкивается» с потенциальным барьером. Большую часть времени энергия атома мала для преодоления барьера, но иногда его энергия благодаря флуктуации может оказаться больше энергии активации ($E_{\text{фл.}} > Q$) и атом преодолевает барьер. Относительное время, в течение которого атом обладает такой энергией, пропорционально $\exp(-Q/kT)$ согласно статистике Больцмана. Следует также учесть, что атом в кристаллической решетке может с равной вероятностью перейти в любое из Z равноценных мест в ближайшей координационной сфере (Z — множитель, зависящий от геометрии элементарной ячейки кристалла). С учетом этого общая частота прыжков $\dot{f}$, совершаемых с исходной позиции, за счет флуктуаций энергии в расчете на один атом будет равна

$$\dot{f} = Z\nu \exp(-Q/kT). \quad (8.7)$$

Каков же детальный механизм отдельных атомных прыжков? Единого универсального механизма диффузионных перемещений для всех материалов и условий нет. Эти механизмы зависят от природы химических связей, типа и компактности решетки, природы диффундирующей примеси, температуры диффузии и других факторов. В некоторых веществах диффузионные перемещения могут происходить по нескольким механизмам одновременно или с изменением условий диффузии может меняться и ее механизм. Наиболее вероятны следующие механизмы диффузии: междоузельный (перемещение атомов по междоузлиям), вакансионный (перемещение атомов по вакансиям), кольцевой или обменный (прямой обмен местами между атомами) и диссоциативный. Рассмотрим эти механизмы более подробно.

8.1.1. Междоузельный механизм

Процесс диффузии атомов в твердых растворах внедрения является простейшим и наиболее наглядным примером этого механизма диффузии. Сам механизм сводится к последовательному переходу атомов из одного междоузлия в другое. В разбавленных твердых растворах можно считать перескоки атома из одного междоузлия в другое независимыми друг от друга. Скорость диффузии атомов небольшого размера по междоузлиям

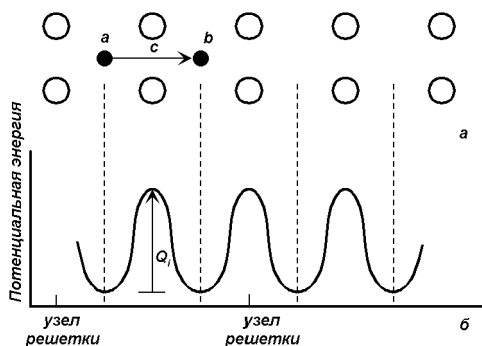


Рис. 8.1. Диффузия по междоузлиям: *а* — схема диффузии; *б* — потенциальная энергия диффундирующего атома в зависимости от его положения в кристаллической решетке.

в твердом теле определяется энергией, необходимой для передвижения атомов внедрения по кристаллической решетке.

Элементарный акт диффузии в этом случае показан на рис. 8.1,*а*. Поскольку в междоузлиях *а* и *б* атомы занимают положение равновесия, эти междоузлия соответствуют минимуму энергии системы. Из соображений симметрии следует, что максимум энергии достигается в точке *с*, посередине между положениями равновесия. Зависимость потенциальной энергии атома от его положения показана на рис. 8.1,*б*. Увеличение энергии диффундирующего атома Q_i при его перемещении в среднее положение вызвано главным образом появлением упругих напряжений, возникающих при его «протискивании» в узком пространстве между атомами решетки, а также происходит за счет частичного разрыва связей. Для твердых тел величина Q_i обычно составляет порядка 1 эВ. Поскольку средняя тепловая энергия атома при умеренных температурах не превышает 0.1 эВ, то для преодоления энергетического барьера Q_i необходима большая флуктуация энергии. Общая частота прыжков, совершаемых с исходной позиции, согласно формуле (8.7) равна

$$\dot{f} = Z\nu \exp(-Q_i/kT). \quad (8.8)$$

Как уже упоминалось, по междоузельному механизму диффундируют примеси внедрения. Например, по такому механизму происходит диффузия Li в Ge. В этом случае диффузия лимитируется только энергией активации миграции примеси Q_i .

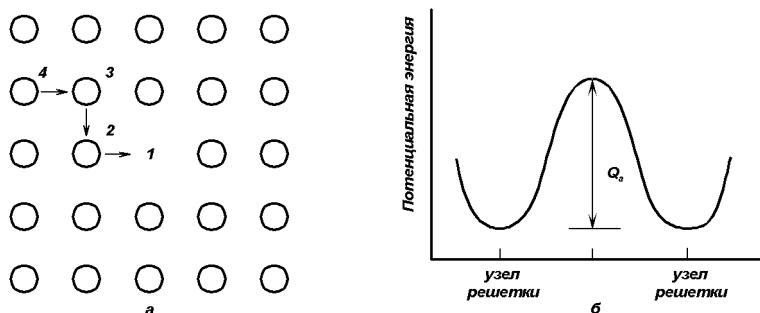


Рис. 8.2. Диффузия за счет движения вакансий: *а* — схема диффузии; *б* — зависимость потенциальной энергии диффундирующего атома от его положения в кристаллической решетке.

8.1.2. Вакансионный механизм

При наличии вакансии в объеме или в поверхностном слое решетки какой-либо из соседних с ней атомов может скачком занять ее место. Это равносильно тому, что вакансия скачком переместится на место атома. Многократное повторение такого акта и будет представлять собой миграцию вакансий и соответственно диффузию атомов в обратном направлении. Однако даже при переходе атома из своего узла в соседнюю вакансию он должен преодолеть определенный энергетический барьер. Переход связан с необходимостью частичного разрыва связей с атомами, соседними в исходном состоянии, и с необходимостью упругого смещения атомов, соседних со стороны вакансии.

На рис. 8.2 показано, как будет изменяться потенциальная энергия атома в решетке в зависимости от положения диффундирующего атома. Энергетический барьер Q_a , который необходимо преодолеть атому для его перехода в соседнюю вакансию, представляет собой энергию активации миграции. Для большинства твердых тел Q_a составляет величину порядка 1 эВ. Относительная доля времени, в течение которого атом обладает энергией, достаточной для преодоления барьера, пропорциональна $\exp(-Q_a/kT)$.

Однако изначально предполагалось, что вакансия уже имеется в решетке. В общем же случае для диффузии по вакансионному механизму необходимо еще затратить энергию Q_v на образование вакансии. Значение Q_v , как правило, немного больше Q_a .

Таким образом, суммарная энергия активации диффузии по ваканси-

онному механизму $Q = Q_a + Q_v$, а частота перескоков

$$f = Z\nu \exp(-Q_a/kT) \exp(-Q_v/kT). \quad (8.9)$$

Вакансионный механизм является основным механизмом самодиффузии и диффузии примесей замещения. По такому механизму, например, происходит диффузия In в Ge и Si.

Следует различать скорость самодиффузии по вакансиям и скорость диффузии самих вакансий. Действительно, по соседству с каждой диффундирующей вакансией всегда имеется атом, с которым она может поменяться местами, тогда как далеко не у каждого диффундирующего атома рядом имеется вакансия. Вероятность того, что вакансия окажется рядом с диффундирующим атомом, пропорциональна концентрации вакансий. Отсюда следует, что коэффициент самодиффузии D_{cd} и коэффициент диффузии вакансий D_v различны, но связаны между собой следующим соотношением:

$$D_{cd} = D_v C_v / C_a, \quad (8.10)$$

где C_v — концентрация вакансий; C_a — концентрация узлов, занятых атомами.

Важнейшим следствием, вытекающим из приведенной зависимости, является высокая чувствительность скорости диффузии по вакансионному механизму к наличию собственных дефектов. Наличие дефектов облегчает установление равновесной концентрации вакансий, способствует ускорению диффузии и уменьшению ее энергии активации. По этой причине в поликристаллических образцах диффузия примесей происходит в основном по межкристаллитным границам.

8.1.3. Обменный механизм

На рис. 8.3,а схематически показана диффузия по кольцевому механизму. В ходе такой диффузии в едином активационном процессе соседние атомы 1, 2, 3 и 4 обмениваются местами.

Этот механизм энергетически наименее выгоден для плотноупакованных решеток. Среди случаев обменного механизма наименее вероятен парный обмен (рис. 8.3,б), так как в этом случае обменивающиеся местами атомы создают очень сильные локальные искажения в решетке в переходных состояниях (атомы, соседние с теми, которые обмениваются местами, должны раздвинуться на расстояния, равные двум атомным диаметрам). Расчеты показали, что потенциальный барьер при диффузии по кольцу из четырех атомов меньше, чем при парном обмене, так

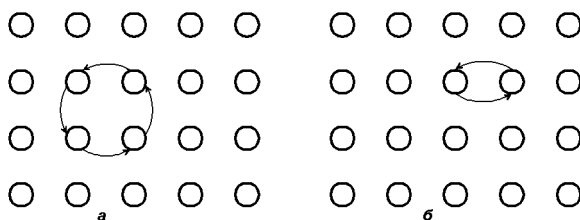


Рис. 8.3. Диффузия по кольцевому механизму (а) и парно-обменному механизму (б).

как в первом случае отсутствует встречное движение атомов, вызывающее наибольшую локальную деформацию решетки. Это дало основание для предположения, что кольцевой механизм вполне вероятен для некомпактных решеток типа алмаза, особенно когда обменная диффузия нескольких атомов осуществляется не в едином активационном процессе, а последовательно, через образование промежуточных дефектов Френкеля.

До сих пор мы обсуждали лишь природу элементарных атомных прыжков и не касались вопроса о перемещении атомов на большие расстояния. Эти перемещения — результат множества прыжков. Рассчитаем теперь смещение атома после того, как он совершил определенное число прыжков. Пусть все прыжки одинаковы по длине (равны межатомному расстоянию) и происходят в решетке с высокой степенью симметрии. Тем не менее предполагается, что движение атома в различных возможных кристаллографических направлениях равновероятно. Атомы могут перескакивать вперед, назад, вверх или вниз. Поэтому нельзя предсказать, какую результирующую траекторию совершит *отдельный* атом после определенного количества прыжков. С достаточной точностью можно определить только смещение, усредненное по *многим* диффундирующим атомам.

Для простоты рассмотрим диффузию в одном направлении. Пусть длина каждого скачка равна δ . При этом среднеквадратичное расстояние $\overline{x^2}$ (среднее из значений x^2 для множества атомов, каждый из которых совершил N прыжков), являющееся мерой общего расстояния, пройденного атомами от начала отсчета, не будет равно 0.

Действительно, для каждого атома

$$x^2 = (\delta_1 + \delta_2 + \dots + \delta_N)(\delta_1 + \delta_2 + \dots + \delta_N) = \delta_1^2 + \delta_2^2 + \dots + \delta_N^2 + 2\delta_1\delta_2 + \dots + 2\delta_{N-1}\delta_N, \quad (8.11)$$

где δ_i — длина i -го прыжка. Если отдельные скачки не коррелированы, то $\langle \delta_i \delta_j \rangle = 0$, и тогда

$$\overline{x^2} = N\delta^2 \quad \text{и} \quad \sqrt{\overline{x^2}} = \sqrt{N}\delta^2. \quad (8.12)$$

Число скачков $N = ft$, где t — время, необходимое для свершения N атомных прыжков, откуда $\sqrt{\overline{x^2}} = \sqrt{ft}\delta^2$. Учитывая (в рамках кинетической теории газов [40, 41]), что для трехмерной диффузии $D = f\delta^2/6$, а для одномерной диффузии $D = f\delta^2/2$, получим соотношение

$$\sqrt{\overline{x^2}} = L = \sqrt{2Dt}, \quad (8.13)$$

где L — диффузионная длина, t — время диффузии, D — коэффициент диффузии.

8.2. Основные параметры диффузии

Чтобы установить зависимость коэффициента диффузии D от температуры, можно использовать уравнение (8.7) для частоты прыжков атома $f = Z\nu \exp(-Q/kT)$. Тогда с учетом того, что $D = f\delta^2/6$ выражение для коэффициента диффузии примет вид:

$$D = (Z/6) \nu \delta^2 \exp(-Q/kT). \quad (8.14)$$

Обозначая $\nu \delta^2$ через D_0 и принимая $Z \approx 6$ (что справедливо для кубических кристаллов), получим формулу, описывающую температурную зависимость коэффициента диффузии,

$$D = D_0 \exp(-Q/kT), \quad (8.15)$$

где Q — энергия активации, а D_0 — предэкспоненциальный множитель, часто называемый частотным фактором.

Обсудим некоторые особенности этого уравнения. Во-первых, из него следует, что коэффициент диффузии D экспоненциально зависит от температуры, если преобладает один механизм диффузии. Многочисленными экспериментами установлено, что такая температурная зависимость имеет место для большинства твердых тел. Кроме того, из уравнения (8.15) следует, что энергию активации можно определить графически или аналитически по найденным из эксперимента значениям D для нескольких температур. Необходимым условием корректности использования формулы (8.15) является независимость значений D_0 и Q

от температуры в исследуемом интервале температур. Если зависимость $\ln D = f(1/T)$ оказывается линейной (рис. 8.4, кривая 1), то требуемое условие выполняется. Определив тангенс угла наклона прямой, можно найти энергию активации диффузионного процесса. Если точки ложатся на ломанную линию (рис. 8.4, кривая 2), то это означает, что по каким-либо причинам механизмы диффузии различны в разных температурных интервалах и для каждого из них должно быть найдено свое значение энергии активации. Точки могут ложиться и на кривую линию (рис. 8.4, кривая 3). Это, как правило, соответствует случаю наложения нескольких механизмов диффузии, роль каждого из которых существенно зависит от температуры. Такой случай требует специального анализа, а найденные для отдельных участков значения Q представляют собой *эффективные энергии активации*.

Во-вторых, из уравнения (8.15) вытекает, что величина D_0 связана с частотой атомных скачков (или с частотой колебаний решетки $\nu \sim 10^{13}$ Гц) и величиной скачка атома примеси. Кроме того, из уравнения (8.15) следует, что в области температур, где диффузия идет с заметной скоростью, коэффициент D_0 должен быть постоянным, не зависеть от температуры и по порядку величины составлять от 10^{-4} см²/с до 10^{-3} см²/с. В ряде случаев эксперимент дает близкие значения D_0 , но чаще они отличаются от расчетных на несколько порядков. Причины этого несоответствия до конца не ясны до сих пор, однако существует ряд предположений, с помощью которых делались попытки объяснить это расхождение [41]. Например, предполагалось, что на величине D_0 может сказаться изменение механизма диффузии в том или ином температурном интервале и соответственно изменение величины Q , то есть появление температурной зависимости Q . Кроме того, существует предположение, что величина D_0 может изменяться с изменением энергии образования вакансий. С повышением температуры уменьшается энергия образования вакансий за счет увеличения расстояния между атомами. В результате Q может убывать с ростом температуры в первом приближении по закону, близкому к линейному: $Q = Q_0 - \gamma T$, где Q_0 — энергия активации, экстраполированная к нулю абсолютной температуры, а γ — численный коэффициент, связанный с ангармоничностью колебаний атомов твердого тела, которая в свою очередь зависит от их массы. Рядом авторов было показано, что элементарный акт диффузии изменяет термодинамический потенциал некоторой области кристаллической решетки. Поэтому диффузия сопровождается не только изменением энергии системы за счет преодоления энергетического барьера, но и изменением энтропии системы. Это также ведет к изменению предэкспоненциально-

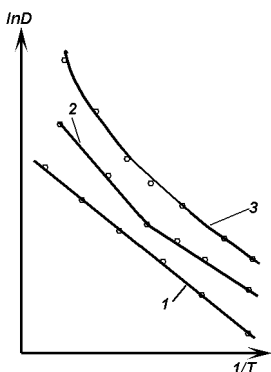


Рис. 8.4. Различные случаи зависимости $\ln D = f(1/T)$.

го множителя в формуле (8.15). Следует отметить, что в предыдущих рассуждениях не учитывалось, что диффундирующий по вакансиям атом может возвращаться в только что покинутую вакансию. Учет этой вероятности приводит к необходимости введения в значение D_0 корреляционного множителя.

8.3. Распределение диффундирующих атомов по глубине кристалла

Рассмотрим два частных, но важных для технологии полупроводников и получения $p - n$ -переходов в полупроводниковых приборах случая решения уравнения одномерной диффузии (8.6).

8.3.1. Диффузия из бесконечно тонкого слоя (из ограниченного источника)

Диффузия из бесконечно тонкого слоя соответствует ситуации, когда небольшое количество легирующего вещества осаждается на поверхности легируемого материала. В этом случае все диффундирующие в течение некоторого времени t с поверхности в объем атомы полностью переходят в кристалл, причем распределяются моноатомно. В любой момент времени полное количество диффундирующих атомов остается постоянным.

Сначала рассмотрим диффузию из конечного слоя толщины h в полуограниченное тело. Под полуограниченным телом будем понимать такое тело, которое с одной стороны ограничено плоскостью $x = 0$, а с другой

стороны простирается в бесконечность. Будем считать тело изотропным, а коэффициент диффузии не зависящим от концентрации.

Распределение атомов по глубине кристалла при таких условиях описывается уравнением

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}, \quad (8.16)$$

с начальным распределением концентрации

$$C(x, 0) = C_0 \text{ для } 0 < x < h \text{ и } C(x, 0) = 0 \text{ для } h < x < \infty, \quad (8.17)$$

где C_0 — начальное распределение концентрации, h — толщина слоя.

Решение уравнения (8.16) (решается методом разделения переменных) имеет вид:

$$C(x, t) = \frac{C_0}{2} \left[\operatorname{erf} \left(\frac{h-x}{2\sqrt{Dt}} \right) + \operatorname{erf} \left(\frac{h+x}{2\sqrt{Dt}} \right) \right], \quad (8.18)$$

где $\operatorname{erf}(Y) = \frac{2}{\pi} \int_0^Y e^{-z^2} dz$ — функция Гаусса, а $Y = \frac{h \pm x}{2\sqrt{Dt}}$. Устремляя в (8.18) h к нулю, получим формулу для диффузии из бесконечно тонкого слоя в полуограниченное тело. Учитывая, что $C_0 h = S = \text{const}$, где S — количество вещества в слое, отнесенное к единице площади слоя (постоянное общее число атомов на единицу площади), получим

$$\begin{aligned} \lim_{h \rightarrow 0} C(x, t) &= \frac{S}{2} \lim_{h \rightarrow 0} \frac{1}{h} \left[\operatorname{erf} \left(\frac{h-x}{2\sqrt{Dt}} \right) + \operatorname{erf} \left(\frac{h+x}{2\sqrt{Dt}} \right) \right] \\ &= \frac{S}{\sqrt{\pi Dt}} \exp \left(-\frac{x^2}{4Dt} \right). \end{aligned} \quad (8.19)$$

Полагая $x = 0$, получим величину поверхностной концентрации примесных атомов

$$C_s = C(0, t) = \frac{S}{\sqrt{\pi Dt}}. \quad (8.20)$$

По мере удаления от поверхности кристалла (источника) концентрация диффундирующих атомов C падает, стремясь к 0 (рис. 8.5). Максимальное значение C достигается при $x = 0$. Со временем начальное распределение постепенно расплывается, величина максимума кривой распределения понижается пропорционально корню квадратному из времени. Таким образом, для диффузии из бесконечно тонкого слоя характерно уменьшение со временем количества диффундирующих атомов у поверхности за счет проникновения их в глубь кристалла.

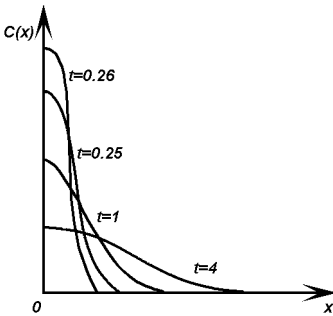


Рис. 8.5. Диффузия из бесконечно тонкого слоя в полуограниченное тело.

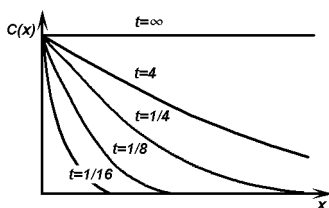
Следует отметить, что так как поток вещества через плоскость $x = 0$ отсутствует, то в любой момент времени $t > 0$ диффундирующее вещество не поступает в тело извне, а целиком уже находится в нем. Однако в реальных системах растворимость вещества, как правило, ограничена и вещество продолжает поступать в диффузионную зону и при $t = 0$, то есть нарушается условие $j(x, t)(x = 0) = 0$ и количество вещества S в диффузионной зоне меняется со временем, а не остается постоянным. Поток вещества через плоскость $x = 0$ играет роль в основном в первые моменты диффузии. С течением времени реальный источник вещества постепенно иссякает и распределение концентраций все более приближается к распределению, описываемому формулой (8.19).

8.3.2. Диффузия из постоянного источника

Термин диффузия из постоянного источника означает, что в полуограниченное тело через плоскость $x = 0$ проникает извне диффундирующее вещество, причем на границе тела поддерживается постоянная, не зависящая от времени поверхностная концентрация C_s этого вещества, то есть запас атомов диффундирующей примеси на поверхности или в питающей паровой фазе, контактирующей с ней, настолько велик, что не меняется в процессе диффузии. Тогда начальные и граничные условия для уравнения (8.6) в этом случае будут следующие: $C(x, 0) = 0$; $C(0, t) = C_s$; $C(\infty, t) = 0$. Решение уравнения (8.16) в этом случае имеет вид:

$$C(x, t) = C_s[1 - \operatorname{erf}(Y)], \text{ где } \operatorname{erf}(Y) = \frac{2}{\pi} \int_0^Y e^{-z^2} dz \quad (8.21)$$

Рис. 8.6. Диффузия из постоянного источника в полуограниченное тело.



— функция Гаусса, $Y = x/(2\sqrt{Dt})$, а C_s — некоторая постоянная поверхностная концентрация вещества (см^{-3}).¹

Соответствующее распределение примесных атомов по глубине кристалла показано на рис. 8.6. Видно, что поток диффузии максимален в первые моменты времени и обращается в нуль при $t \rightarrow \infty$. Для таких условий диффузии характерен рост концентрации примеси в приповерхностном слое со временем и стремление ее к постоянной величине $C(x, t) = C_{im}$.

8.4. Экспериментальное определение коэффициентов диффузии

Для изучения диффузии в твердых телах в настоящее время разработано большое число физических и физико-химических методов исследования. Эти методы основаны на измерении распределения концентрации диффундирующего вещества в исследуемом образце в зависимости от времени и температуры диффузионного отжига. При этом распределение концентрации определяется либо прямыми измерениями ее в различных частях исследуемого образца с помощью химических, спектроскопических, рентгеновских, электронографических, радиоактивных и других методов анализа, либо же косвенным образом — посредством изучения характера изменений некоторых физических свойств вещества, вызванных проникновением диффундирующей примеси.

Широкое развитие получили методы изучения диффузионных процессов, основанные на использовании радиоактивных изотопов. Им присуща высокая чувствительность, универсальность и сравнительная простота. В настоящее время наиболее перспективным методом измерения диффузионных профилей является масс-спектрометрия вторичных ионов

¹С физической точки зрения C_s не может превышать величины растворимости диффундирующего вещества в растворителе, то есть $C_{smax} = C_{im}$, где C_{im} — предельная растворимость примеси в данном веществе при температуре диффузии.

(МСВИ). Этот метод обеспечивает измерение низкой концентрации примесей (например, $\sim 5 \cdot 10^{15} \text{ см}^{-3}$ бора и мышьяка в кремнии), обладает высоким разрешением по глубине (несколько нанометров), универсальностью. Метод МСВИ наряду с радиоактивным методом позволяет определить общее количество введенной примеси, поэтому, если необходимо определить электрически активную часть примеси, то следует воспользоваться электрическими методами. К настоящему времени разработаны и широко используются специфические полупроводниковые методы измерения диффузионных профилей электрически активных примесей (или их электрически активной части) и их коэффициентов диффузии в полупроводниках. Эти методы основаны главным образом на исследовании изменений электрических свойств в различных частях полупроводникового образца, обусловленных проникновением туда диффундирующих атомов. Эти методы не столь универсальны, как радиоактивные и МСВИ, но их преимуществом является незначительная трудоемкость и отсутствие специфики, связанной с применением радиоактивных изотопов. Однако следует иметь в виду, что практическое применение полупроводниковых методов исследования возможно только при использовании материалов высокой химической чистоты. Кроме того, диффундирующее вещество должно быть электрически активной примесью и оказывать влияние на электрические свойства исследуемого полупроводника. К полупроводниковым методам относятся метод электронно-дырочного перехода, метод фото-э.д.с., метод электропроводности и емкостный метод [39, 41]. Мы ниже рассмотрим лишь один из них — метод $p - n$ -перехода, позволяющий непосредственно определять концентрацию даже при очень малых глубинах проникновения примесей, когда измерение эффекта Холла невозможно.

Метод электронно-дырочного перехода

Известно, что характерной особенностью полупроводников является то, что введением донорных или акцепторных примесей можно придавать полупроводниковому материалу электронный или дырочный характер проводимости. Введение соответствующих примесей (акцепторных в случае полупроводников n -типа и донорных — в случае полупроводников p -типа) в поверхностный слой образца приводит к возникновению достаточно узких переходных областей, где характер проводимости вещества изменяется с электронного на дырочный или наоборот. Эти переходные области ($p - n$ -переходы) при диффузионном отжиге перемещаются в направлении градиента концентрации диффундирующей примеси. При этом

скорость перемещения перехода будет определяться скоростью диффузии примеси. Поэтому, исследуя скорость перемещения границы электронно-дырочного перехода в зависимости от температуры и времени диффузионного отжига, можно по полученным данным вычислить коэффициент диффузии. Глубина залегания $p-n$ -перехода и его форма (резкий или размытый переход) при заданной температуре и времени диффузионного отжига в каждом случае определяются начальными и граничными условиями опыта (диффузия из тонкого слоя, нанесенного на поверхность образца; диффузия из постоянного источника и т. д.).

Пусть в исходный материал, например n -типа, с известной концентрацией, равномерно распределенной по объему донорной примеси C_d , диффузией при температуре T в течение времени t вводится акцепторная примесь. Затем кристалл достаточно быстро охлаждается, поверхность его сошлифовывается под малым углом α (рис. 8.7) и с помощью точечного зонда исследуется вид вольт-амперной характеристики или определяется знак термо-э.д.с. вдоль сошлифованной поверхности. Если диффундирующее вещество электрически активно и каждый атом, проникший в исследуемый образец, создает там электрон или дырку, а также известны начальные и граничные условия опыта, то можно найти характер распределения концентрации введенных носителей тока в исследуемом образце. В образце $p-n$ -переход возникает в том месте, где концентрация носителей тока, вносимых примесными атомами, становится равной концентрации противоположных по знаку заряда исходных носителей тока в исследуемом образце $C_d = C_a(x, t)$. При диффузии из ограниченного источника ($C_d < S/\sqrt{\pi Dt}$) и при диффузии из постоянного источника ($C_d < C_{as}$, C_{as} — растворимость акцепторной примеси в данном веществе при температуре диффузии) искомый коэффициент диффузии D может быть рассчитан согласно уравнению (8.19) или уравнению (8.21) соответственно, если точно известна глубина залегания $p-n$ -перехода $x = b \sin \alpha$ — расстояние до $p-n$ -перехода от граничной поверхности образца. Величина C_d определяется из измерений постоянной Холла R_H перед диффузионным отжигом образца. При использовании уравнения (8.19) необходимо знать S — количество вещества в слое, приходящееся на единицу площади, а при использовании уравнения (8.21) — растворимость C_{is} примеси при данной температуре. Если эти величины неизвестны или трудно определяемы, то их можно исключить из рассмотрения, воспользовавшись *методом «двух образцов»*.

Сущность этого метода заключается в одновременном определении границ $p-n$ -перехода в двух образцах с разными начальными значениями концентрации носителей тока C_{d1} и C_{d2} (предполагается, что подвиж-

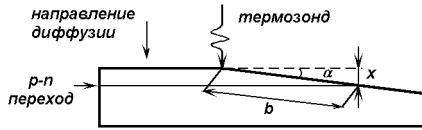


Рис. 8.7. Схема метода электронно-дырочного перехода для определения коэффициента диффузии.

ность носителей тока в обоих образцах одинакова). После проведения одновременной диффузии в течение времени t $p-n$ -переходы в двух образцах будут расположены на разных глубинах x_1 и x_2 . Из уравнения (8.19) получим

$$D = \frac{x_2^2 - x_1^2}{4t \ln(C_{d1}/C_{d2})}, \quad (8.22)$$

а из уравнения (8.21) получим

$$\frac{C_{d1}}{C_{d2}} = \frac{1 - \operatorname{erf}(Y_1)}{1 - \operatorname{erf}(Y_2)}, \quad (8.23)$$

где $Y_i = x_i/(2\sqrt{Dt})$.

При особо точных измерениях коэффициентов диффузии следует иметь в виду следующее обстоятельство. При определении границы $p-n$ -перехода с помощью термозонда полученное значение x , как правило, несколько отличается от действительного значения, при котором концентрация носителей тока, обусловленных диффузией примесных атомов, в точности равна концентрации исходных носителей тока в образце. Такое различие обусловлено тем, что при наличии носителей двух знаков (что имеет место в переходной области) коэффициент термо-э.д.с. α в полупроводниках зависит не только от концентрации, но также от подвижностей и эффективных масс носителей тока [6,41]. Поскольку различие в эффективных массах мало сказывается на α (из-за логарифмической зависимости), то равенство нулю коэффициента термо-э.д.с. определяется условием: $n\mu_n = p\mu_p$, где n и p , μ_n и μ_p — концентрации и подвижности электронов и дырок соответственно. Поэтому для точного определения D величину C_d в уравнениях (8.19) и (8.21) следует умножить на отношение подвижности вводимых носителей тока к подвижности основных носителей тока.

При измерении коэффициентов диффузии методом электрон-дырочных переходов также следует обращать особое внимание на наличие

освещения поверхности образца, так как последнее приводит к появлению фото-э.д.с., которая, накладываясь, например на измеряемую термо-э.д.с., часто совершенно искажает результаты измерений и приводит к ошибочным выводам.

8.5. Диффузия в германии и кремнии

К настоящему времени наиболее полные данные о коэффициентах диффузии различных примесей получены для германия и кремния. Первоначально основное внимание при исследовании диффузии было сосредоточено на определении коэффициентов диффузии тех примесей, которые существенно изменяли электрические свойства этих полупроводников и приводили к возникновению $p - n$ -переходов. К таким примесям прежде всего относятся элементы IIIA и VA подгрупп таблицы Менделеева, которые образуют с германием и кремнием, как правило, твердые растворы замещения и создают в запрещенной зоне «мелкие» водородоподобные энергетические уровни. Эти примеси сравнительно легко ионизируются, поэтому в широком интервале температур являются основными источниками носителей тока; при этом приводимость Ge и Si изменяется в большом диапазоне (на несколько порядков).

Дальнейшие исследования были сосредоточены на изучении диффузии Li, Cu, Au, Ag, Zn, Ni и Fe. Эти элементы, за исключением лития, также образуют с германием и кремнием, как правило, твердые растворы замещения, но создают в запрещенной зоне глуболежащие многозарядные, как правило, акцепторные уровни, число которых соответствует разности между числом валентных электронов германия (кремния) и атома примеси (см. гл. 3). Такие «глубокие» уровни выступают как уровни захвата или рекомбинации носителей тока и играют большую роль в неравновесных процессах, уменьшая время жизни и ускоряя рекомбинацию неосновных носителей тока в полупроводниках. Литий с германием и кремнием образует твердый раствор внедрения и является донором с одним уровнем.

В табл. 8.1 приведены данные исследований температурных зависимостей коэффициента диффузии D для различных примесей в Ge, которые практически все удовлетворительно описываются формулой $D = D_0 \exp(-Q/kT)$.

Из таблицы видно сильное различие в значениях коэффициентов диффузии D и параметров D_0 и Q у элементов IIIA–VA групп и некоторых элементов из II и VI групп, с одной стороны; и у элементов IB, VIII (переходные металлы) групп таблицы Менделеева и Li, с другой сто-

Таблица 8.1. Диффузионные параметры различных примесей в Ge [19,24,41].

Примесь	Группа	D_0 , см ² /с	Q , эВ	T , °C	D , см ² /с
Na	I	0.07	2.08	700–850	$1.2 \cdot 10^{-11}$
Be	II	0.5	2.5	720–920	$9.1 \cdot 10^{-13}$
Zn	II	0.65	2.55	825–918	$6.9 \cdot 10^{-13}$
Cd	II	$1.75 \cdot 10^9$	4.42	760–915	$3.0 \cdot 10^{-12}$
B	III	$6.0 \cdot 10^8$	45	760–850	$4.4 \cdot 10^{-13}$
Al	III	$1.0 \cdot 10^3$	3.45	554–905	$6.2 \cdot 10^{-14}$
Ga	III	10	30	650–900	$8.1 \cdot 10^{-14}$
In	III	$1.8 \cdot 10^4$	3.64	554–919	$1.6 \cdot 10^{-13}$
Tl	III	$1.7 \cdot 10^3$	3.4	800–920	$1.8 \cdot 10^{-13}$
Ge	IV	13.6	3.09	535–904	$4.2 \cdot 10^{-14}$
Sn	IV	70	3.05	798–846	$5.7 \cdot 10^{-13}$
Pb	IV	$1.6 \cdot 10^3$	3.6	800	$2.0 \cdot 10^{-14}$
P	V	1.2	24	650–900	$6.4 \cdot 10^{-12}$
As	V	3	2.43	750–900	$1.2 \cdot 10^{-11}$
Sb	V	3.5	2.45	750–850	$1.1 \cdot 10^{-11}$
Bi	V	3.3	2.5	750–850	$6.0 \cdot 10^{-12}$
O	VI	0.4	2.08	285–760	$6.8 \cdot 10^{-11}$
Te	VI	5.6	2.43	770–900	$2.2 \cdot 10^{-11}$
Li	I	$1.3 \cdot 10^{-3}$	0.46	350–800	$9.0 \cdot 10^{-6}$
Fe	VIII	0.13	1.1	775–930	$8.9 \cdot 10^{-7}$
Ni	VIII	0.8	0.91	700–875	$4.3 \cdot 10^{-5}$
Co	VIII	0.16	1.12	750–850	$8.8 \cdot 10^{-7}$
Cu	I	$1.9 \cdot 10^{-4}$	0.18	700–900	$2.8 \cdot 10^{-5}$
Ag	I	$7.4 \cdot 10^{-2}$	1.0	730–900	$8.8 \cdot 10^{-7}$
Au	I	$3.5 \cdot 10^{-6}$	0.63	800–900	$3.8 \cdot 10^{-9}$
H	I	$2.72 \cdot 10^{-3}$	0.38	780–930	$4.5 \cdot 10^{-5}$
He	0	$6.1 \cdot 10^{-3}$	0.69	795–872	$3.5 \cdot 10^{-6}$
Xe	0	$5 \cdot 10^{-6}$	1.2	250–980	$1.2 \cdot 10^{-11}$

роны. Первая группа примесей имеет D на несколько порядков меньше («медленно» диффундирующие примеси), чем вторая группа («быстро» диффундирующие примеси); энергии активации диффузии Q у первой группы больше, чем у второй; D_0 тоже больше в первой группе, чем во второй. Однако из таблицы видно, что существует большой разброс в значениях D_0 и Q в самих группах. Здесь следует иметь в виду, что интервал надежного определения параметров диффузии, особенно для примесей первой группы, довольно узок из-за их малых глубин проникновения в кристалл. Поэтому к приведенным значениям D_0 , полученным экстраполяцией, следует относиться как к оцененным с точностью до порядка. Тем не менее полученные данные позволяют сделать вывод о различных механизмах диффузии для этих двух групп примесей. Аналогичная ситуация наблюдается и при исследовании диффузии различных примесей в Si [39, 41].

Так как примеси первой группы образуют с германием твердые растворы замещения, то можно полагать, что и диффузия в них носит вакансионный характер. В этом случае Q можно отождествить с соответствующей энергетической характеристикой $Q_a + Q_v$ в формуле (8.9).

Диффузию примесей второй группы можно трактовать как междоузельную. С этим полностью согласуются экспериментальные данные по Li. Он образует с германием твердый раствор внедрения и в кристаллической решетке является донором, создающим электронные уровни, расположенные на 0.01 эВ ниже дна зоны проводимости. Действительно, Li имеет наименьший по сравнению с другими элементами ионный радиус $r_{Li^+} = 0.68 \text{ \AA}$, поэтому ионизованный атом Li^+ может свободно перемещаться по междоузлиям решетки Ge. Относительно других «глубоких» примесей второй группы (Cu, Ni, Fe, Au и т. д.) вопрос о междоузельном механизме диффузии не столь очевиден. Эти примеси являются, как правило, акцепторами в Ge, а такой характер их электрической активности возможен только в том случае, когда эти примеси расположены в узлах решетки. Предположение о их междоузельной диффузии может быть подтверждено лишь в том случае, если будет экспериментально доказано, что диффундируют они не в виде ионизованных акцепторов A^- , а в виде ионизованных доноров D^+ . Для этого необходимо определить зарядовое состояние диффундирующих примесных ионов.

8.6. Определение зарядового состояния диффундирующих примесных ионов

8.6.1. Диффузия Li в Ge

Одним из методов определения зарядового состояния быстро диффундирующих примесных ионов является наблюдение их дрейфа в электрическом поле. Впервые такой эксперимент был проведен на литии в германии. Суть его состоит в следующем (рис. 8.8). Диффундирующая примесь наносится на поверхность германия p -типа проводимости кратковременным вплавлением ее в поверхностный слой. При этом реализуется случай «точечного» источника с неограниченным запасом примесных атомов — капля, вплавленная в кристалл и имеющая радиус много меньший характерных расстояний диффузии. Далее образец прогревается при заданной температуре T время t_1 для формирования четкого фронта диффузии, представляющего собой полусферу радиуса r_1 . Затем образец охлаждается до комнатной температуры, а исходный источник примеси удаляется шлифовкой и специальным травлением. После травления на поверхности образца остается лунка, концентрично с которой находится диффузионная область, обогащенная литием, которая имеет проводимость n -типа. Затем определяется положение $p-n$ -перехода, располагающегося на поверхности полусферы радиуса r_1 , с которой в дальнейшем пойдет диффузия. Граница $p-n$ -перехода выявляется, например, химическим окрашиванием в специальном красителе² или электрическим осаждением титаната бария. Затем образец помещается в постоянное электрическое поле (напряженностью 1–10 В/см), а диффузию проводят при той же температуре T , что и первый раз. Ток, который пропускается через образец (1–10 А), одновременно используется для его нагрева. Ввиду отсутствия источника дальнейшая диффузия примеси происходит аналогично рассмотренному выше случаю диффузии из ограниченного источника (уравнение (8.19)), то есть радиус полусферы увеличивается за счет обеднения областей прилегающих к бывшему источнику. Одновременно все диффундирующие ионы в соответствии со своим знаком заряда q будут дрейфовать в электрическом поле со скоростью $V_{др.} = \mu E$, где μ — эффективная подвижность ионов, связанная с их коэффициентом диффузии соотношением Эйнштейна $\mu = (q/kT)D$. Таким образом, центр полусферы после соответствующего прогрева переместится в новое

²Если образец, в котором сформирован $p-n$ -переход, погрузить в специальный краситель и выдержать на свету несколько минут, то область p -типа проводимости будет иметь цвет, отличный от цвета n -типа проводимости.

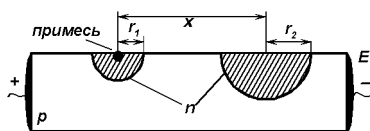


Рис. 8.8. Схема метода определения заряда диффундирующих ионов (x — смещение в электрическом поле; r_1 (без поля) и r_2 (с полем) — радиусы p – n -переходов после прогрева).

положение, отстоящее от исходного на расстояние $x = \mu Et$.

В случае лития смещение пятна в электрическом поле оказалось направленным от «+» к «–» (от анода к катоду), то есть полностью подтвердило знак заряда ионизованного донора Li^+ . Кроме того, опыты с литием прямо доказали справедливость соотношения Эйнштейна для процессов диффузии и дрейфа ионов, а не только для электронов и дырок.

8.6.2. Диффузия Cu в Ge. Диссоциативный механизм диффузии

Результаты исследования диффузии меди в германии, полученные в широком интервале температур (600–900°C), приведены на рис. 8.9. Видно, что механизмы диффузии меди в области высоких температур (750–900°C) и в области сравнительно низких температур (600–700°C) различаются. Можно предположить, что в первой области преобладает междоузельный механизм диффузии, а при температурах ниже 700°C — в основном вакансионный механизм. Это предположение подтвердили эксперименты по диффузии меди в Ge в электрическом поле, проведенные при разных температурах. Было показано, что в области температур 750–900°C медь в Ge в основном присутствует в виде положительных ионов. С другой стороны, акцепторные свойства меди в Ge при умеренных и комнатной температуре указывают, что при этих температурах медь в основном находится в виде отрицательных ионов.

Идея о двойственном механизме диффузии меди в германии была развита Мэзеном и Бренкманом [42]. Они исходили из предположения, что в германии всегда существует динамическое равновесие между медью в узлах Cu_y и медью в междоузлиях $\text{Cu}_{\text{мy}}$: $\text{Cu}_{\text{мy}} \rightleftharpoons \text{Cu}_y + \Delta Q$, где ΔQ — теплота обменной реакции. С повышением температуры реакция сдвигается в сторону увеличения $\text{Cu}_{\text{мy}}$, а с понижением температуры — в сторону увеличения Cu_y . Подтверждением этому предположению служило сравнение концентрации электрически активной (акцепторной) меди

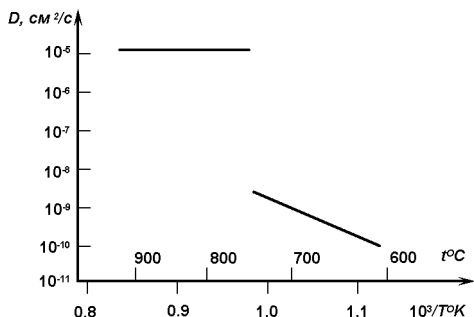


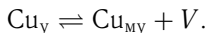
Рис. 8.9. Диффузия меди в германии.

Cu_y , измеренной электрическим методом, с общей концентрацией меди ($\text{Cu}_y + \text{Cu}_{\text{my}}$), определенной радиоактивными методами. Оказалось, что при температуре 600°C эти две концентрации почти равны, в то время как при 800°C общая концентрация меди почти в два раза больше концентрации электрически активной меди Cu_y .

Позднее большим числом исследований было установлено, что медь в германии относится к диссоциативным амфотерным примесям (см. гл. 3), то есть находится в решетке германия как в узлах, так и в междоузлиях в разных зарядовых состояниях, причем в сравнимых количествах. Концентрация меди в разных кристаллохимических позициях зависит от температуры. Каждое из двух возможных состояний атомов меди в решетке Ge — медь в междоузлиях и медь в узлах — характеризуется своим коэффициентом диффузии, причем естественно считать, что $D_{\text{my}} > D_y$, и в процессе диффузии происходит перераспределение между этими двумя состояниями меди. Представление о существовании динамического равновесия между медью в узлах и междоузлиях было использовано для интерпретации наблюдающейся зависимости скорости диффузии от степени совершенства кристаллов.

Действительно, накопленный экспериментальный материал о диффузии меди в германии позволил вполне обоснованно предположить, что быстрая диффузия меди осуществляется путем перемещения ее по междоузлиям кристаллической решетки германия в виде положительных ионов до встречи с вакансией. При встрече с вакансией медь переходит из междоузлия в узел. В узлах кристаллической решетки германия медь находится в виде отрицательных ионов. Тепловое движение может снова перебросить часть атомов меди из узлов в междоузлия (то есть происходит диссоциация узловых (акцепторной) меди Cu_y с образованием междоузельной (донорной) меди Cu_{my} и вакансии, которые затем диф-

фундируют отдельно с разной скоростью). Поэтому в решетке германия при постоянной температуре всегда существует динамическое равновесие



Очевидно, что чем больше вакансий в кристалле, тем больше смещается это равновесие в сторону преобладания меди в узлах. Концентрация вакансий в германии при заданной температуре зависит от степени совершенства кристалла.

В несовершенных кристаллах с высокой плотностью дислокаций быстро достигается высокая равновесная концентрация вакансий $[V]$, то есть можно записать

$$[\text{Cu}_y] = [V][\text{Cu}_{\text{my}}]/k(T) = \beta[\text{Cu}_{\text{my}}], \quad (8.24)$$

где $k(T)$ — константа равновесия, а $\beta = [V]/k$. Если диффузия определяется только междоузельной примесью, то ее поток выражается согласно уравнению (8.2) в виде:

$$M = -D_{\text{my}} \partial[\text{Cu}_{\text{my}}]/\partial x, \quad (8.25)$$

где D_{my} — коэффициент диффузии по междоузлиям. Эффективный коэффициент диффузии в несовершенном кристалле $D_{\text{н.к}}^*$ представляет собой отношение этого потока к общему градиенту концентрации примеси, то есть

$$D_{\text{н.к}}^* = -M \left(\frac{\partial([\text{Cu}_{\text{my}}] + [\text{Cu}_y])}{\partial x} \right)^{-1}, \quad (8.26)$$

откуда

$$D_{\text{н.к}}^* = D_{\text{my}} [\text{Cu}_{\text{my}}] / ([\text{Cu}_{\text{my}}] + [\text{Cu}_y]) = D_{\text{my}} / (1 + \beta). \quad (8.27)$$

Формула (8.27) отражает тот факт, что каждый ион меди находится в состоянии высокой диффузионной подвижности по междоузлиям лишь небольшую долю времени, соответствующую отношению $[\text{Cu}_{\text{my}}] / ([\text{Cu}_{\text{my}}] + [\text{Cu}_y])$. Эффективный коэффициент диффузии для этого случая лимитируется диффузией меди по междоузлиям.

В совершенных кристаллах с низкой плотностью дислокаций присутствие меди в узлах зависит от притока вакансий, поступающих в кристалл диффузией от поверхности. Концентрация вакансий будет меньше ее равновесного значения для данной температуры. В этом случае эффективный коэффициент диффузии в совершенном кристалле $D_{\text{с.к}}^*$ будет лимитироваться диссоциативным механизмом, при котором Cu_y переходит

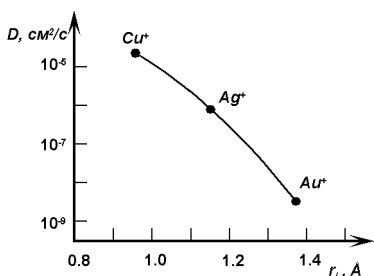


Рис. 8.10. Зависимость коэффициента диффузии D при 850°C от ионных радиусов быстро диффундирующих примесей.

в Cu_{My} , оставляя за собой вакансию. Далее Cu_{My} и вакансии диффундируют независимо. Эффективный коэффициент диффузии в этом случае определяется коэффициентом диффузии вакансий и временем, в течение которого вакансии находятся в состоянии диффузионной подвижности, и будет равен

$$D_{\text{с.к}}^* = D_v [V] / ([V] + [\text{Cu}_y]), \quad (8.28)$$

где D_v — коэффициент диффузии вакансий, $[V]$ — концентрация вакансий. Так как $D_v \ll D_{\text{My}}$ и $[V] / ([V] + [\text{Cu}_y]) < [\text{Cu}_{\text{My}}] / ([\text{Cu}_y] + [\text{Cu}_{\text{My}}])$, то $D_{\text{с.к}}^* < D_{\text{н.к}}^*$, то есть скорость диффузии в совершенных кристаллах должна быть значительно меньше, чем в кристаллах с высокой плотностью дислокаций, и определяется скоростью диффузии вакансий.

Опыты по определению зарядового состояния примеси были выполнены и с другими глубокими примесями в германии при температурах выше 700°C и дали те же результаты, то есть подтвердили донорный тип примесей в процессах высокотемпературной диффузии. Таким образом, по-видимому, в области высоких температур эти элементы диффундируют по междоузлиям в виде положительных ионов, а в области умеренных температур — по вакансиям в виде отрицательных ионов. На рис. 8.10 показана зависимость коэффициентов диффузии Cu, Ag и Au в зависимости от их ионных радиусов r_i . Как видно, с ростом r_i коэффициент диффузии уменьшается. С учетом того, что ковалентный радиус германия равен $1.22 \text{ \AA}$, становится понятной причина очень малого, по сравнению с другими примесями этой группы, коэффициента диффузии золота.

8.7. Влияние состава сплава и природы диффундирующего вещества на скорость и параметры диффузии

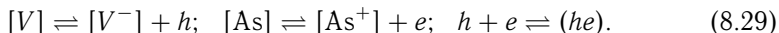
Диффузия в полупроводниковых материалах имеет ряд особенностей. Важнейшей из этих особенностей является наличие в полупроводниках электрически активных примесей и собственных дефектов, прежде всего вакансий. Кулоновское взаимодействие между ними изменяет подвижность, концентрацию и характер распределения дефектов и соответственно условия и скорость диффузии. Важно также, что влияние примесей в полупроводниках проявляется при весьма малых концентрациях. Кроме того, на процессы диффузии в полупроводниках сказывается и низкая компактность решеток последних.

8.7.1. Самодиффузия

Самодиффузия в элементарных веществах, соединениях и твердых растворах чаще всего осуществляется по вакансионному механизму. Скорость ее обычно меньше скорости гетеродиффузии.

На самодиффузию в полупроводниках влияют примеси. Характер влияния определяется не только размерным фактором $\Delta R/R$, где R — ковалентный радиус атомов основного вещества, а ΔR — разность ковалентных радиусов основного вещества и примеси (упругие напряжения ведут к перераспределению точечных дефектов). Существенную роль играет непосредственное влияние примесей на концентрацию вакансий. Коэффициент диффузии тем больше, чем выше концентрация вакансий (см. (8.10)). В полупроводниках между концентрацией электрически неактивных и электрически активных вакансий и примесей существует динамическое равновесие. Концентрации заряженных вакансий и примесей взаимосвязаны.

Действительно, в случае донорных примесей, например мышьяка в германии, равновесие и взаимодействие можно представить так:



Здесь he — электронно-дырочная пара. Образование этих пар сдвигает равновесие вправо, то есть увеличивает концентрацию акцепторных вакансий $[V^-]$; чтобы при этом сохранилась равновесная для данной температуры концентрация электрически неактивных вакансий, должна возрасти общая концентрация вакансий, что должно способствовать ускорению самодиффузии.

Ясно, что при введении акцепторной примеси картина будет обратной. Концентрация дырок возрастает за счет ионизации акцепторной примеси, уменьшается концентрация $[V^-]$, а с ней и общая концентрация вакансий, что в результате приводит к замедлению самодиффузии.

Особо следует остановиться на самодиффузии в полупроводниковых соединениях, которая осложняется наличием ионной составляющей связи, возможной диссоциацией некоторых соединений при температуре диффузии или испарением летучих компонент.

В полупроводниковых соединениях самодиффузия каждого из компонентов соединения осуществляется в большинстве случаев только в своей подрешетке. Переход вакансий из одной подрешетки в другую практически исключен, так как при этом ближайшими соседями атома, перешедшего в вакансию, стали бы одноименно заряженные атомы. Чем больше ионная составляющая связи в соединении, тем менее вероятен такой переход. Поэтому переход атома в вакансию может осуществляться только через междоузлия. Этим объясняется и высокое значение энергии активации самодиффузии в соединениях.

Во всех известных случаях для соединений $A^{III}B^V$ и $A^{II}B^{VI}$ энергия активации диффузии элемента A^{III} (A^{II}) значительно меньше, чем B^V (B^{VI}). Это связано с тем, что уход из узла атома B^V или B^{VI} требует больших энергетических затрат, чем уход атомов A^{III} или A^{II} .

На скорость и параметры самодиффузии в соединениях существенно сказывается также отклонение от стехиометрии. Последнее всегда ускоряет самодиффузию того элемента, содержание которого меньше стехиометрического и в подрешетке которого имеются вакансии.

8.7.2. Гетеродиффузия

Рассмотрим отличительные особенности гетеродиффузии в германии и кремнии. Как следует из табл. 8.1, элементы IB и VIII (переходные металлы) групп диффундируют в них с большей скоростью, чем элементы IIIA и VA групп. Механизм диффузии примесей IB и VIII групп, как мы уже установили, — междоузельный для примесей внедрения (литий) или диссоциативный, если атомы примеси могут размещаться как в узлах, так и в междоузлиях.

Особый практический интерес представляет поведение элементов IIIA и VA групп в германии и кремнии. В Si элементы IIIA группы (акцепторы) диффундируют на 1–1.5 порядка быстрее, чем примеси VA группы (доноры). В Ge, наоборот, коэффициенты диффузии элементов IIIA группы примерно на два порядка меньше, чем элементов VA группы.

Разность ковалентных радиусов примеси и основного вещества (ΔR) в общем случае для примесей из IIIA группы больше, чем для примесей из VA группы. Поэтому локальные упругие искажения в решетке матрицы должны быть больше вокруг примесных атомов IIIA группы (сюда будут стремиться вакансии) и скорость их диффузии тоже должна быть больше. Такое толкование хорошо согласуется с экспериментальными наблюдениями диффузии в кремнии. Аналогичная ситуация должна наблюдаться и в германии. Однако в действительности, как уже упоминалось выше, коэффициенты диффузии примесей из IIIA группы в Ge меньше, чем для примесей из VA группы. При диффузии в германии, по-видимому, появляется дополнительный фактор, который перекрывает влияние размерного фактора на скорость диффузии. Этим фактором являются заряженные вакансии V^- .

Вероятность образования акцепторных вакансий тем больше, чем меньше энергия их образования. О последней можно судить по расположению акцепторных уровней вакансий в запрещенной зоне. Эти уровни для Ge и Si существенно различны. В Ge акцепторный уровень вакансии расположен в нижней половине запрещенной зоны на 0.26 эВ выше потолка валентной зоны, а в Si в верхней половине — на 0.16 эВ ниже дна зоны проводимости. Отсюда следует, что доля электрически заряженных вакансий в Ge намного больше, чем в Si, где они все практически электрически неактивны.

В силу кулоновского взаимодействия между электрически заряженными вакансиями и примесями вокруг положительно заряженных донорных примесей (VA группы) в решетке Ge создаются области, обогащенные вакансиями, а вокруг отрицательно заряженных акцепторных примесей (IIIA группы) — области, обедненные вакансиями. Это значит, что перемещение акцепторных примесей будет затруднено, что и наблюдается в Ge.

При наличии нескольких примесей были установлены следующие закономерности. Диффузия донорной примеси замедляется при предварительном легировании образца акцепторной примесью и ускоряется при предварительном легировании донором. Причиной этого также является кулоновское взаимодействие доноров и акцепторов (образование донорно-акцепторных пар), которое уменьшает эффективную скорость диффузии.

Особенностью гетеродиффузии в полупроводниковых соединениях является то, что большинство примесей диффундируют по подрешеткам тех компонентов соединения, которые ближе к примесям по химическим свойствам. В некоторых случаях известную роль играет и размерный

фактор $\Delta R/R$.

Параметры диффузии примесных элементов, диффундирующих в одной подрешетке полупроводникового соединения, очень близки между собой и резко отличаются от параметров диффузии примесей, диффундирующих по другой подрешетке.

8.8. Диффузия и растворимость газов

Выращивание монокристаллов полупроводников, а также высокотемпературный прогрев при различных технологических операциях часто проводятся в атмосферах водорода, инертных газов (гелий, аргон и т. д.) или в вакууме. Кроме того, при нагреве камера роста кристаллов «дегазируется», выделяя в значительных количествах кислород и другие адсорбированные на стенках камеры газы. Все эти газы способны растворяться в расплаве и проникать в кристалл.

Вопрос о влиянии газов на свойства полупроводниковых материалов до сих пор нельзя считать окончательно решенным, во-первых, потому, что это влияние оказывается зачастую не прямым, а косвенным; а во-вторых, потому, что это косвенное влияние в разных полупроводниках оказывается существенно различным.

Для характеристики процессов диффузии и растворения газов в полупроводниках, как правило, пользуются следующими величинами: 1) коэффициентом диффузии D ; 2) растворимостью, определяющей равновесную концентрацию газа в полупроводнике при данной температуре и внешнем давлении газа в одну атмосферу. Эти две величины тесно взаимосвязаны друг с другом.

В Ge и Si, как правило, газы непосредственно не влияют на электрические параметры материала, поэтому сведения о их содержании в кристаллах получают обычно не полупроводниковыми (электрическими) методами. Согласно результатам масс-спектрального анализа количество (отдельно) водорода, кислорода и углерода в кристаллах Ge и Si, выращенных с помощью обычных методов, примерно одинаково и составляет $3 \cdot 10^{18} \text{ см}^{-3}$ в Ge и $2 \cdot 10^{19} \text{ см}^{-3}$ в Si.

Коэффициенты диффузии для водорода, гелия, кислорода и ксенона в германии приведены в табл. 8.1. Близкие по порядку величины значения коэффициентов диффузии наблюдаются и для соответствующих газов в кремнии. Представляется вероятным, что диффузия водорода и гелия в германии и кремнии носит междоузельный характер. Также предполагается, что водород в этих полупроводниках может быть частично связан с имеющимся там кислородом (носителей заряда водород не дает).

Высокие концентрации растворенного кислорода приводят к появлению характерных полос поглощения в Ge при $\lambda = 11.6$ мкм, а в Si при 9 мкм. Относительная величина коэффициента поглощения меняется в зависимости от содержания кислорода (введенного, например, в камеру роста путем дозированного давления). Было установлено, что величина коэффициента диффузии кислорода при 1300°C в Si $D \cong 10^{-10}$ см²/с [41]. В отличие от водорода, гелия и углерода, которые не оказывают заметного влияния на электрические свойства Ge и Si, кислород вызывает появление электрически активных центров в Si и Ge, подвергнутых термической обработке (см. гл. 3). В Ge и Si кислород присутствует либо в атомарном виде, либо образует комплексы $\text{Si}(\text{Ge})\text{O}_x$. Атомы кислорода, размещающиеся в междоузлиях, нейтральны, а комплексы $\text{Si}(\text{Ge})\text{O}_x$ обладают донорными свойствами.

В полупроводниковых соединениях газы тоже, как правило, не влияют на электрические параметры материала, исключение составляет кислород. Например, в GaAs и GaP с кислородом связываются наиболее глубокие донорные уровни, лежащие приблизительно посередине запрещенной зоны; в результате кристаллы с концентрацией активных центров, превышающей 10^{17} см⁻³, проявляют проводимость, близкую к собственной.

При выращивании полупроводников следует обратить внимание на такой газ, как азот. Он играет существенную роль в неравновесных процессах и поглощении, например, в GaP. Замещение атомов фосфора в узлах атомами азота, которые являются «изоэлектронными», не влияет на электрические параметры кристаллов. Однако, имея большую по сравнению с фосфором электроотрицательность и высокую растворимость (до 10^{19} см⁻³), примесь азота приводит к появлению интенсивных линий в спектрах излучательной рекомбинации, связанных не только с одиночными атомами азота, но и с парами $N - N$ -центров, расположенных на достаточно близких расстояниях [43].

8.9. Температурные зависимости растворимости примесей

Процесс диффузии может быть использован для легирования кристаллов полупроводников: объемного в случае быстро диффундирующих примесей и приповерхностного в случае медленно диффундирующих примесей (формирование $p - n$ -перехода).

Важными параметрами диффузионного легирования, как и легирования вообще, является предельная растворимость электрически актив-

ных примесей C_{im} и ее температурная зависимость. Как уже говорилось, основой для определения C_{im} может служить уравнение (8.21) для случая диффузии из неограниченного источника. Для быстро диффундирующих примесей (объемное легирование) C_{im} может быть определена с помощью холловских измерений концентрации носителей заряда, обусловленной введением в образец электрически активных примесей, при различных временах прогрева и при постоянной температуре. В случае медленно диффундирующих примесей нужно провести исследование глубины залегания $p - n$ -перехода в различных образцах с одинаковой исходной концентрацией C_i при постоянной температуре и при постоянном времени прогрева в зависимости от количества нанесенной на поверхность образца примеси. В основе этого способа определения C_{im} лежит установление перехода от случая диффузии из ограниченного источника примеси к случаю диффузии из неограниченного источника. Для первого случая (уравнение (8.19)) характерно изменение глубины залегания $p - n$ -перехода с ростом S при прогреве в течение одного и того же времени. Во втором случае глубина залегания $p - n$ -перехода определяется предельной растворимостью примеси C_{im} и, следовательно, не меняется с изменением количества примеси на поверхности. Таким образом, переход от первого случая ко второму должен проявиться в том, что изменение глубины залегания $p - n$ -перехода при некотором S прекращается. Это и будет доказательством справедливости уравнения (8.21), из которого при $x = 0$ находится C_{im} .

На рис. 8.11 представлены температурные зависимости предельной растворимости медленно и быстро диффундирующих примесей в Ge. Из этого рисунка можно сделать следующие выводы.

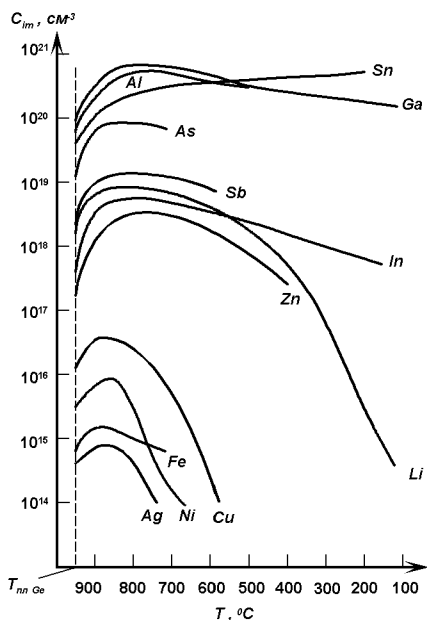
1. Для всех примесей характерен первоначальный рост растворимости с понижением температуры начиная с температуры плавления основного вещества.

2. Примеси IIIA и VA групп в Ge имеют предельную растворимость $10^{19} - 10^{21} \text{ см}^{-3}$, слабо или совсем не меняющуюся при дальнейшем понижении температуры (после первоначального возрастания).

3. Быстро диффундирующие примеси имеют значительно меньшие растворимости в соответствии с их коэффициентами разделения. У этих примесей наблюдается максимум растворимости в области $T < T_{пл.}$ при дальнейшем уменьшении растворимости с понижением T .

Поясним наблюдаемые зависимости. Для определенности рассмотрим в качестве примера примесный полупроводник n -типа, находящийся в равновесии с донорной примесью в газовой (внешней) фазе. Для простоты будем считать, что эффективные массы электронов и дырок равны

Рис. 8.11. Температурные зависимости растворимости примесей в германии.



массе свободного электрона. Поскольку в объеме полупроводника донорная примесь C_d частично диссоциирует на ионы C_d^+ и электроны e , то этот процесс может быть записан в виде следующей реакции:



где $[C_d^0]$ — концентрация нейтральных доноров в полупроводнике, а $[C_d^+]$ — концентрация ионизованных доноров. Одновременно с этим в полупроводнике происходит процесс «диссоциации» и «рекомбинации» электронов и дырок самого полупроводника: переход электронов из валентной зоны в зону проводимости и обратная рекомбинация. Этот процесс также можно выразить в виде реакции



Запишем условие электронейтральности для рассмотренной системы:

$$C_d^+ + p_v = n_c \quad \text{или} \quad C_d = n_c + n_d - p_v \approx n_c - p_v, \quad \text{так как } n_d \ll C_d, \quad (8.32)$$

где p_v — концентрация дырок в валентной зоне; n_c — концентрация электронов в зоне проводимости; n_d — концентрация электронов на донорных уровнях; C_d — полная концентрация введенных доноров.

Применение закона действующих масс к этим двум реакциям (8.30) и (8.31) и использование условия нейтральности для рассмотренной системы позволяет связать концентрацию доноров C_d с константами равновесия для реакции (8.30) K_+ и для реакции (8.31) K_i [41]:³

$$C_d = K_d / \sqrt{K_d + n_i^2}, \quad (8.33)$$

где $K_d = K_+ n_d$, а $n_i = K_i = n_c p_v$. Здесь K_d является функцией только температуры и в первом приближении $K_d \sim T^{3/2} \exp(-B/T)$.

Аналогичная формула может быть получена и для концентрации акцепторов:

$$C_a = K_a / \sqrt{K_a + n_i^2}; \quad K_a \sim T^{3/2} \exp(-B/T), \quad (8.34)$$

где $K_a = K_- n_a$; $K_- = K_+$; n_a — концентрация электронов на акцепторных уровнях.

Таким образом, в области низких температур $K_{d,a} > n_i^2$ и растворимость доноров C_d (акцепторов C_a) с ростом температуры возрастает пропорционально $K_d^{1/2}$ ($K_a^{1/2}$). В области высоких температур, когда $n_i^2 > K_{d,a}$, растворимость начинает убывать. Таков общий характер температурной зависимости растворимости электрически активной примеси.

Характерная температурная зависимость растворимости примесей с явно выраженным максимумом носит название *ретроградной*. Ретроградная растворимость ярко выражена для быстро диффундирующих примесей и проявляется (хотя и слабо) у некоторых медленно диффундирующих примесей (Ga, As, Sb, Zn на рис. 8.11). Отсутствие ее у остальных примесей может быть обусловлено «замораживанием» атомов в узлах решетки и малыми коэффициентами диффузии.

Выше было показано, что существует резкое различие в значениях коэффициентов диффузии в Ge и Si для примесных элементов IIIA, VA групп и примесных элементов IB, VIII (переходные металлы) групп таблицы Менделеева. С другой стороны, эти две группы примесей резко различаются и по значениям растворимости в Ge и Si (см. рис. 8.11). Те примесные элементы, которые обладают аномально высокой скоростью диффузии, имеют очень малую растворимость и малые коэффициенты разделения в Ge и Si. Такая корреляция не является случайной. Она

³Из закона действующих масс следует, что для реакции (8.30) $K_+ = [C_d^+][e]/[C_d^0]$. Так как $C_d^+ = C_d - n_d$; $[e] = n_c$; $[C_d^0] = n_d$, то $K_+ = (C_d - n_d)n_c/n_d$. Предполагая, что $n_d \ll C_d$, получим $K_+ = C_d n_c/n_d$. Применение закона действующих масс к реакции (8.31) дает $K_i = [e][h] = n_c p_v = n_i^2$. Тогда $n_c = K_+ n_d / C_d$, а $p_v = n_i^2 / n_c = n_i^2 C_d / K_+ n_d$. Подставим в (8.32) полученные n_c и p_v , тогда $C_d = n_c - p_v = K_+ n_d / C_d - n_i^2 C_d / K_+ n_d$.

обусловлена тем, что как диффузия, так и растворимость в конечном счете определяются одними и теми же факторами: концентрацией вакансий и энергией связи примесных атомов в узлах кристаллической решетки. При данной температуре кристалла (определяющей концентрацию вакансий) состояние примесных атомов в вакантных узлах кристалла будет тем устойчивее, а коэффициент диффузии тем меньше, чем больше их энергия связи, иначе — чем глубже потенциальные ямы, в которых находятся примесные атомы. С другой стороны, этот же фактор определяет и растворимость примесных атомов в кристалле: чем больше энергия связи примесных атомов в вакантных узлах кристаллической решетки, тем больше и вероятность занятия этих узлов, а следовательно, и больше растворимость.

Известно, что помимо соблюдения «благоприятного» размерного фактора, необходимым условием образования твердых растворов замещения является наличие у замещающих атомов валентных электронов в таких же квантовых состояниях, как и у атомов основной решетки. В Ge и Si ковалентная тетраэдрическая связь осуществляется посредством sp^3 -орбиталей. При этом основную роль в образовании ковалентной связи здесь играют p -электроны. Поэтому, для того чтобы примесные элементы могли образовывать с Si и Ge стабильные твердые растворы замещения, их атомы должны обладать валентными электронами на s - и p -орбиталях. У элементов IB и VIII групп, в отличие от элементов IIIA и VA групп, на внешней электронной оболочке совершенно отсутствуют p -электроны, обуславливающие основную связь в кристаллической решетке Ge или Si. С другой стороны, разность электроотрицательностей этих элементов в Si или Ge, как правило, невелика. Следовательно, ионная составляющая связи здесь не играет значительной роли. Все это способствует «неустойчивому» состоянию атомов IB и VIII групп в узлах кристаллической решетки Ge и Si, что приводит к их малой растворимости и высокой скорости диффузии.

Таким образом, аномально высокая скорость диффузии, малая растворимость и малые коэффициенты разделения в Si и Ge элементов IB и VIII групп вызвана существенным различием в электронной конфигурации валентных оболочек атомов этих элементов и атомов Si и Ge. Для элементов IIIA и VA групп это различие менее существенно, поэтому им присуща большая растворимость и малая скорость диффузии в Ge.

Эффекты, обусловленные ретроградной растворимостью

Представленные на рис. 8.11 температурные зависимости растворимости быстро диффундирующих примесей могут обуславливать эффекты обратимых изменений концентрации электрически активных примесей. Например, если примесь введена диффузией при достаточно высокой температуре, когда ее растворимость близка к максимальной, то последующее понижение температуры делает соответствующий твердый раствор пересыщенным. В таких условиях, как и любая неравновесная система, раствор стремится перейти в термодинамически равновесное состояние, соответствующее меньшей концентрации электрически активной примеси. Избыток ее должен выделиться во вторую фазу, то есть перейти в электрически неактивное состояние. Подобный эффект в полупроводниках носит название «осаждения примесей», а параметром, контролирующим его скорость, является коэффициент диффузии соответствующей примеси.

В качестве примера рассмотрим процесс осаждение лития в германии. Для лития, имеющего значительный коэффициент диффузии в широком интервале температур, процессы осаждения в германии достаточно интенсивны даже при комнатной температуре, потому наиболее исследованы [41]. Было выяснено, что зародыши второй фазы возникают вследствие попадания иона лития из междоузлия в вакансию V_{Ge} , при этом образуется комплекс: $Li^+ + e + V_{Ge} = LiV_{Ge}$. Таким образом, концентрация зародышей должна определяться концентрацией Li и концентрацией вакансий в германии. Поскольку концентрация вакансий зависит от совершенства кристалла — наличия дислокаций и примесей (помимо лития), то эти факторы должны также влиять на скорость распада твер-

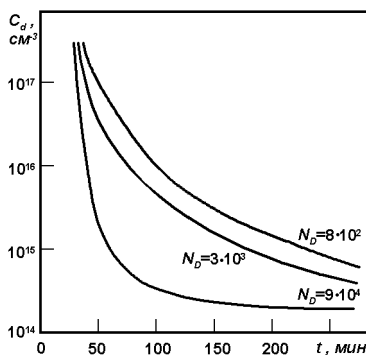


Рис. 8.12. Изменение концентрации электрически активного лития при $59^{\circ}C$ (после насыщения при $450^{\circ}C$) в образцах Ge с разной плотностью дислокаций N_D ($см^{-2}$).

дого раствора. Действительно, эксперименты показали, что на процессы осаждения лития значительное влияние оказывает, кроме коэффициента диффузии, наличие в кристаллах «центров осаждения» — дислокаций и порождаемых ими вакансий. На рис. 8.12 показана зависимость концентрации доноров от времени после насыщения литием при 425°C в образцах Ge с разной плотностью дислокаций N_D . Четко видно резкое увеличение скорости осаждения с ростом N_D и быстрое приближение к равновесной растворимости. Эффекты осаждения и образования акцепторных центров также весьма интенсивны при температурах $T > 500^{\circ}\text{C}$ в Ge, легированном Si [41].

Глава 9

Монокристаллические пленки

9.1. Эпитаксия

Слово эпитаксия состоит из двух греческих слов: «эпи» — «над» и «так-сис» — «упорядочивание». Поэтому термин *эпитаксия* означает наращивание кристаллографически ориентированных монокристаллических слоев на монокристаллические подложки или друг на друга. Монокристаллическая подложка в процессе выращивания играет роль затравочного кристалла.

Можно выделить два вида эпитаксии: гомоэпитаксию (автоэпитаксию) и гетероэпитаксию.

Гомоэпитаксия — это наращивание монокристаллической пленки на подложку из того же вещества, что и пленка. Например, наращивание эпитаксиальной пленки кремния на подложку (объемный кристалл) из монокристаллического кремния.

Гетероэпитаксия — это наращивание пленки на инородную подложку. Например, наращивание эпитаксиальной пленки PbTe на монокристаллическую подложку BaF_2 . Однако при гетероэпитаксии кристаллические решетки пленки и подложки должны быть подобны, а параметры решеток близки для обеспечения роста монокристаллического слоя.

Методы эпитаксиального выращивания монокристаллических пленок получили широкое распространение в технологии полупроводниковых приборов только тогда, когда научились получать пленки заданного состава, обладающие параметрами, близкими к параметрам объемных кристаллов. Можно сказать, что научно-технический прогресс в области микроминиатюризации интегральных схем в значительной мере обязан внедрению в производство методов эпитаксиальной кристаллизации. За-

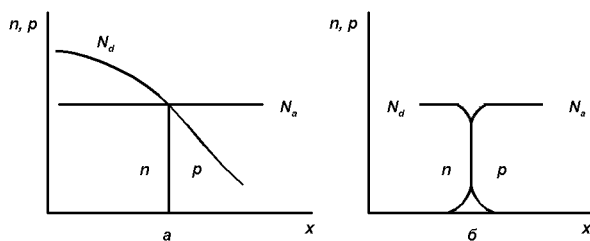


Рис. 9.1. Схемы окрестностей $p-n$ -перехода в случаях диффузионного (а) и эпитаксиального (б) легирования. При создании $p-n$ -перехода методом диффузии примесей в кристалл по обе стороны от перехода располагаются довольно широкие области сильно скомпенсированного материала.

мена объемных кристаллов эпитаксиальными пленками позволила повысить выход годных приборов, сократить длительность операций, снизить их себестоимость.

К основным преимуществам эпитаксиальной технологии можно отнести следующие преимущества.

Эпитаксиальное наращивание полупроводниковых пленок осуществляется, как правило, при температурах более низких, чем температуры получения объемных монокристаллов. При этом упрощается контроль за процессом кристаллизации и обеспечивается лучшая воспроизводимость свойств. Понижение температуры роста сопровождается замедлением диффузии примесей (в том числе и загрязняющих) в процессе получения эпитаксиальных пленок. В итоге улучшаются выходные параметры полупроводниковых материалов.

Эпитаксиальные методы роста позволяют достаточно просто осуществлять легирование монокристаллических пленок непосредственно в процессе их выращивания, обеспечивают однородное распределение легирующих элементов в пленках, дают возможность выращивать резкие $p-n$ -перехода со скачкообразно меняющейся концентрацией (рис. 9.1, б). Применение эпитаксиальных слоев на подложке предоставляет разработчику приборов возможность изменения профиля легирования в изготавливаемой структуре в гораздо более широких пределах, чем это возможно при использовании диффузии или ионной имплантации.

Существенной особенностью эпитаксиального наращивания из газовой фазы является возможность осаждения чистого материала на сильно легированных подложках. Такая возможность чрезвычайно важна для производства полупроводниковых квантовых генераторов света

(лазеров). С другой стороны, для ряда приложений (например, для изготовления транзисторов) необходимы тонкие эпитаксиальные слои полупроводниковых соединений на высокоомных подложках. Эпитаксиальные методы позволяют это довольно просто реализовать.

Использование эпитаксиальных пленок в электронной промышленности позволило существенно улучшить характеристики туннельных и лазерных диодов,¹ разработать технологию получения транзисторов с высоким коэффициентом усиления на высоких частотах,² мощных и высоковольтных транзисторов. На применении эпитаксиальных слоев основано производство таких приборов, как планарные полевые транзисторы, выполненные на структуре металл–окисел–полупроводник с изоляцией V-образными канавками (V-MOP). Эпитаксиальные структуры также используются для улучшения характеристик памяти с произвольным доступом и комплементарных интегральных МОП-схем. Новые перспективы в технике открыло применение эпитаксиальных гетероструктур, создание которых другими методами затруднено, в полупроводниковых приборах (например, для изготовления инжекционных лазеров). Кроме того, эпитаксия дает возможность получения многослойных структур со свойствами каждого слоя, практически не зависящими от свойств предыдущего слоя. Это открывает широкие возможности для разработки качественно новых типов электронных приборов.

Следует отметить, что интенсивная разработка технологических методов тонкопленочной эпитаксии, обеспечивающих прецизионное управление процессом роста и контроль качества получаемых структур, позволила совершить качественный скачок в развитии физики полупроводников. История развития физики полупроводников такова, что если основными объектами исследования лет 30 назад были монокристаллы, а лет 15 назад — эпитаксиальные пленки, то сейчас — это многослойные гетероструктуры, сверхрешетки, структуры с квантовыми нитями и точками.

Действительно, в ходе исследования свойств очень тонких пленок был обнаружен ряд новых интересных физических эффектов, которые откры-

¹Известно, что во многих оптоэлектронных приборах «объем» кристалла часто играет роль балласта, уменьшающего интенсивность выходящего излучения из-за оптического поглощения и рекомбинации. Отсюда естественное стремление уменьшить этот объем.

²Стремление к увеличению напряжения пробоя база–коллектор при формировании транзистора в объеме кремния требовало применения материала с высоким удельным сопротивлением, что при большой толщине кремния приводило к чрезмерному увеличению сопротивления коллектора, вело к увеличению рассеиваемой мощности и уменьшению коэффициента усиления. Использование высокоомных эпитаксиальных слоев на подложках с низким удельным сопротивлением позволило решить эти проблемы. Таким же образом были решены аналогичные проблемы и при производстве биполярных интегральных схем.

ли возможность создания принципиально нового класса полупроводниковых приборов — приборов, разработанных на основе свойств структур с квантовыми ямами. Не удивительно, что в настоящее время структуры с квантовыми ямами, нитями, точками являются одними из наиболее интенсивно исследуемых объектов в физике полупроводников, а характеристика стран как «технологических держав» в значительной мере определяется именно уровнем развития технологии получения таких структур.

В настоящей главе будут рассмотрены основы существующих представлений о первых стадиях зарождения и дальнейшем росте эпитаксиальных пленок из газообразных фаз, а также некоторые методы эпитаксиального наращивания тонких слоев полупроводников, включая и жидкостную эпитаксию.

9.2. Основные закономерности роста эпитаксиальных пленок при выращивании из газообразной фазы

В литературе принято разделять газообразные фазы на паровые и газовые. *Паровой* фазой называют газообразную фазу, состав которой совпадает с составом выращиваемого из нее нелегированного вещества или соединения. *Газовой* называют газообразную фазу, состав которой отличается от состава выращиваемого из нее нелегированного вещества или соединения.

Эпитаксиальное наращивание может осуществляться по схеме пар (газ) — кристалл [п(г)–к], а также по схеме пар (газ) — жидкость — кристалл [п(г)–ж–к]. В обоих схемах перенос вещества к месту конденсации происходит через газообразную фазу, поэтому скорость роста эпитаксиальных пленок оказывается невысокой.

Движущей силой процесса конденсации из газообразных фаз, как и любого фазового перехода, является разность термодинамических потенциалов газообразной и твердой фаз, причем величина ΔG определяется ΔP , где ΔP — абсолютное пересыщение газообразной фазы, равное разности между фактическим давлением газообразной фазы и равновесным давлением пара при данной температуре. Величина пересыщения определяет скорость роста кристаллической пленки. Процессы конденсации эпитаксиальных пленок из газообразной фазы в основном аналогичны процессам, которые подробно обсуждались в гл. 4 и 6, посвященных процессам зарождения центров новой фазы, механизмам роста и мето-

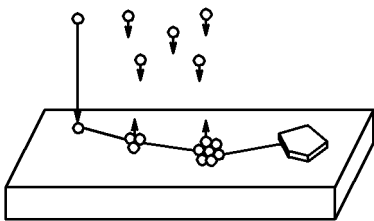


Рис. 9.2. Схема осаждения атомов и образования ориентированных зародышей на монокристаллическую пленку.

дам выращивания объемных монокристаллов из газообразной фазы.

Ниже мы рассмотрим особенности ростовых процессов при эпитаксии.

При росте по схеме п(г)–к процесс осаждения атомов на подложку с образованием и ростом монокристаллических зародышей можно представить следующим образом. При соударении атома с поверхностью подложки он либо адсорбируется на ней, либо возвращается обратно в газообразную фазу (рис. 9.2). Процесс адсорбции характеризуется *коэффициентом термической аккомодации* α_T , который определяется как

$$\alpha_T = \frac{T_1 - T_a}{T_1 - T_n} = \frac{E_1 - E_a}{E_1 - E_n}, \quad (9.1)$$

где T_1 и E_1 — температура и энергия поступающих на поверхность атомов, T_a и E_a — температура и энергия атомов адсорбированного слоя, T_n и E_n — температура и энергия атомов на поверхности подложки. Очевидно, что T_a зависит от температуры подложки и от температуры падающих частиц, то есть $T_a = f(T_1, T_n)$.

Установлено, что, как правило, влияние температуры пучка на температуру адсорбированного слоя мало. В этом случае процессы, происходящие на подложке, в значительной степени определяются величиной T_n . Если $T_a = T_n$, то $\alpha_T = 1$, что соответствует «прилипанию» атомов к подложке — полной аккомодации. При этом обычно происходит неориентированный рост пленки. Если $T_a > T_n$, то $\alpha_T < 1$, а на поверхности подложки происходит миграция атомов, которая может закончиться либо их десорбцией, либо их встраиванием в растущий зародыш (рис. 9.2). Следует отметить, что миграция (поверхностная диффузия) крайне необходима для получения монокристаллических пленок. Поверхностная диффузия приводит к тому, что атом «находит» свое равновесное положение и в результате на подложке образуются макроскопические ориентированные зародыши кристалла.

Режим, при котором происходит адсорбция значительной части падающих на подложку атомов совместно с их миграцией, — $T_{\text{п}} < T_a < T_1$. В таком режиме после установления термодинамического равновесия в системе время жизни t_n адсорбированного атома на поверхности подложки или растущей пленки будет

$$t_n \sim \frac{1}{\nu_0} \exp \left(\frac{E_{\text{дес}}}{kT_k} \right), \quad (9.2)$$

где ν_0 — частота колебаний адсорбированного атома вблизи положения равновесия на подложке, $E_{\text{дес}}$ — энергия десорбции, $T_k \equiv T_a$ — температура конденсации (см. гл. 4). Скорость поверхностной диффузии адсорбированного атома

$$V_{\text{диф}} \sim \nu_0 \exp \left(-\frac{E_{\text{диф}}}{kT_k} \right), \quad (9.3)$$

где $E_{\text{диф}}$ — энергия активации диффузии. Среднее число мест, занимаемых атомом, участвующем в процессе поверхностной диффузии, за время его жизни на поверхности

$$N_0 \sim \frac{t_n}{t_{\text{диф}}} = \exp \left(\frac{E_{\text{дес}} - E_{\text{диф}}}{kT_k} \right). \quad (9.4)$$

После достижения равновесия в адсорбированном слое (равновесие между адсорбированными частицами и зародышами) начинается образование стабильных зародышей конденсируемой фазы (центров новой фазы). Рост и срастание центров новой фазы приводит к образованию эпитаксиальных слоев. Образование первых стабильных зародышей требует больших пересыщений, чем их последующий рост (см. гл. 4).

Таким образом, адсорбированный на поверхности подложки атом может участвовать в одном из трех процессов (см. рис. 9.2):

- 1) десорбции с поверхности (обычно при больших $T_{\text{п}}$);
- 2) образовании вместе с другими атомами стабильного зародыша;
- 3) присоединении к уже существующему зародышу.

Для полного описания процесса зародышеобразования необходимо знать следующие основные характеристики: а) размер критического зародыша; б) скорость образования центров новой фазы и концентрацию критических зародышей с учетом возможного их изменения со временем; в) структуру и ориентацию критических зародышей. Все эти величины тесно связаны с условиями роста — степенью пересыщения газообразной фазы, температурой конденсации, скоростью конденсации, скоростью поверхностной диффузии, структурой и чистотой поверхности подложки и т. д.

Найдем основные зависимости. При этом учтем, что возможно два подхода к рассмотрению механизма зародышеобразования: термодинамический и молекулярно-кинетический.

Термодинамический подход целесообразен в том случае, когда стабильными являются зародыши с большими критическими размерами (50–100 частиц в зародыше). Эти зародыши обладают свойствами макрофазы. Обычно такое состояние процесса реализуется при небольших пересыщениях. Молекулярно-кинетический подход является плодотворным в том случае, когда стабильны зародыши столь малых размеров (1–10 атомов в зародыше), что к ним неприменимо представление о зародышах критических размеров в термодинамическом смысле. Обычно такое состояние реализуется при больших пересыщениях.

Рассмотрим процесс образования зародышей критического размера при эпитаксиальном росте в рамках термодинамического подхода. В процессе роста в условиях постоянного пересыщения ($\Delta P = \text{const}$) устанавливается равновесие в адсорбированном слое между концентрацией адсорбированных частиц, числом «дозародышей» и числом зародышей критического размера

$$n_k = N \exp(-E_{\text{кр}}/kT_k), \quad (9.5)$$

где n_k — концентрация критических зародышей, N — концентрация адсорбированных частиц, $E_{\text{кр}}$ — энергия образования критического зародыша (зависит от степени пересыщения).

Зародыши размеров, больших чем критические, в этом соотношении не участвуют, так как им термодинамически более выгодно расти, чем распадаться. Зародышам критического размера также термодинамически выгоднее расти, чем распадаться на «дозародыши». Поэтому с того момента, когда из «дозародыша» образовался зародыш критических размеров и он начал расти, он перестает участвовать в соотношении (9.5), а равновесие в нем смещается в сторону образования зародышей критического размера. Таким образом, в условиях роста идет непрерывное образование зародышей критического размера.

Скорость образования центров новой фазы

$$V \sim \omega n_k, \quad (9.6)$$

где ω — частота, с которой критический зародыш захватывает атомы или молекулы и становится центром новой фазы. Если предположить, что присоединение атомов к критическому зародышу происходит в основном

путем их поверхностной диффузии, то скорость образования центров новой фазы будет

$$V \sim \omega n_k \sim \exp \left(\frac{E_{\text{дес}} - E_{\text{диф}} - E_{\text{кр}}}{kT_k} \right). \quad (9.7)$$

Таким образом, мы нашли как концентрация критических зародышей n_k и скорость образования центров новой фазы V зависят от условий роста в рамках термодинамического подхода.³

Однако термодинамическая теория зародышеобразования неприменима к случаю, когда критический размер зародышей мал. Между тем из многих экспериментальных данных следует, что в реальных случаях эпитаксиального роста зародыши критических размеров состоят всего лишь из нескольких атомов. В этом случае необходимо применять молекулярно-кинетический подход к описанию процесса зародышеобразования.

Родин и Уолтон [44] предложили теорию образования критических зародышей, которые состоят всего из нескольких частиц. Следуя в основном общепринятой модели, эти авторы вместо свободной энергии ввели статистические суммы и потенциальные энергии, а затем при определенных допущениях вывели выражение для скорости образования зародышей, состоящих из i частиц. В рамках этой молекулярно-кинетической теории было найдено, что скорость образования центров новой фазы, состоящих из i частиц, описывается формулой

$$V \sim J(J/\nu_0 N)^i \exp \left[\frac{(i+1)E_{\text{адс}} + E_{\text{дис}} - E_{\text{диф}}}{kT_k} \right], \quad (9.8)$$

где J — интенсивность потока падающих атомов, ν_0 — собственная частота колебаний атома на поверхности, N — концентрация адсорбированных частиц, $E_{\text{адс}}$ — энергия адсорбции, $E_{\text{дис}}$ — энергия диссоциации скопления из i атомов, $E_{\text{диф}}$ — энергия активации диффузии.

Наиболее существенный результат этой теории заключается в возможности определения пороговых температур изменения типа ориентации растущей пленки. При больших пересыщениях стабильным зародышем может оказаться даже одиночный атом. Тогда, когда образуется комплекс из двух атомов, он скорее будет расти, чем распадаться. Снижение пересыщения ведет к тому, что вероятность присоединения к одному

³Размер критического зародыша в зависимости от условий роста в рамках термодинамического подхода определяется формулами, приведенными в гл. 4. Структура и ориентация критических зародышей определяется температурой подложки T_n , ее структурой и чистотой.

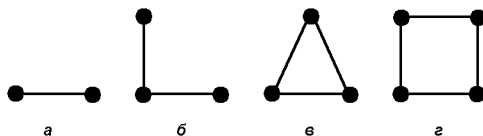


Рис. 9.3. Схематическое изображение конфигураций атомов в двухмерных зародышах: *а*, *б* — зародыши с одной и двумя связями на атом; *в*, *г* — наименьшие стабильные зародыши.

атому второго сравнивается с вероятностью распада пары. В этом случае стабильными будут конфигурации с двумя связями на атом, то есть при понижении пересыщения стабильность приобретают конфигурации из трех и четырех атомов. Такими конфигурациями являются треугольник и квадрат. Ясно, что в первом случае при дальнейшем росте центры новой фазы имеют ориентацию (111), во втором — (100) (рис. 9.3). Таким образом, изменение пересыщения должно сопровождаться изменением числа атомов в стабильных зародышах, что, как видно из рисунка, приводит к изменению ориентации. Температуру перехода от одной ориентации к другой можно узнать, приравнявая между собой выражения для скорости образования критических зародышей из 2 и 3 частиц, из 3 и 4 частиц и т. д. (рис. 9.4).

Температура конденсации T_k также оказывает значительное влияние на критические размеры и ориентацию зародышей. Так, при больших пересыщениях и низких T_k преобладает неориентированный рост. С повышением T_k при некоторой определенной температуре T происходит изменение числа атомов в критическом зародыше (по крайней мере от 1 до 2), и тогда уже должен наблюдаться ориентированный рост. Таким образом, при низких T_k велика вероятность образования стабильных критических зародышей, но отсутствует ориентация. При повышении T_k скорость зародышеобразования падает, но начинает преобладать ориентированный рост.

При больших пересыщениях и невысоких температурах конденсации, когда вероятность десорбции мала, скорость конденсации v_k приблизительно равна интенсивности потока пара J . Поэтому можно установить связь между температурой начала ориентированного роста или температурой смены ориентации и скоростью конденсации, что весьма удобно для экспериментальной проверки теории. Сопоставление экспериментальных данных о величинах V при выбранной T_k с расчетными, полученными из зависимости (9.8), при указанных условиях роста позволи-

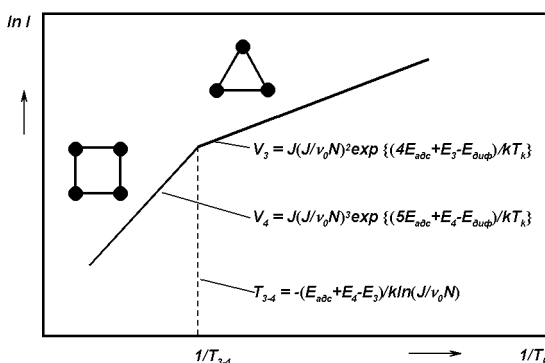


Рис. 9.4. Зависимость скорости образования зародышей от температуры конденсации.

ло подтвердить правильность молекулярно-кинетического подхода. Кроме того, формула 9.8 позволяет оценить число атомов в стабильном зародыше.

Реальные температуры источника и подложки при выращивании ориентированных монокристаллических пленок выбираются с учетом оптимизации всех параметров.

Большое внимание при эпитаксиальном росте пленок должно уделяться присутствию в системе загрязняющих газов. При наличии инородного газа, адсорбированного на поверхности подложки или растущей пленки, эпитаксия возможна лишь при условии, что скорость адсорбции инородного газа не превышает скорости присоединения атомов к ориентированному зародышу.

При росте пленок по механизму п(г)–ж–к на границе раздела фаз подложка — пленка возможны следующие ситуации:

1) На фронте кристаллизации образуется сплошной слой жидкой фазы значительной толщины. Процессы роста эпитаксиальной пленки в этом случае подобны процессам роста кристаллов из расплавов (растворов) (см. гл. 6).

2) На фронте кристаллизации образуется сплошная тонкая пленка жидкой фазы толщиной l , причем l находится в пределах от l_{min} до l_{max} . При $l < l_{min}$ пленка существует в виде отдельных островков; при $l > l_{max}$ сплошность пленки нарушается и на поверхности подложки образуются капли жидкой фазы.

Процессы роста эпитаксиальной пленки из капель аналогичны случаю (1); различие заключается лишь в том, что фронт кристаллизации не сплошной.

При кристаллизации тонкой пленки жидкой фазы (в том числе и тонкопленочных островков) процессы роста могут отличаться от процессов роста кристаллов из капель жидкой фазы в связи с тем, что строение пленки жидкой фазы и жидкой макрофазы (жидкости того же состава) может быть различным.

9.3. Экспериментальное изучение процессов зародышеобразования, роста и совершенства эпитаксиальных пленок в зависимости от условий роста

Экспериментальное изучение процессов зародышеобразования, роста эпитаксиальных пленок и их качества в зависимости от условий выращивания сейчас проводится с помощью самой современной аппаратуры непосредственно в процессе конденсации. На основании этих исследований можно сделать выводы о местах преимущественной адсорбции атомов, изучить влияние загрязняющих газов на образование зародышей, оценить плотность, совершенство и ориентацию зародышей, а также количественно оценить $E_{\text{диф}}$ и $E_{\text{дес}}$.

Наблюдения в вакууме 10^{-9} – 10^{-10} мм рт. ст. показывают, что, как правило, зародыши образуются преимущественно на неровных участках подложки, на плоскостях с большими индексами Миллера и с низкой ретикулярной плотностью. Поперечный размер зародыша увеличивается пропорционально корню квадратному из времени, что характерно для роста по механизму поверхностной диффузии.

При низких T_k зародыши образуются на многослойном адсорбате (промежуточный слой толщиной 1–5 монослоев), который либо аморфен, либо представляет собой однородный слой, повторяющий структуру подложки. С увеличением T_k толщина адсорбата уменьшается. Существуют два возможных подхода (для соответствующих случаев) объяснения механизма образования зародышей ориентированной пленки на многослойном адсорбате: 1) диффузионные параметры адсорбированных атомов на поверхности подложки сильно отличаются от параметров поверхностной самодиффузии; 2) связи атомов в первом осажденном слое с атомами подложки гораздо прочнее, чем межатомные связи в осаждаемом веществе.

Электронная микроскопия высокого разрешения в сочетании с методами электронной дифракции позволяет исследовать процесс образования конденсатов на стадии частичного заполнения первого монослоя. В большинстве случаев ориентированного роста на ранней стадии конденсации образуются изолированные трехмерные зародыши, так как количество питающего вещества недостаточно для образования монослоя. Эти зародыши обычно становятся видимыми в электронном микроскопе после того, как их размер достигает $10 \text{ \AA}$. Случаи роста конденсатов монослоями весьма редки. При гомоэпитаксии, когда растущий кристалл продолжает структуру подложки, в принципе, возможен рост без образования изолированных зародышей. Однако, как правило, образование полупроводниковых эпитаксиальных пленок происходит посредством возникновения и дальнейшего роста именно изолированных зародышей, что связано с присутствием загрязнений на поверхности подложки, которые служат центрами кристаллизации.

Электронографические и электронно-микроскопические исследования показали, что на начальной стадии роста эпитаксиальных пленок могут образовываться зародыши с различными ориентациями. Это явление широко распространено при эпитаксиальном росте и оказывает значительное влияние на структурное совершенство выращиваемых пленок. Вопрос о причинах образования зародышей с разными ориентациями является до сих пор предметом многочисленных теоретических и экспериментальных исследований. Исходят из того, что при образовании зародышей определяющая роль принадлежит поверхностной энергии, поэтому чаще всего образуются зародыши с минимальной поверхностной энергией. Наличие разнотипных зародышей означает, что существует несколько типов зародышей, отличающихся разной ориентацией и удовлетворяющих такому требованию (одинаковая минимальная поверхностная энергия при одинаковых условиях роста). Было исследовано влияние различных факторов (температуры подложки, скорости конденсации, структуры и чистоты подложки, вакуума и т. д.) на ориентацию зародышей. В результате были найдены условия роста, при которых образуются зародыши лишь с одной ориентацией (см. гл. 9.2).

Электронная микроскопия дает сведения и о влиянии условий роста на образование и плотность зародышей. Количественное определение зависимости плотности зародышей от условий кристаллизации (T_k , скорости конденсации v_k , состояния подложки) обнаружило удовлетворительное согласие между экспериментальными данными и молекулярно-кинетической теорией образования зародышей и позволило оценить такие параметры роста, как $E_{\text{дес}}$ и $E_{\text{диф}}$.

Наличие на поверхности подложки адсорбированных газов и загрязнений сопровождается рядом эффектов, влияющих на плотность, на совершенство и на ориентацию зародышей, а следовательно, и на их последующий рост. В конечном счете было показано, что адсорбированные газы и загрязнения на подложке приводят к изменению структуры и свойств конденсатов. Действительно, наличие инородных адсорбированных атомов способствует увеличению плотности зародышей, так как увеличивается число активных центров кристаллизации на поверхности подложки, и заметно изменяет подвижность адсорбированных атомов растущей эпитаксиальной пленки. Кроме того, адсорбированные газы способны существенно изменять структуру растущего слоя, так как при наличии на поверхности подложки газовых примесей может происходить рост зародышей с ориентацией, отличающейся от ориентации зародышей, образующихся в условиях сверхвысокого вакуума. В итоге на стадии коалесценции происходит срастание различным образом ориентированных зародышей, что ухудшает структурное совершенство пленки.

Таким образом, вопрос о получении атомно-чистых поверхностей подложки заслуживает особого внимания. Чтобы создать и сохранить чистую поверхность, необходима длительная термическая обработка или ионная (электронная) бомбардировка с последующим поддержанием сверхвысокого вакуума порядка 10^{-10} мм рт. ст. Без такой обработки и в обычном вакууме (10^{-6} – 10^{-7} мм рт. ст.) практически все кристаллы, используемые в качестве подложек, покрыты слоем адсорбированных газов. При комнатной температуре такой слой создается за несколько секунд.

Следует отметить, что различные виды обработки поверхности (термическая или ионная чистка, скалывание в вакууме, электронная бомбардировка и др.) сами по себе так же приводят к заметному изменению плотности зародышей. Это связано с изменением таких характеристик, как $E_{\text{диф}}$ и $E_{\text{дес}}$. Например, электронная бомбардировка на начальных стадиях роста пленок приводит 1) к повышению плотности зародышей (наблюдается скопление вакансий в местах образования зародышей); 2) к улучшению совершенства структуры и уменьшению электрического сопротивления пленок, так как растущие из зародышей островки получают глаже, а их коалесценция начинается гораздо раньше; 3) к образованию хорошо ориентированных монокристаллических пленок.

Таким образом, структура пленок зависит от начальной структуры зародышей, от скоростей их роста и особенностей коалесценции. Загрязнения на поверхности подложки сказываются на совершенстве выращиваемых пленок как в момент образования зародышей, так и при их

срастании. Повышение чистоты эксперимента обычно способствует повышению совершенства эпитаксиальных пленок и понижению температуры эпитаксии.

9.4. Особенности гетероэпитаксии

До сих пор мы не разделяли гомо- и гетероэпитаксию. Однако при гетероэпитаксиальном выращивании пленок существуют специфические проблемы. Одна из них — это вопрос о характере сопряжения кристаллических решеток выращиваемой пленки и подложки на границе раздела. Приведем основные из существующих представлений по этому вопросу.

1. Характер сопряжения решеток определяется степенью соответствия периодов решеток плоских сеток, по которым происходит сращивание: сетки должны быть подобны, а относительное различие периодов решеток не должно превышать определенную величину. Впервые эти представления были сформулированы Руайэ (1928–1935 гг.); в качестве предельного было названо несоответствие периодов решеток, равное 15%. Однако в ряде случаев, как показывает практика, возможно наращивание пленок и при различии периодов решеток срастающихся фаз более чем на 15%.

Для того чтобы объяснить этот результат в рамках структурно-геометрических представлений, было сформулировано положение о *поверхностном псевдоморфизме*. В кристаллическом поле подложки структура нарастающей фазы упруго деформируется, между подложкой и нарастающей фазой образуется псевдоморфный слой, который и обеспечивает эпитаксиальное наращивание. Межатомные расстояния в псевдоморфном слое переменны: от характерных для поверхности срастания до межатомных расстояний наращиваемой фазы. Толщина псевдоморфного слоя определяется природой срастающихся фаз. Основоположником теории о псевдоморфном слое был Ван дер Мерве. Его теория основана на представлении о существовании сильного физического взаимодействия атомов в сопряженных решетках. В ее рамках вычисляется энергия поверхности раздела двух различных кристаллов и выясняются условия, при которых эта энергия принимает минимальное значение. Знание минимальной энергии поверхности раздела позволяет определить предельные толщины псевдоморфных слоев, что весьма важно с практической точки зрения.

Так как величины упругой деформации сопрягаемых кристаллических решеток ограничены, то на поверхности раздела между псевдоморфным слоем и подложкой (геометрия плоских сеток которых подобна, но меж-

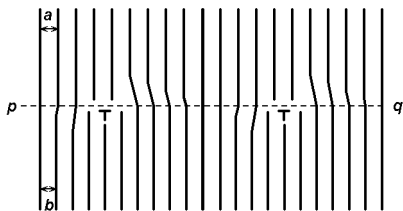


Рис. 9.5. Схема возникновения дислокаций несоответствия.

атомные расстояния разные) могут возникать дислокации несоответствия (рис. 9.5), которые компенсируют напряжения, возникающие в двух кристаллических решетках из-за различия их периодов. Иными словами, решетки сопрягающихся кристаллов упруго деформируются таким образом, чтобы частично (при большом несоответствии их периодов) или полностью (при малом несоответствии) ликвидировать геометрическое различие между ними на определенных участках поверхности, а оставшаяся доля несоответствия, выходящая за пределы упругой деформации, компенсируется возникновением дислокаций несоответствия. Плотность дислокаций несоответствия зависит от степени несоответствия периодов решеток срастающихся фаз: чем выше различие периодов решеток, тем больше плотность дислокаций несоответствия.

Дислокации несоответствия могут возникать при образовании стабильных зародышей, причем при боковом (тангенциальном) росте последних дислокации несоответствия могут перемещаться и взаимодействовать одна с другой, так как степень упругой деформации зародыша зависит от его размеров. Окончательная плотность дислокаций несоответствия фиксируется при слиянии стабильных зародышей (при образовании сплошной пленки).

Отметим еще, что явление псевдоморфизма на самой ранней стадии эпитаксии сводится не только к геометрическому искажению решетки, но и к изменению нормальной для решетки конденсата химической связи. Под влиянием поля подложки возникают такие явления как поляризация, ионизация, диссоциация атомов или молекул и т. п. По мере роста пленок в псевдоморфном слое происходит постепенное ослабление искажения связи.

2. Другой подход к проблеме сопряжения кристаллических решеток в плоскости раздела основывается на том, что определяющее влияние на эпитаксию оказывают активные центры роста на подложке. Предполагается, что явления поверхностного псевдоморфизма не существует:

кристаллическая структура наращиваемых слоев отвечает структуре, характерной для объемного материала. Эпитаксия и степень совершенства наращиваемых слоев определяются действием активных центров роста. Природа этих центров до сих пор окончательно не выяснена. Возможно, что они связаны с дефектами, находящимися в приповерхностном слое подложки и создающими локальные искажения потенциального рельефа подложки. В местах искажений и образуются стабильные зародыши. Эти представления не исключают возможность появления дислокаций на границе срастания зародышей, однако в этом случае их нельзя считать дислокациями несоответствия.

3. Третий подход основан на том, что между срастающимися фазами образуется переходный слой, состав (а следовательно, и структура) которого определяется равновесной фазовой диаграммой срастающихся фаз.

Приведенные краткие представления по вопросу о характере сопряжения кристаллических решеток выращиваемой пленки и подложки опираются на экспериментальные факты и каждое описывает процесс сопряжения кристаллических решеток для определенного круга веществ. В общем случае возможность сопряжения кристаллических решеток определяется возможностью соответствующей перестройки химических связей.

9.5. Причины образования структурных дефектов

Одной из важнейших практических задач при выращивании эпитаксиальных монокристаллических пленок является выяснение причин образования дефектов. Большая часть исследований посвящена изучению различных стадий роста эпитаксиальных пленок с целью нахождения способов, позволяющих снизить число дефектов в них.

Исследование структуры пленок в процессе их роста позволило определить основные источники дефектов в пленках. К ним относятся следующие.

1. Дефекты подложки. Пленка может наследовать дислокации подложки.
2. Дислокации несоответствия.
3. Малоугловые границы, образующиеся при срастании стабильных крупных зародышей, имеющих различную ориентацию.
4. Дефекты упаковки, возникающие при коалесценции зародышей. Причина их возникновения — несоответствие решеток зародышей, связанное с нарушением порядка чередования атомных плоскостей при слия-

нии стабильных зародышей по сравнению с ненарушенным монокристаллом. Дефекты упаковки могут возникать и в месте контакта зародыша и подложки.

5. Дислокации, возникающие из-за эффектов пластической деформации растущей пленки. После коалесценции зародышей и образования сплошной пленки механизм возникновения дислокаций несоответствия перестает действовать. Вместе с тем с ростом толщины пленки в связи с уменьшением доли несоответствия, компенсируемой за счет упругой деформации, напряжение, при котором начинается пластическая деформация, уменьшается, что ведет к возрастанию плотности дислокаций.

6. В случае сильно легированных подложек дефекты в эпитаксиальных слоях могут вызываться выделениями и предвыделениями второй фазы в подложке.

7. Загрязнения поверхности подложки также могут приводить к образованию дефектов в эпитаксиальном слое.

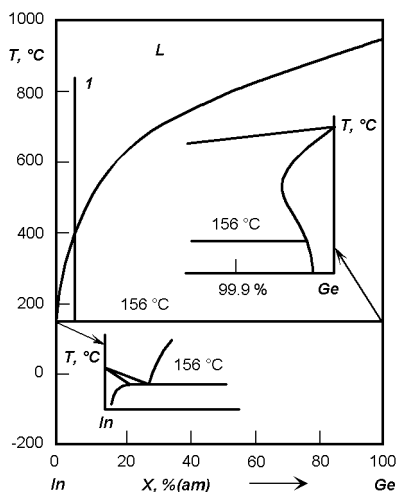
9.6. Методы эпитаксии

Существует целый ряд методов получения эпитаксиальных монокристаллических пленок, из которого принято выделять следующие группы: 1) осаждение из жидких растворов (жидкостная эпитаксия); 2) конденсация из газовой фазы (газовая эпитаксия); 3) конденсация из паровой фазы.

9.6.1. Жидкостная эпитаксия

Метод жидкостной эпитаксии по своей сути ничем не отличается от метода выращивания объемных монокристаллов из растворов, который уже подробно рассматривался (см. гл. 6). Основным преимуществом метода жидкостной эпитаксии является то, что рост эпитаксиальной пленки происходит при температурах более низких, чем температура плавления исходного вещества. Разберем подробнее этот метод выращивания эпитаксиальных пленок на примере системы Ge–In. Для этого рассмотрим диаграмму состояния системы Ge–In со стороны In. Температура плавления индия значительно ниже, чем у Ge (рис. 9.6), поэтому он и используется в качестве растворителя. Из диаграммы следует, что, меняя процентное содержание индия в сплаве Ge–In в соответствии с линией ликвидуса, можно менять температуру кристаллизации сплава. При охлаждении жидкой фазы вдоль линии 1 при температуре 400°C (точка пересечения с линией ликвидуса) раствор переходит в пересыщенное

Рис. 9.6. Диаграмма состояния системы Ge-In.



состояние и из него начинает выделяться твердая фаза, состав которой, согласно фазовой диаграмме, будет соответствовать чистому Ge, а состав жидкой фазы будет определяться по точке пересечения коноды с линией ликвидуса, то есть будет содержать 92% In и 8% Ge. По мере дальнейшего охлаждения составы твердой и жидкой фаз будут изменяться по линиям солидуса и ликвидуса (при равновесном охлаждении) соответственно, а их количество будет определяться правилом рычага.

Практически метод жидкостной эпитаксии реализуется следующим образом (рис. 9.7). Подложка Ge *n*-типа с ориентацией $\langle 111 \rangle$ закрепляется в верхней части наклонного графитового тигля, в нижней части которого находится раствор Ge в In. После достижения температуры 520°C нагрев тигля прекращается (раствор находится в перегретом состоянии) и его наклоняют так, чтобы раствор покрыл подложку. Далее начинается медленное охлаждение тигля, в процессе которого подложка Ge успевает частично раствориться (≈ 8 мкм) до установления равновесной концентрации Ge в растворе; при этом ликвидируются поверхностные нарушения подложки. Затем при фиксированной температуре происходит эпитаксиальное наращивание Ge на подложке с образованием пленки толщиной ≈ 25 мкм. После осаждения при температуре 350°C тигель возвращают в исходное положение, извлекают из печи и остаток In удаляют с поверхности кристалла. При необходимости легирования примесь добавляют в раствор с учетом ее коэффициента разделения в условиях

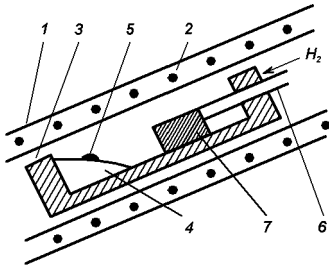


Рис. 9.7. Схема установки для выращивания эпитаксиальных пленок Ge методом жидкостной эпитаксии: 1 — герметичная камера; 2 — нагреватель; 3 — тигель; 4 — расплав; 5 — растворяемое вещество; 6 — держатель; 7 — подложка.

роста пленки. В качестве подложки для Ge оптимальными материалами служат сам Ge и GaAs, находящийся с ним в одном изоэлектронном ряду и имеющий очень близкий параметр кристаллической решетки.

Широкое распространение для выращивания пленок посредством жидкостной эпитаксии получил метод движущегося растворителя (см. гл. 6). Для выращивания эпитаксиальных пленок, например, GaAs (система Ga–As, Ga — растворитель) используются следующие ростовые условия: температура тигля $\approx 900^\circ\text{C}$, а градиент температуры вдоль оси растущей пленки $\Delta T \approx 1.5\text{--}2$ град/мм.

Метод жидкостной эпитаксии (выращивание из раствора) может использоваться и для получения сверхрешеточных структур [45]. Рост слоя достигается охлаждением жидкой фазы, контактирующей с подложкой, ниже температуры насыщения. Сам эпитаксиальный процесс состоит из ряда ростовых операций, в каждой из которых используются различные растворы, состав которых подбирается с целью обеспечить рост требуемых слоев в сверхрешетке. Операции роста чередуются с операциями переноса подложки между различными жидкими фазами, используемыми для выращивания сверхрешетки.

9.6.2. Эпитаксия из газообразной фазы

Процессы, происходящие при выращивании эпитаксиальных пленок методом газовой эпитаксии с помощью химических реакций, по существу, уже обсуждались при описании выращивания монокристаллов из газовой фазы (см. гл. 6). Рассматривались два основных способа выращивания эпитаксиальных пленок из газовой фазы с помощью химических реакций: 1) метод диссоциации и восстановления газообразных химических соединений и 2) метод газотранспортных реакций.

Первый способ отличается от второго тем, что в нем источником

материала для роста эпитаксиальной пленки служат легколетучие химические соединения, которые подвергаются термической диссоциации или восстановлению соответствующим газообразным восстановителем на поверхности подложки, то есть для роста эпитаксиальной пленки используются химические реакции, происходящие только на поверхности подложки. При этом поверхность подложки играет роль катализатора. Каталитическая активность подложки зависит от строения поверхности (кристаллографии) и наличия активных центров на ней. Выбираются такие химические реакции, продукты которых, за исключением кристаллизующегося вещества, являются газами и легко удаляются из зоны реакции. Для достижения стационарного процесса роста образующиеся в результате химических реакций газы необходимо непрерывно удалять, для чего всегда целесообразно использовать проточные системы.

Во втором методе в основе выращивания пленок из газовой фазы лежат обратимые химические реакции, то есть используется возможность изменения направления реакции в зависимости от температуры. В зоне источника химические реакции идут с образованием летучих химических соединений, содержащих кристаллизующее вещество. Затем летучие соединения за счет конвекции переносятся в зону конденсации, где происходит обратная химическая реакция с выделением кристаллизующего вещества. В этом методе принципиальна обратимость химических реакций в отличие от первого.

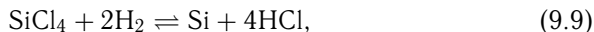
Скорость роста эпитаксиального слоя определяется выходом химических реакций и поэтому зависит от концентрации взаимодействующих компонентов в газовой смеси, давления в системе, скорости прохождения газовой смеси над подложкой, каталитической активности и температуры подложки. Эти параметры можно регулировать в процессе эпитаксиального наращивания. Для каждого материала или группы схожих материалов подбираются такой тип реакций и такие условия конденсации, которые обеспечили бы наилучшие структурные и электрические параметры выращиваемых эпитаксиальных пленок. Рассмотрим примеры выращивания эпитаксиальных пленок этими способами.

Метод химических реакций

Метод химических реакций широко используется при выращивании эпитаксиальных пленок Si. Как правило, используются реакции восстановления галогенидных соединений кремния (SiCl_4 , SiBr_4 , SiI_4 и SiHCl_3) водородом.

Наибольшее распространение получил метод восстановления SiCl_4

водородом. При этом водород обычно выполняет роль и газа-носителя, и восстановителя. Преимуществом этого метода является возможность достаточно простой очистки исходных реагентов. Процесс восстановления водородом осуществляется по схеме, изображенной на рис. 9.8. Реакционную камеру изготавливают, как правило, из кварцевого стекла и охлаждают водой или воздухом, чтобы исключить реакцию восстановления газовой смеси на стенках камеры, ограничив ее протекание поверхностью нагретых подложек. Реакционная камера соединяется с основной газовой магистралью. Поступающие в камеру газы проходят над поверхностью подложек и удаляются через стеклянную трубку в водоохлаждаемую ловушку. Вся система в целом должна тщательно продуваться. Поток водорода регулируется краном K_1 , который пропускает водород непосредственно в реакционную камеру, чтобы обеспечить требуемое молярное соотношение $\text{SiCl}_4/\text{H}_2 = 1:20 - 1:50$, и краном K_2 , направляющим водород в термостатированный испаритель, содержащий SiCl_4 (для насыщения H_2 парами SiCl_4). Было установлено, что количество выделяющегося кремния при фиксированной температуре конденсации может быть заметно увеличено, если в газовую фазу дополнительно ввести молекулы HCl . Поэтому в систему включен кран K_3 , через который вводится обезвоженный HCl . Кремниевые подложки вырезают из монокристаллов с высоким структурным совершенством в определенном кристаллографическом направлении (как правило, перпендикулярно направлению роста $\langle 111 \rangle$) и перед помещением их в реакционную камеру предварительно обрабатывают (механическая и химическая полировка). Осаждение ведут обычно при температуре подложек $1200 - 1290^\circ\text{C}$. Были опробованы различные методы нагрева подложек, однако в подавляющем большинстве вариантов промышленных установок используют высокочастотный нагрев. Благодаря локализованному нагреву в этом случае осаждение кремния на стенках камеры сводится к минимуму. Таким образом, подложки нагреваются излучением и посредством теплопроводности от нагревателя, питаемого индукционными токами. При этом идут основная реакция осаждения



реакции



и конкурирующая реакция травления



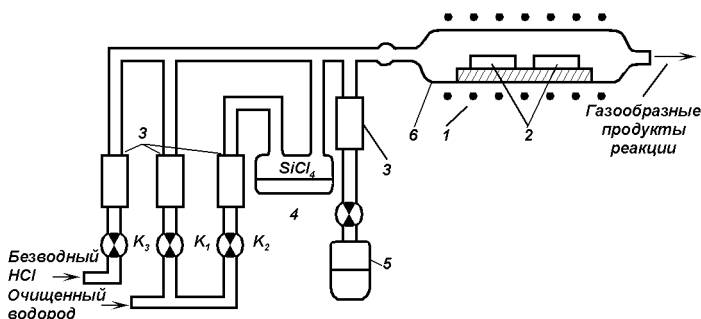


Рис. 9.8. Схема установки для получения эпитаксиальных пленок Si методом химических реакций: 1 — нагреватель; 2 — подложки; 3 — расходомеры; 4 — испаритель; 5 — источник легирующей примеси; 6 — реакционная камера.

Кроме того, в системе протекает ряд других реакций, определяемых возможными равновесиями в системе Si–H–Cl. Разность скоростей осаждения и газового травления подложки определяет скорость роста эпитаксиального слоя, характерное значение которой составляет 60–300 мкм/ч. Кроме того, скорость роста и качество получаемых эпитаксиальных пленок зависят от температуры подложки, относительной концентрации $SiCl_4/H_2$, скоростей газовых потоков, концентрации примесей, длительности процесса и геометрических характеристик системы.

Этим методом получают Si с удельным сопротивлением $1000 \text{ Ом} \cdot \text{см}$ и плотностью дислокаций $10^2\text{--}10^3 \text{ см}^{-2}$. Легирование эпитаксиальных пленок Si, получаемых восстановлением $SiCl_4$, обычно осуществляется либо 1) непосредственным добавлением легирующей примеси в испаритель к жидкости $SiCl_4$, либо 2) введением газообразных соединений примеси непосредственно в газовую фазу (PH_3 , $SbCl_3$, $AsCl_3$, BBr_3 или BCl_3).

Технология получения эпитаксиальных пленок Ge вполне аналогична описанной технологии выращивания пленок Si как с точки зрения используемого оборудования, так и самой сущности процесса.

Метод газотранспортных реакций

В процессе эпитаксиального роста с использованием обратимых химических реакций перенос вещества от источника к подложке осуществляется благодаря сдвигу равновесия используемой химической реакции в зонах источника и подложки, которые имеют разные температуры. Реагентами-переносчиками могут служить I_2 , Se, Te, H_2 , пары воды и

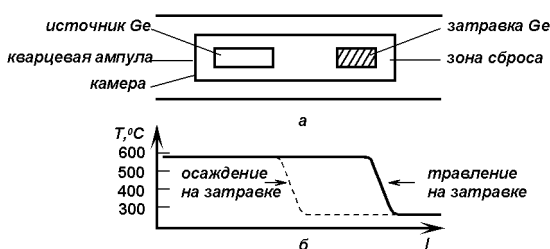
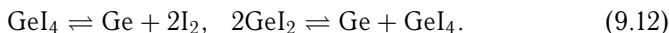


Рис. 9.9. *a* — Схема установки для получения эпитаксиальных пленок Ge методом газотранспортных реакций в закрытой системе; *b* — температурный профиль в реакционной камере в течение процесса роста (*l* — источник вещества; 2 — реакционная камера; 3 — затравка; 4 — зона сброса; 5 — кварцевая ампула; 6 — эпитаксиальная пленка).

др. Например, получение пленок Ge с помощью этого метода основано на обратимых реакциях



При фиксированном полном количестве йода равновесие в этих реакциях смещается вправо при понижении температуры.

В рассматриваемом случае реакционная камера загружается затравкой и источником Ge, а затем откачивается. Далее в нее вводят пары I_2 в количестве, необходимом для создания оптимального давления. Затем реакционная камера запаивается и помещается в печь (рис. 9.9). Вначале температурный профиль в реакционной камере устанавливается таким образом, чтобы при 500–550°C происходило травление материала источника и затравки, а Ge в составе летучего соединения GeI_4 удалялся и осаждался в зоне сброса, находящейся при температуре 300°C. В замкнутой системе перенос Ge в составе летучего соединения из области, находящейся при высокой температуре, в более холодную область осуществляется диффузионным образом (молекулярная и (или) конвективная диффузия). Затем газообразный продукт разложения I_2 вновь диффузионным образом переносится в зону источника. Далее температурный профиль изменяется так, чтобы температура в области подложки резко падала до 300–400°C. После этого начинается осаждение на подложку при температуре 300–400°C. Скорость переноса, определяемая скоростью диффузии и давлением йода, зависит также от температуры источника и затравки.

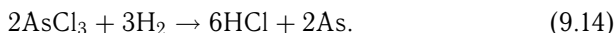
Аналогичные процессы происходят при выращивании эпитаксиальных

пленок Si по реакциям

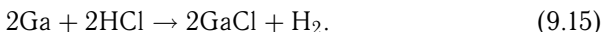


Характерные температуры для получения Si составляют $T_{\text{ист}} = 1150^\circ\text{C}$ и $T_{\text{подл}} = 900\text{--}950^\circ\text{C}$. Направление реакции сильно зависит от давления паров йода. Так, при низких давлениях преобладает реакция травления ($\text{Si} + 2\text{I}_2 \rightarrow \text{SiI}_4$) и кремний переносится из холодной зоны в горячую. При более высоких давлениях I_2 (выше 60–100 мм рт. ст.) перенос Si идет из горячей зоны в холодную.

Метод газотранспортных реакций может быть использован для получения эпитаксиальных пленок и в проточных системах. Например, пленки GaAs, обладающие хорошими электрофизическими свойствами, получают в проточных системах. В качестве исходных материалов используются следующие реагенты: AsCl_3 , Ga. Процесс проводится в установке, изображенной на рис. 9.10. Исходный Ga и подложки помещают в кварцевую трубу (реактор), которая расположена в двухзонной печи. В реактор подается смесь водорода и паров AsCl_3 . Для дозирования соотношения AsCl_3/H_2 поток водорода разделяют на две «нити»: по одной дозированное количество H_2 подается непосредственно в реактор, а по другой — водород сначала проходит через испаритель, содержащий AsCl_3 , насыщается его парами, а затем смешивается с основным потоком H_2 . Паро-газовая смесь, попадая в реактор, нагретый до температуры $T_1 = 800\text{--}850^\circ\text{C}$, подвергается реакции



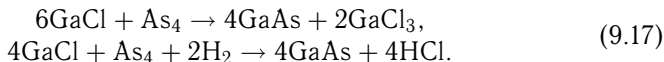
Образующийся хлористый водород реагирует с Ga, находящимся в лодочке,



Пары мышьяка полностью поглощаются расплавленным галлием, находящимся в лодочке в зоне с температурой $T_1 = 800\text{--}850^\circ\text{C}$, до тех пор, пока не образуется насыщенный раствор As в Ga. Газообразный GaCl переносится в более холодную часть камеры ($T_2 = 750\text{--}800^\circ\text{C}$) и частично диссоциирует по реакции



После того, как образовался насыщенный раствор As в Ga (2.3% мышьяка), свободный мышьяк переносится вместе с GaCl и во второй зоне происходят реакции



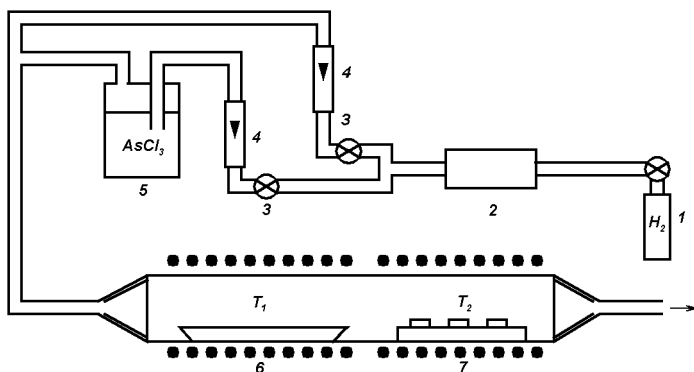


Рис. 9.10. Схема установки для получения эпитаксиальных пленок GaAs методом газотранспортных реакций в проточной системе: 1 — водород; 2 — блок тонкой очистки водорода; 3 — игольчатые натекатели; 4 — расходомеры; 5 — испаритель с $AsCl_3$; 6 — расплав галлия; 7 — подложки.

В качестве подложек используют пластины с ориентацией $\langle 110 \rangle$, вырезанные из монокристаллических слитков полуизолирующего GaAs (который получается введением глубоких примесей Cг и O).

Все рассмотренные методы эпитаксиального наращивания используются и для получения гетероэпитаксиальных слоев.

Рост из газовой фазы с использованием металлорганических соединений

В последнее время усиленно развивается технология выращивания сверхрешеточных структур из газовой фазы с использованием металлорганических соединений (MOCVD⁴). В основе этого метода лежит метод химических реакций.

MOCVD представляет собой метод выращивания, в котором необходимые компоненты доставляются в камеру роста в виде газообразных металлорганических алкильных соединений, и рост слоя осуществляется при термическом разложении (пиролизе) этих газов и последующей химической реакции между возникающими компонентами на нагретой пластине-подложке [46]. В настоящее время посредством этого метода можно выращивать большинство полупроводниковых соединений $A^{III}B^V$, $A^{II}B^{VI}$, $A^{IV}B^{IV}$.

⁴MOCVD — metalloorganic chemical vapor deposition.

Рассмотрим более подробно этот метод получения сверхрешеточных структур на примере выращивания сверхрешеток на основе GaAs. В работе [47] впервые было показано, что разложение газовой смеси металлорганического соединения триметилгаллия $(\text{CH}_3)_3\text{Ga}$ и гидрида мышьяка (арсина) AsH_3 , проводимое при температуре 600–700°C в атмосфере H_2 , можно использовать для выращивания тонких эпитаксиальных пленок GaAs в открытом (проточном) реакторе (рис. 9.11). Следует отметить, что MOCVD хорошо подходит для массового производства приборов с субмикронными слоями и имеет ряд преимуществ перед другими технологиями (см. ниже и [48, 49]).

Эпитаксия в процессе MOCVD осуществляется при пропускании однородной газовой смеси реагентов с газами-носителями над нагретой подложкой в реакторе с холодными стенками. Энергия для нагревания газовой смеси обычно создается мощным радиочастотным генератором с частотой порядка 450 кГц. Этот же генератор нагревает и графитовый держатель, на котором размещена монокристаллическая подложка. При этом газовая смесь вблизи поверхности подложки нагревается до высоких температур, а стенки реакционной камеры остаются относительно холодными, что приводит к осаждению полупроводниковой пленки на поверхности подложки (реакция идет при повышенных температурах) при малых потерях реагирующих веществ на поверхности реактора. Тщательно контролировать необходимо только температуру подложки, причем небольшие ее изменения не играют роли, поскольку большинство свойств пленок, полученных MOCVD, малочувствительны к таким изменениям температуры.

Парциальными давлениями различных газовых компонентов можно управлять с помощью электроники, контролирующей скорость потока от каждого из источников. Это позволяет с высокой точностью управлять всеми параметрами роста, обеспечивая надежную воспроизводимость результатов.

Металлорганические соединения, используемые для получения полупроводниковых пленок при комнатной температуре, как правило, являются жидкостями, обладают высоким давлением паров и могут быть легко доставлены в зону реакции путем пропускания газа-носителя, например H_2 , через жидкости, которые играют роль источников (рис. 9.11). Кроме того, используемые для роста полупроводниковых пленок гидриды при комнатной температуре являются газами и обычно берутся в качестве добавок к H_2 . Эти металлорганические и гидридные компоненты смешиваются в газовой фазе и пиролизуются (при 600–800°C, как правило) в потоке водорода в открытом реакторе при атмосферном или несколько

пониженном давлении.

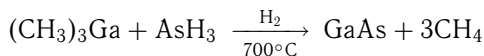
В большинстве случаев при выращивании полупроводниковых эпитаксиальных пленок используется метиловая и этиловая металлорганика. Такие соединения относительно просты в изготовлении и легко пиролизуются в атмосфере водорода, выделяя атомы соответствующего металла и в качестве побочных продуктов газообразные метан или этан, которые могут быть легко удалены из ростовой камеры.

Многослойные, многокомпонентные эпитаксиальные структуры могут быть последовательно получены в едином ростовом цикле. Кроме того, поскольку в процессе роста не участвуют травящие вещества, то возможно получение резких границ между различными слоями, причем возрастает однородность слоев по толщине и составу, так как процесс роста не является результатом конкуренции между осаждением и травлением, как в некоторых других методах газофазной эпитаксии.

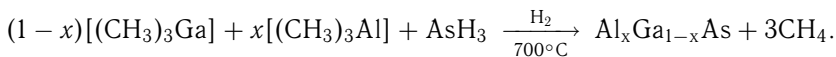
И, наконец, рассматриваемый технологический процесс может быть воспроизведен в большом объеме. При необходимости могут быть достигнуты высокие скорости роста, удовлетворяющие требованиям промышленности.

1. Механизм роста и эпитаксия в процессе МOCVD

Процесс МOCVD полупроводниковых соединений осуществляется, как правило, в несколько последовательных этапов, происходящих в различных зонах реактора. Рассмотрим этот процесс на примере выращивания GaAs и $\text{Al}_x\text{Ga}_{1-x}\text{As}$. Основные химические реакции в этом случае могут быть записаны в виде



и



При МOCVD намеренно создается неравновесная ситуация, когда поступающая газовая смесь содержит большую концентрацию реагентов, чем концентрация реагентов в газовой смеси в случае равновесия с подложкой при температуре роста.

При этом исходят из следующих соображений. При равновесии химические потенциалы каждого из компонентов μ_i в газовой и твердой фазах должны быть одинаковыми. При отклонении от равновесия система будет стремиться восстановить его со скоростью, пропорциональной

$\Delta\mu$ (разности химических потенциалов). Таким образом, максимальная скорость роста оказывается ограниченной требованиями термодинамики и скоростью пропускания пересыщенного газа через реактор.

Было выяснено, что скорость роста, за исключением случая очень высоких температур, оказывается существенно меньше своего термодинамически вычисленного значения. По-видимому, кинетика процесса роста является недостаточно быстрой для установления равновесия в системе. Обычно равновесие существует вблизи растущей поверхности, но отсутствует в газовой фазе вдали от этой поверхности. При этом скорость роста пропорциональна разности химических потенциалов между газовой фазой на входе и той частью газа, которая находится в равновесии с растущей пленкой.

Термодинамические соображения определяют не только скорость роста, но часто и состав растущего тройного или четверного твердого раствора $A^{III}B^V$. Если на границе роста устанавливается термодинамическое равновесие, то состав вещества, растущего из газовой фазы данного состава при данной температуре, диктуется термодинамикой. Заметим, что даже если газовую фазу считать идеальным газом, то термодинамическое описание твердой фазы через ее химический потенциал требует учета взаимодействия компонентов твердого раствора между собой. Этим взаимодействием определяются области существования твердых растворов. Понятно, что для получения информации о зависимости состава растущего твердого раствора от состава газовой фазы требуется, по- существу, знание фазовой диаграммы взаимодействующих бинарных соединений (см. гл. 4).

Итак, было выяснено, что процесс эпитаксии определяется термодинамическим состоянием системы, но реальная скорость роста часто меньше максимального значения, допускаемого термодинамикой. Это означает, что процесс лимитируется либо кинетическими эффектами на поверхности подложки, либо скоростью доставки питающего материала к ростовой поверхности. К поверхностным кинетическим эффектам можно отнести адсорбцию на поверхности, поверхностные реакции, встраивание атомов в растущую эпитаксиальную пленку. Однако было установлено, что процесс MOCVD лимитируется поверхностной кинетикой лишь при очень низких температурах. Обычно самым медленным процессом является диффузия реагентов через пограничный слой, и именно она чаще всего определяет скорость роста в этом методе.

Динамика газового потока в реакторе при близком к атмосферному давлении приводит к образованию пограничного слоя у неподвижной поверхности подложки. Из-за пониженной скорости газа в данной области

этот пограничный слой представляет собой квазинеподвижную газовую прослойку, параметры которой определяют скорость переноса реагентов к поверхности подложки. При достижении поверхности подложки реагенты связываются на ней. Пограничный слой также характеризуется большим градиентом температуры, диффундируя через который реагенты могут разлагаться, высвобождая атомы веществ, необходимые для роста слоя. Если скорость диффузии через пограничный слой превышает скорость разложения реагентов, то заметное влияние на рост пленки могут оказывать каталитические эффекты на поверхности подложки.

Легирование растущего эпитаксиального слоя при MOCVD осуществляется путем введения в газовый поток соответствующего реагента. По-видимому, этот реагент будет диффундировать через пограничный слой так же, как основные реагенты. Встраивание же примесей в кристаллическую решетку может определяться либо поверхностной кинетикой, либо скоростью протекания химических реакций реагентов.

Итак, подведем итог в вопросе о влиянии условий выращивания на скорость роста эпитаксиального слоя, получаемого MOCVD. Эта скорость определяется тремя факторами [50]: термодинамическим состоянием системы, диффузией реагентов через пограничный слой и поверхностной кинетикой. На все три фактора влияют такие условия роста, как скорость газового потока, парциальное давление реагентов в газовом потоке и температура подложки.

Скорость газового потока в реакторе при фиксированных парциальных давлениях реагентов влияет на определяемую термодинамикой скорость роста, изменяя количество «строительного материала», проходящего через реактор. Эта скорость также влияет на толщину пограничного слоя, то есть на время диффузии через него. В обоих случаях увеличение скорости газового потока ускоряет рост эпитаксиального слоя.

Увеличение парциального давления реагентов в газовом потоке также увеличивает скорость роста.

Повышение температуры подложки увеличивает скорость роста эпитаксиальной пленки в случае, когда она лимитируется поверхностной кинетикой. Но для экзотермических процессов, как правило, имеющих место при MOCVD, повышение температуры подложки уменьшает равновесный выход материала, то есть если скорость роста определяется термодинамикой, то увеличение температуры подложки уменьшает ее. Если процесс роста лимитируется диффузией, то температура мало влияет на скорость роста.

Влияние ориентации подложки существенно лишь в случае роста, лимитируемого поверхностными кинетическими явлениями.

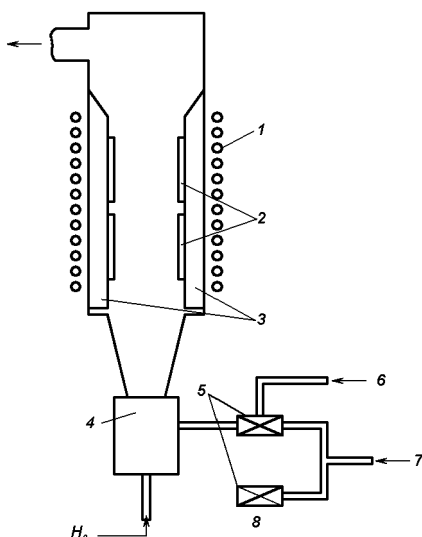
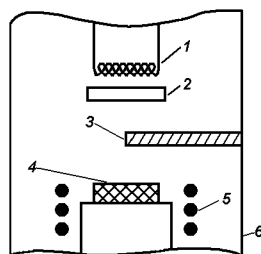


Рис. 9.12. Схема оптимизированного реактора для МОСVD (в этом вертикальном реакторе с восходящим потоком выпуск газа осуществляется снизу; подложки закреплены на внутренней стороне полого графитового держателя): 1 — катушка высокочастотного генератора; 2 — подложки; 3 — графитовый держатель; 4 — смесительная камера; 5 — пневматические клапаны; 6 — линия продувки; 7 — компоненты растущей пленки; 8 — обходная линия.

II. Выращивание сверхрешеточных структур

Рост многослойных структур в методе МОСVD осуществляется путем изменения газовой атмосферы в реакторе. Скорость, с которой осуществляется такое изменение, зависит от величины потока и геометрии реактора. Метод МОСVD позволяет при больших потоках производить замену газа достаточно быстро и получать при этом резкие гетеропереходы [49]. Так как время, необходимое для замены, определяется величиной потока, то резкость границ оказывается связанной со скоростью роста и зависит от концентрации реагентов в газовом потоке через реактор. Для получения минимальной толщины переходных слоев в эпитаксиальных структурах при МОСVD состав газовой фазы вблизи подложки должен меняться максимально быстро. Однако даже идеально резкое изменение газового состава перед реакционной камерой не позволяет получить резкого скачка состава вблизи подложки из-за размытия состава газовой смеси в процессе подхода к поверхности подложки. Это ведет к плавному изменению состава слоистой эпитаксиальной структуры. Для преодоления этой проблемы была создана наиболее оптимальная конструкция реактора (рис. 9.12) и предпринят ряд технологических мер, позволяющих свести к минимуму влияние факторов, размывающих скачок концентрации [51].

Рис. 9.13. Схема установки для получения эпитаксиальных пленок из паровой фазы: 1 — нагреватель; 2 — подложка; 3 — заслонка; 4 — источник; 5 — нагреватель; 6 — рабочая камера, обеспечивающая хороший вакуум.



9.6.3. Конденсация из паровой фазы

Выше нами были рассмотрены некоторые методы газовой эпитаксии, использующие обратимые химические реакции. Следующая группа методов эпитаксии основана на кристаллизации пленок из паровой фазы. Вырашивание пленок из паровой фазы, образованной из атомов и молекул вырабатываемого материала, производится преимущественно в откачанных вакуумных камерах (рис. 9.13). Процесс сводится к созданию потока паров, испускаемых источником, нагретым до выбранной температуры возгонки или испарения. Пары, пройдя некоторый путь, конденсируются на подложке. Давление насыщенного пара элементарного вещества, образующего одноатомные молекулы, как функция температуры описывается уравнением Клаузиуса-Клапейрона. Однако поскольку процесс проводится в динамическом режиме, реальное давление паров над поверхностью источника несколько ниже и описывается выражением

$$P = \alpha_0 P_s, \quad (9.18)$$

где P_s — давление насыщенных паров при температуре источника $T_{\text{ист.}}$, α_0 — коэффициент обратной аккомодации, равный отношению числа атомов, которые реконденсировали с поверхности источника, к числу атомов, столкнувшихся с поверхностью источника.

В случае двухкомпонентного пара (А и В) его давление $P_{\text{общ.}}$ над источником, содержащим оба компонента, (в соответствии с законом Рауля, который выполняется для идеальных газов⁵) равно

$$P = P(A) + P(B) = \alpha_0(A)P_s(A)N(A) + \alpha_0(B)P_s(B)N(B), \quad (9.19)$$

где $P(A)$ и $P(B)$ — давления паров компонентов А и В; $\alpha_0(A)$ и $\alpha_0(B)$ — коэффициенты обратной аккомодации компонентов А и В, $P_s(A)$ и

⁵Идеальный газ — газ, в котором пренебрегается взаимодействием между атомами или молекулами, входящими в его состав.

$P_s(B)$ — давление насыщенных паров компонентов А и В при $T_{\text{ист}}$, а $N(A)$ и $N(B)$ — мольные доли компонентов А и В ($N(A) + N(B) = 1$). Приведенные формулы могут быть обобщены и на случай многокомпонентного пара.

В процессе кристаллизации из двух- или многокомпонентного пара принимаются специальные меры к сохранению его постоянного состава. Следует отметить, что создание многокомпонентной паровой фазы заданного состава сопряжено со значительными трудностями. Чем больше различие в давлении насыщенных паров компонентов соединения, тем труднее оказывается управлять составом паровой фазы и возникает необходимость использовать раздельное испарение компонентов кристаллизующего вещества. При этом давление паров компонентов регулируется температурами источников.

Другой, не менее трудной, задачей при конденсации из паровой фазы является управление составом пленки, которая образуется при конденсации паров нескольких элементов. Коэффициент конденсации зависит от природы конденсирующихся атомов, материала и состояния подложки, а значит и состав образующейся пленки может быть не идентичен составу паровой фазы.

При проведении процессов роста эпитаксиальных пленок в вакууме, а особенно в непрерывно откачиваемых вакуумных системах, наименее контролируемым и наименее изученным является влияние присутствующих в рабочей камере остаточных газов и паров. При давлении остаточных газов 10^{-6} – 10^{-4} мм рт. ст. поток газовых молекул, бомбардирующих поверхность роста, часто сравним с потоком атомов конденсируемого пара (10^{14} – 10^{15} атомов/см² · с). Остаточные газы, способные вступать в химические реакции с веществом подложки и входить в решетку кристалла, оказывают, как правило, вредное влияние на скорость роста, структуру и электрофизические параметры пленок (см. выше разд. 9.3). Влияние же инертных газов, по-видимому, незначительно, а иногда даже может способствовать улучшению характеристик эпитаксиальной пленки. Поэтому выращивание эпитаксиальных пленок из паровой фазы проводится в тщательно дегазированных герметичных системах с остаточным давлением химически активных газов (кислород, углеводороды и др.) не более 10^{-8} – 10^{-10} мм рт. ст.; остаточное давление инертных газов может быть порядка 10^{-6} мм рт. ст.

Если эпитаксиальные пленки выращиваются в вакууме в режиме молекулярных (атомных) пучков, то вакуум в камере, где происходит конденсация, поддерживается таким, чтобы длина свободного пробега молекул (атомов) превышала расстояние от источника до подложки (см.

гл. 6). Пересыщение над подложкой определяется потоком пара кристаллизующего компонента и температурой подложки. Регулируя температуру источника и подложки, можно управлять давлением паров над источником, пересыщением, а следовательно, и скоростью роста пленки.

Основными технологическими факторами, определяющими возможность получения эпитаксиальных пленок с контролируемыми свойствами методом конденсации из паровой фазы, являются: 1) природа, кристаллографическая ориентация и состояние поверхности подложки; 2) выбор температуры источника $T_{\text{ист}}$ (величины пересыщения) и температуры подложки $T_{\text{под}}$, при которых обеспечивается, с одной стороны, закономерное встраивание атомов в решетку растущей пленки, а с другой стороны, заданный химический состав растущего кристалла.

Существует довольно большое число различных методов выращивания эпитаксиальных пленок из паровой фазы: метод испарения, возгонка, катодное распыление, лазерная эпитаксия и т. д. Мы здесь остановимся только на двух методах, обладающих определенными преимуществами перед другими методами и детально разработанных в связи с возможностью получения с их помощью сверхрешеточных структур. Это метод горячей стенки и молекулярно-лучевая эпитаксия (МЛЭ).

Метод горячей стенки

Метод горячей стенки представляет собой разновидность технологии вакуумного осаждения пленок, характеризуемую тем, что рост происходит в условиях, очень близких к термодинамическому равновесию. Основной особенностью метода является наличие нагретого экрана (горячей стенки), служащего для сосредоточения и направления на подложку испаряемого вещества. При этом исключаются потери испаряемого материала, создается возможность поддержания высокого давления паров вещества или его различных компонент и сводится до минимума разность температур источника и подложки.

К концу 70-х годов метод горячей стенки был детально разработан в применении к целому ряду соединений $A^{IV}B^{VI}$ и твердых растворов на их основе. Это позволило плавно, без каких-либо серьезных модификаций технологических установок и режимов роста, перейти от эпитаксии отдельных слоев к изготовлению сверхрешеток. Сверхрешетки на основе полупроводников $A^{IV}B^{VI}$ были впервые получены в начале 80-х годов в университетах Японии [52, 53] и Австрии [54] методом горячей стенки.

Типичные установки для изготовления сверхрешеток этим методом показаны на рис. 9.14. В качестве примера используются соединения

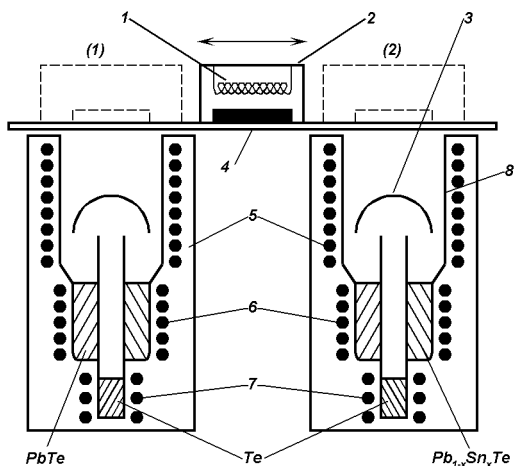


Рис. 9.14. Схема установки для эпитаксиального выращивания сверхрешеток PbTe-PbSnTe методом горячей стенки (во всей системе вакуум $\sim 10^{-6}$ мм рт. ст.; резервуары, содержащие Te , используются при росте $p\text{-PbTe}$ и $p\text{-PbSnTe}$; избыток Te создает отклонение от стехиометрии): 1 — нагреватель подложки; 2 — головка с подложкой, периодически передвигаемая мотором; 3 — отражатель паров Te ; 4 — подложка; 5 — нагреватели стенок; 6 — нагреватели источников; 7 — нагреватели резервуаров Te ; 8 — стенка из кварцевого стекла.

PbTe и $\text{Pb}_{1-x}\text{Sn}_x\text{Te}$. В состав установки входят несколько (обычно два) реакторов, которые используются для получения слоев разных материалов или слоев разного типа проводимости, а также подвижный подложкодержатель, перемещение которого между реакторами позволяет сформировать сверхрешетку. Выращивание слоев производится в вакууме порядка 10^{-6} мм рт. ст. Условия роста определяются температурами двух источников-испарителей, стенок и подложки. Для управления типом проводимости отдельных слоев используют два способа: регулируемое отклонение состава от стехиометрии, то есть легирование электрически активными собственными дефектами, и легирование примесями [53]. В первом случае в дополнительные резервуары реакторов загружают чистый теллур, давлением паров которого и определяется тип проводимости материала, так как избыток халькогена в PbTe дает дырочную проводимость, а избыток свинца — электронную. Уровень легирования при этом определяется температурой резервуара с теллуrom и существенно зависит от температуры подложки.

Перед выращиванием сверхрешетки на подложку BaF_2 обычно наносится буферный слой PbTe толщиной порядка 500 нм. Толщины слоев сверхрешетки для этого метода определяются двумя ограничивающими факторами: 1) нижняя граница определяется градиентом состава за счет взаимодиффузии компонентов; 2) верхняя граница, при превышении которой свойства сверхрешетки перестают отличаться от свойств объемного материала, определяется длиной свободного пробега носителей.

Этот метод выращивания применяется и для получения сверхрешеточных структур на основе $\text{A}^{\text{II}}\text{B}^{\text{VI}}$, например ZnS–ZnSe.

Молекулярно-лучевая эпитаксия (МЛЭ)

МЛЭ или МВЕ⁶ представляет собой процесс эпитаксиального роста слоев различных соединений, происходящий за счет реакций между термически создаваемыми молекулярными или атомными пучками соответствующих компонентов на поверхности подложки, находящейся в сверхвысоком вакууме при повышенной температуре. В основе этой технологии лежит возможность роста эпитаксиальных пленок, по существу, в динамическом режиме в отличие от более традиционных методов, где рост идет в условиях, близких к термодинамическому равновесию. МЛЭ характеризуется:

1) относительно низкой температурой роста (на 100–200°C меньше, чем при жидкофазовой и газотранспортной эпитаксии);

⁶МВЕ — molecular beam epitaxy.

Рассмотрим более подробно этот метод получения сверхрешеточных структур на примере выращивания сверхрешеток на основе GaAs. В работе [47] впервые было показано, что разложение газовой смеси металлорганического соединения триметилгаллия $(\text{CH}_3)_3\text{Ga}$ и гидрида мышьяка (арсина) AsH_3 , проводимое при температуре 600–700°C в атмосфере H_2 , можно использовать для выращивания тонких эпитаксиальных пленок GaAs в открытом (проточном) реакторе (рис. 9.11). Следует отметить, что MOCVD хорошо подходит для массового производства приборов с субмикронными слоями и имеет ряд преимуществ перед другими технологиями (см. ниже и [48, 49]).

Эпитаксия в процессе MOCVD осуществляется при пропускании однородной газовой смеси реагентов с газами-носителями над нагретой подложкой в реакторе с холодными стенками. Энергия для нагревания газовой смеси обычно создается мощным радиочастотным генератором с частотой порядка 450 кГц. Этот же генератор нагревает и графитовый держатель, на котором размещена монокристаллическая подложка. При этом газовая смесь вблизи поверхности подложки нагревается до высоких температур, а стенки реакционной камеры остаются относительно холодными, что приводит к осаждению полупроводниковой пленки на поверхности подложки (реакция идет при повышенных температурах) при малых потерях реагирующих веществ на поверхности реактора. Тщательно контролировать необходимо только температуру подложки, причем небольшие ее изменения не играют роли, поскольку большинство свойств пленок, полученных MOCVD, малочувствительны к таким изменениям температуры.

Парциальными давлениями различных газовых компонентов можно управлять с помощью электроники, контролирующей скорость потока от каждого из источников. Это позволяет с высокой точностью управлять всеми параметрами роста, обеспечивая надежную воспроизводимость результатов.

Металлорганические соединения, используемые для получения полупроводниковых пленок при комнатной температуре, как правило, являются жидкостями, обладают высоким давлением паров и могут быть легко доставлены в зону реакции путем пропускания газа-носителя, например H_2 , через жидкости, которые играют роль источников (рис. 9.11). Кроме того, используемые для роста полупроводниковых пленок гидриды при комнатной температуре являются газами и обычно берутся в качестве добавок к H_2 . Эти металлорганические и гидридные компоненты смешиваются в газовой фазе и пиролизуются (при 600–800°C, как правило) в потоке водорода в открытом реакторе при атмосферном или несколько

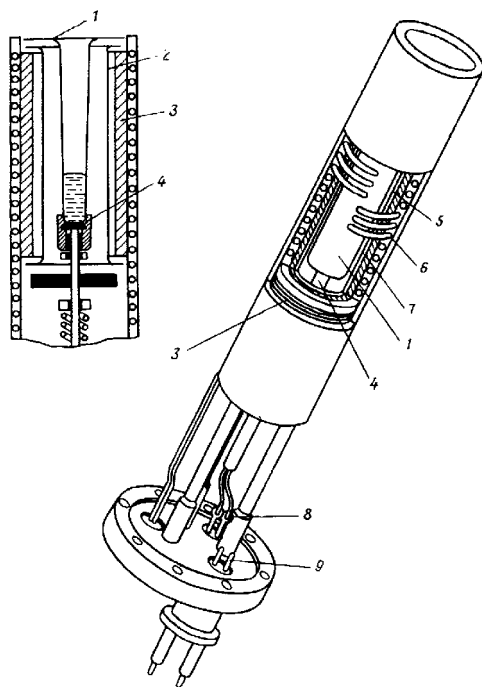


Рис. 9.16. Схематическое изображение эффузионной ячейки, используемой в современной системе МЛЭ: 1 — конический тигель из пиролитического BN, графита или стеклоуглерода; 2 — нагревательное устройство; 3 — радиационные экраны; 4 — термopара; 5 — оболочка нагревателя; 6 — водяное охлаждение; 7 — трубка из фольги; 8 — выводы термopары; 9 — силовой ввод.

зоны, в которых происходят различные физические явления. Первая — зона создания молекулярных пучков. Вторая — зона пересечения пучков от различных источников, где смешиваются испаряемые вещества, образуя паровую фазу, вступающую в контакт с подложкой, на которой происходит кристаллизация. Поверхность подложки, где происходят основные реакции между компонентами, является третьей зоной в системе МЛЭ.

Эффузионные ячейки, используемые в системах МЛЭ, обычно представляют собой либо цилиндрические, либо конические тигли с характерным диаметром 1–2 см, длиной 5–10 см, диаметром эффузионного отверстия 5–8 мм и его толщиной ≈ 0.5 мм. Для изготовления тиглей

используется пиролитический графит высокой чистоты или пиролитический нитрид бора BN. Тигель помещается внутрь нагревателя из окиси алюминия. В стандартных конструкциях эффузионных ячеек (рис. 9.16) проволочная спираль танталового нагревателя фиксируется двумя концентрическими цилиндрами из Al_2O_3 . Внутренний цилиндр особой формы содержит тигель, а внешний цилиндр окружается несколькими слоями тонкой гофрированной танталовой фольги для эффективной теплоизоляции отдельных ячеек. К тиглю подведена термопара, обеспечивающая контроль за температурой в эффузионной ячейке. Каждая эффузионная ячейка монтируется на молибденовом держателе, обеспечивающем возможность точной юстировки и быстрого демонтажа. Вся ячейка имеет общую водяную рубашку, эффективно теплоизолирующую ее от окружающей среды.

Первый теоретический анализ распределения потоков в молекулярном пучке принадлежит Кнудсену, который изучал эффузию из малого отверстия в эффузионной ячейке, содержащей идеальный газ. Идеальная кнудсеновская ячейка содержит пар, находящийся в равновесии с испаряемым веществом, и имеет эффективную апертуру в виде малого отверстия площадью S (отверстие не имеет стенок). По одну сторону отверстия находится газ под давлением p_1 , по другую сторону давление газа равно p_2 , причем $p_1 \gg p_2$. Линейные размеры отверстия не превосходят длин свободного пробега при обоих указанных давлениях. Газовые резервуары вне (p_2) и внутри (p_1) ячейки столь велики, что молекулы сталкиваются друг с другом значительно чаще, чем со стенками. Если справедливы следующие предположения: 1) молекулы имеют одинаковые точечные массы, 2) силы взаимодействия между молекулами отсутствуют, 3) распределение молекул по скоростям является максвелловским, 4) газ является изотропным, то полное число молекул N , прошедших в единицу времени через отверстие наружу, равно

$$N = ((p_1 - p_2)S/m)\sqrt{M/(2\pi RT)}, \quad (9.20)$$

где m — масса молекул, M — молекулярный вес, R — газовая постоянная, T — температура по обе стороны эффузионного отверстия.

Результаты, полученные Кнудсеном, полезны и при использовании эффузионных ячеек с отверстиями большого диаметра. При этом необходимо рассматривать большое отверстие как комбинацию малых точечных отверстий, а создаваемые ими потоки складывать. Однако при значительной толщине отверстия оно начинает оказывать коллимирующее воздействие на молекулярный пучок, что вызывает дополнительные осложнения.

В реальных системах МЛЭ редко используются идеальные кнудсеновские ячейки, так как необходимо проводить эпитаксиальный рост на большой площади. Реальные эффузионные ячейки имеют значительно большую апертуру, чем необходимо для равновесной ячейки. Поэтому в формулу (9.20) должен быть введен поправочный коэффициент, который, к сожалению, не может быть вычислен непосредственно. Он находится эмпирически.

Однородность толщины пленок, выращенных методом МЛЭ, зависит от однородности молекулярного пучка по площади подложки. Однородность молекулярного пучка определяется взаимным расположением ячеек и подложек и соосностью их расположения (см. гл. 6). Задача подбора оптимальной геометрии системы является достаточно сложной задачей, подходы к решению которой освещены в специальной литературе [5, 58].

О физических процессах, происходящих в третьей зоне (рост на подложке), рассказывалось в начале этой главы в разд. 9.2, посвященном механизмам роста эпитаксиальных пленок. Напомним основные этапы эпитаксиального роста полупроводниковых соединений, в том числе и для метода МЛЭ:

- 1) подвод питающего вещества к поверхности роста;
- 2) адсорбция составляющих вещество атомов и молекул на поверхности подложки;
- 3) миграция атомов и молекул по поверхности подложки, реиспарение;
- 4) встраивание составляющих вещество атомов и молекул в кристаллическую решетку, образование стабильных зародышей;
- 5) коалесценция зародышей.

Образование пленки на твердой подложке неразрывно связано с фазовым превращением, характерным для выбранного метода роста. В методе МЛЭ это фазовое превращение представляет собой переход из сильно неравновесной паровой фазы в тонкопленочную твердую фазу.

Анализ процессов, происходящих на поверхности роста, показал, что способность атомов двигаться по поверхности, их среднее время пребывания там, поток атомов и молекул на растущую поверхность пленки являются кинетическими параметрами, оказывающими решающее влияние на процессы роста при МЛЭ.

Для исследования особенностей роста в технологии МЛЭ используют соответствующие методы анализа. К этим методам относятся дифракция быстрых электронов, Оже-электронная спектроскопия (AES), масс-спектрометрия вторичных ионов (SIMS) и другие современные методы исследования. В настоящее время большинство исследовательских

и значительная часть технологических установок МЛЭ содержат встроенное перечисленное выше аналитическое оборудование, позволяющее проводить анализ процессов роста непосредственно в ходе выращивания эпитаксиальных пленок.

Дифракция быстрых электронов применяется для исследования поверхностной структуры, микроструктуры и степени гладкости выращиваемых пленок. Она используется также для контроля за изменением структуры верхних атомных слоев в ходе эпитаксиального роста пленок при различных параметрах роста и для наблюдения топографии поверхности.

Метод электронной Оже-спектроскопии используется для изучения химического состава поверхности подложки и определения химического состава растущего слоя.

Масс-спектрометрия вторичных ионов служит мощным методом определения химического состава поверхностных слоев твердого тела. Ее преимуществами перед Оже-спектроскопией является высокая абсолютная чувствительность по многим элементам, хорошая элементная избирательность и высокое разрешение по массам. Этот метод можно использовать для анализа объемного материала и исследования профиля поверхности, для определения остаточных газов в ростовой камере, анализа состава материалов источников и для изучения десорбционных явлений на поверхности эпитаксиального слоя.

Данные по поверхностной кинетике роста и легирования могут быть получены с помощью методики модулированного молекулярного пучка [59, 60]. В этой методике изменение падающего молекулярного пучка приводит к изменению во времени концентрации адсорбированных молекул на поверхности подложки и благодаря этому — к изменению скорости десорбции, что может быть измерено с помощью масс-спектрометрии. С другой стороны, модулируя поток десорбции, можно непосредственно определить, какие атомы и молекулы уходят с поверхности, выделив их на фоне остаточных газов. Дополнительное использование методов электронной Оже-спектроскопии и дифракции электронов позволяет получать дополнительные сведения о поверхности подложки на любой стадии процесса, а в совокупности с программированной термодесорбцией дает также дополнительную информацию о стехиометрии поверхности.

Разновидностью метода МЛЭ является лазерная МЛЭ, в которой испарение исходных компонентов осуществляется с помощью импульсного излучения мощного лазера [61, 62]. Этот метод был применен для выращивания сверхрешеток HgTe-CdTe . Пучок атомов ртути создается испарителем эффузионного типа. Поликристаллические образцы CdTe

и Те помещаются в танталовые лодочки, установленные на поворотной площадке, и поочередно подвергают воздействию сфокусированного лазерного луча. Особенности процесса лазерной МЛЭ, в первую очередь, связаны с более высокой температурой источников. В случае лазерной МЛЭ CdTe испарялся в виде атомов кадмия и теллура в отличие от обычной МЛЭ, при которой испарение происходило в виде атомов кадмия и молекул теллура. Таким образом, состав паровой фазы оказывался различным, что сказывалось на характере процесса эпитаксии, а именно ориентация подложки в случае лазерной МЛЭ оказывала существенно меньшее влияние на рост эпитаксиальной пленки, чем в случае обычной МЛЭ.

Литература

- [1] *К. фон Клитцинг*. УФН, 1986, т. 150, в. 1, с. 107.
- [2] Квантовый эффект Холла. Сб. статей под ред. Ю. В. Шмарцева, М., 1986.
- [3] *А. П. Силин*. УФН, 1985, т. 147, в. 3, с. 485.
- [4] *В. Т. Маслюк, П. А. Феннич*. Зарубежная электронная техника, 1981, N 8(241), с. 3.
- [5] *М. Герман*. Полупроводниковые сверхрешетки, М., Мир, 1989.
- [6] *В. Л. Бонч-Бруевич, С. Г. Калашников*. Физика полупроводников. Изд. 2-е. М., Наука, 1990.
- [7] *М. И. Клигер*. Аморфные диэлектрики и полупроводники, М., Изд. МИФИ, 1985.
- [8] *А. И. Попов, Н. И. Михалев*. Атомная структура некристаллических полупроводников, М., Изд. МЭИ, 1992.
- [9] *З. Метфессель, Д. Маттис*. Магнитные полупроводники, М., Мир, 1972.
- [10] *Э. Л. Нагаев*. Физика магнитных полупроводников, М., 1979.
- [11] *В. М. Фридкин*. Сегнетоэлектрики-полупроводники, М., 1976.
- [12] *М. Лайнс, А. Гласс*. Сегнетоэлектрики и родственные им материалы, М., 1981.
- [13] Органические полупроводники. Сб. статей под ред. М. В. Курика. Киев, 1976.

- [14] *Ф. Гутман, Л. Лайонс*. Органические полупроводники, М., 1970.
- [15] *К. Сайто*. Химия и периодическая таблица, М., Мир, 1982.
- [16] *Ч. Киттель*. Введение в физику твердого тела. Изд. 4-е. М., Наука, 1978.
- [17] *С. С. Горелик, М. Я. Дашевский*. Материаловедение полупроводников и металловедение. М., Metallurgia, 1973.
- [18] *Д. М. Васильев*. Физическая кристаллография. М., Metallurgia, 1981.
- [19] Физические величины: Справочник, под ред. И. С. Григорьева, Е. З. Мейлихова, М., Энергоатомиздат, 1991.
- [20] *R. C. Evans*. Crystal Chemistry. 2nd edition. Cambridge University Press, 1966.
- [21] *Ж. П. Сюше*. Физическая химия полупроводников. М., Metallurgia, 1969.
- [22] *С. А. Медведев*. Введение в технологию полупроводниковых материалов. М., Высшая школа, 1970.
- [23] *Б. Ридли*. Квантовые процессы в полупроводниках. М., Мир, 1986.
- [24] *Landolt-Bornstein*. Numerical Data and Functional Relationships in Science and Technology. New Series, v. 17, subv. a and b; Editor in Chief: K. H. Hellwege. Springer-Verlag, Berlin, Heidelberg, New York, 1982.
- [25] Полупроводниковые вещества. Вопросы химической связи. Сб. статей под ред. В. П. Жузе. М., Иностранная литература, 1960.
- [26] *Н. А. Горюнова*. Сложные алмазоподобные полупроводники. М., Советское радио, 1968.
- [27] *М. П. Шаскольская*. Кристаллография. М., Высшая школа, 1976.
- [28] *В. С. Вавилов, Н. П. Кекелидзе, Л. С. Смирнов*. Действие излучений на полупроводники, М., Наука, 1988.
- [29] *Ч. Уэрт, Р. Томсон*. Физика твердого тела. Изд. 2-е. М., Мир, 1969.

- [30] *Э. М. Омеляновский, В. И. Фистуль*. Примеси переходных металлов в полупроводниках. М., Металлургия, 1983.
- [31] *В. И. Фистуль*. Амфотерные примеси в полупроводниках. М., Металлургия, 1992.
- [32] *А. Милнс*. Примеси с глубокими уровнями в полупроводниках. М., Мир, 1977.
- [33] *G. W. Ludwig, H. H. Woodbury*. Phys. Rev. Lett., 1960, v. 5, p. 98.
- [34] *А. Б. Ройцин, Л. А. Фирштейн*. Теор. и эксп. химия, 1966, т. 2, с. 747.
- [35] Полупроводники. Под ред. Н. Б. Хэннея. М., ИЛ, 1962.
- [36] *В. И. Иверонова, А. А. Кацнельсон*. Ближний порядок в твердых растворах. М., Наука, 1977.
- [37] *Ю. М. Таиров, В. Ф. Цветков*. Технология полупроводниковых и диэлектрических материалов. М., Высшая школа, 1983.
- [38] *Ю. М. Шашков*. Металлургия полупроводников. М., Металлургия, 1960.
- [39] Технология СБИС. Под ред. С. Зи. М., Мир, 1986.
- [40] *И. В. Савельев*. Курс физики (в 5 кн.). М., Наука, 1998.
- [41] *Б. И. Болтакс*. Диффузия в полупроводниках. М., Наука, 1961.
- [42] *F. van der Maesen, J. Brenkmann*. J. Electrochem. Soc., 1955, т. 102, p. 229.
- [43] *В. В. Остробородова*. Основы технологии и материаловедения полупроводников. М., Изд. Моск. ун-та, 1988.
- [44] *Т. Р. Родин, Д. Уолтон*. Монокристаллические пленки. М., Мир, 1966.
- [45] *В. Б. Уфимцев, Р. Х. Акчурин*. Физико-химические основы жидкофазной эпитаксии. М., Металлургия, 1983.
- [46] *Н. М. Manasevit*. J. Cryst. Growth, 1981, v. 55, p. 1.
- [47] *Н. М. Manasevit*. Appl. Phys. Lett., 1968, v. 12, p. 156.

- [48] *P. Balk, E. Veuhoff*. J. Cryst. Growth, 1981, v. 55, p. 35.
- [49] *P. D. Dapkus*. J. Cryst. Growth, 1984, v. 68, p. 345.
- [50] *G. B. Stringfellow*. J. Cryst. Growth, 1984, v. 68, p. 111.
- [51] *M. R. Leys, C. van der Opdorp, M. P. A. Vieggers, H. Y. Talen, van der Mheen*. J. Cryst. Growth, 1984, v. 68, p. 431.
- [52] *H. Kinoshita, H. Fujiyasu*. J. Appl. Phys., 1980, v. 51, p. 5845
- [53] *H. Kinoshita, T. Sakashita, H. Fujiyasu*. J. Appl. Phys., 1981, v. 52, p. 2869.
- [54] *H. Clemens, E. J. Fantner, G. Bauer*. Rev. Sci. Instrum., 1983, v. 54, p. 685.
- [55] *I. A. Curless*. J. Vacuum Sci. Technol., 1985, v. 3, p. 531.
- [56] *Sc. M. Stanchak, H. Morkoc, L. C. Witkowski, T. J. Drummond*. Rev. Sci. Instrum., 1981, v. 52, p. 438.
- [57] *S. P. Svensson, T. G. Andersson*. J. Phys. E: Sci. Instrum., 1981, v. 14, p. 1076.
- [58] *M. A. Herman*. Vacuum, 1982, v. 32, p. 555.
- [59] *C. T. Foxon, B. A. Joyce*. Surf. Sci., 1977, v. 64, p. 293.
- [60] *E. A. Kurz, J. B. Hudson*. Appl. Surface Sci., 1984, v. 17, p. 485.
- [61] *J. M. Ballingall, D. J. Leopold, M. L. Wroge, D. J. Peterman, B. J. Morris, J. G. Broerman*. Appl. Phys. Lett., 1986, v. 49, p. 871.
- [62] *F. J. Boero, N. P. Ong, J. T. Cheung*. Solid State Commun., 1985, v. 54, p. 35.

Предметный указатель

- Акцепторы, 117
- Амфотерные примеси, 118
- Антиструктурные дефекты, 91
- Арсенид галлия
 - методы получения
 - MOCVD, 344
 - с помощью газотранспортных реакций, 343
 - образование химической связи, 62
 - энергетические уровни примесей, 132
- Атмосферы Коттрелла, 105
- Атом водорода
 - классификация состояний, 16
 - уравнение Шредингера, 16
 - энергия состояний, 16, 18
- Атомные радиусы, 52
 - ван-дер-ваальсов радиус, 57
 - ковалентный радиус, 53
 - металлический радиус, 55
- Вакансии, 89
- Вектор Бюргерса, 98
- Вицинальные грани, 183
- Включения второй фазы, 116
- Выращивание кристаллов из газовой фазы, 250
 - лимитирующая стадия, 254
 - массоперенос в вакууме, 251
 - метод сублимации-конденсации, 257
 - метод химических реакций, 259
 - метод химического транспорта, 261
- Выращивание кристаллов из расплава, 221
 - вертикальная зонная плавка, 232
 - вытягивание кристаллов из расплава (метод Чохральского), 225
 - горизонтальная зонная плавка, 232
 - лимитирующая стадия, 221
 - метод Бриджмена, 222
 - нормальная направленная кристаллизация, 222
- Выращивание кристаллов из раствора, 235
 - лимитирующая стадия, 236
 - метод движущегося растворителя, 237
 - методы роста, 237
- Газовая фаза, 323
- Германий
 - диффузия Cu, 305
 - диффузия Li, 304

- коэффициенты диффузии различных примесей, 301
- методы получения
 - жидкостная эпитакия, 337
 - метод химических реакций, 339
 - метод химического транспорта, 341
- предельная растворимость и коэффициент разделения, 282
- температурные зависимости растворимости примесей, 313
- энергетические уровни примесей, 121
- Двойники, 113
 - влияние условий выращивания, 244
- Деградация, 88
- Дефекты, 87
 - влияние условий выращивания, 239
 - неоднородности состава кристалла, обусловленные технологическими причинами, 245
 - двумерные, 87, 111
 - двойники, *см.* Двойники
 - дефекты упаковки, *см.* Дефекты упаковки
 - малоугловые границы, *см.* Малоугловые границы
 - линейные, *см.* Дислокации
 - объемные, 88, 114
 - включения второй фазы, 116
 - зональные упругие напряжения, 114
 - поры, 115
 - примеси, *см.* Примеси
 - связь с условиями выращивания
 - канальная неоднородность, 247
 - концентрационное переохлаждение, 248
 - собственные, 87
 - точечные, *см.* Точечные дефекты
 - Дефекты упаковки, 113
 - влияние условий выращивания, 244
 - Дефекты Френкеля, 89
 - Дефекты Шоттки, 89
 - Диаграмма состояния
 - $P - T - X$, 143, 167
 - конода, 145
 - линия ликвидуса, 145
 - линия солидуса, 145
 - правило рычага, 146
 - с неограниченной растворимостью компонентов, 145
 - типы диаграмм, 147
 - с ограниченной растворимостью компонентов, 155
 - с перитектическим превращением, 157
 - с химическими соединениями, 159
 - с эвтектическим превращением, 155
 - фигуративная точка, 145
 - Дивакансии, 91
 - Дислокации, 87, 95
 - вектор Бюргерса, 98
 - взаимодействие с точечными дефектами, 105

- влияние на свойства кристалла, 109
- влияние условий выращивания, 240
- линия винтовой дислокации, 97
- линия краевой дислокации, 96
- методы наблюдения, 110
- простые, 96
 - винтовые, 96
 - краевые, 96
- свойства, 100
 - механизмы движения, 101
- смешанные, 96, 97
- Диффузия, 283
 - восходящая диффузия, 286
 - газов, 312
 - гетеродиффузия, 284
 - влияние состава сплава и природы диффундирующего вещества на диффузию, 310
 - из бесконечно тонкого слоя, 294
 - из постоянного источника, 296
 - коэффициент диффузии, 284
 - методы определения, 297
 - механизмы, 286
 - вакансионный, 289
 - диссоциативный, 305
 - междоузельный, 287
 - обменный, 290
 - основные параметры, 292
 - самодиффузия, 284
 - влияние состава сплава и природы диффундирующего вещества на диффузию, 309
 - термодиффузия, 286
 - электродиффузия, 286
- Доноры, 117
- Зародыш, 173
- Зонная плавка, 232
- Изоэлектронные примеси, 120
- Инконгруэнтно плавящиеся соединения, 160
- Ионная связь, 30
 - поляризуемость ионов, 31
 - постоянная Маделунга, 30
 - пределы устойчивости структур, 32
- Ионные радиусы, 52, 54
- Ковалентная связь, 24
 - гибридные орбитали, 38
 - длина связи и углы между связями, 37
 - направленность, 37
 - насыщаемость, 37
 - правило Юм-Розери, 37
- Конгруэнтно плавящиеся соединения, 159
- Конденсация, 171
- Коэффициент разделения примесей, 195
 - равновесный, 196
 - в Ge и Si, 198
 - в полупроводниковых соединениях, 200
 - метод определения, 196
 - эффективный, 201
- Коэффициент термической аккомодации, 324
- Кремний
 - диффузия различных примесей, 301
 - методы получения

- метод химических реакций, 339
- метод химического транспорта, 342
- предельная растворимость и коэффициент разделения, 282
- энергетические уровни примесей, 121
- Кристаллизация, 171
- Кристаллизация в неравновесных условиях, 162
- Критический зародыш, 173
- Критическое пересыщение, 178
- Легирование, 264
 - в процессе роста из жидкой фазы, 266
 - методы выравнивания состава вдоль кристалла, 268
 - требования к чистоте легирующих примесей, 282
 - ионная имплантация, 265
 - радиационное, 264
- Ликвация внутрикристаллическая, 164
- Лимитирующая стадия роста, 221
- Линия ликвидуса, 145
- Линия солидуса, 145
- Малоугловые границы, 111
 - влияние условий выращивания, 243
- Междоузельные атомы, 89
- Метастабильная фаза, 166
- Методы определения коэффициентов диффузии, 297
 - метод электронно-дырочно-го перехода, 298
- Методы очистки, 190
- вертикальная бестигельная зонная плавка, 215
- зонная плавка, 205
- контроль чистоты материала, 215
- нормальная направленная кристаллизация, 205
- химические, 191
- Методы эпитаксии, 336
- Механизм роста
 - нормальный, 186
 - послойный, 184
 - слоисто-спиральный, 186
- Молекулярная связь
 - силы Ван-дер-Ваальса, 34
- Неограниченная растворимость компонентов друг в друге, 145
- Несингулярные грани, 183
- Образование зародышей новой фазы
 - гетерогенное, 180
- Образование центров новой фазы
 - гомогенное, 173
- Орбиталь
 - атомная, 17
 - d*-типа, 17, 19
 - f*-типа, 17, 19
 - p*-типа, 17, 19
 - s*-типа, 17, 19
 - молекулярная, 24
 - π -типа, 26
 - σ -типа, 26
 - разрыхляющая, 25
 - связывающая, 25
- Параметр ближнего порядка, 153
- Параметр дальнего порядка, 152

- Паровая фаза, 323
 Переохлаждение, 171
 Пересыщение, 171
 Поверхностный псевдоморфизм, 333
 Полупроводниковые соединения, 57
 закономерности образования бинарных соединений, 72
 правило Музера-Пирсона, 73
 правило нормальных валентностей, 73
 изоэлектронные ряды, 66
 равновесная ионность полупроводникового соединения, 63
 степень ионности полупроводникового соединения, 63
 химическая связь в $A^N B^{8-N}$, 62
 химические связи в полупроводниках $A^{IV} B^{VI}$, 81
 химические связи в полупроводниках $A_2^V B_3^{VI}$, 82
 химические связи в полупроводниках, производных от $A^N B^{8-N}$, 71
 соединения с дефектной структурой, 80
 соединения с резонансными связями, 78
 эффективные заряды, 64
 Получение кристаллов из жидкой фазы, 219
 Поры, 115
 Правило Клечковского, 22
 Правило Хунда, 22
 Примеси, 116
 в полупроводниковых соединениях, 132
 в элементарных полупроводниках, 120
 внедрения, 116
 замещения, 116
 комплексообразование, 137
 электрически активные, 116, 121
 акцепторы, 117
 амфотерные, 117
 доноры, 117
 многозарядные, 117
 простые, 117
 электрически неактивные, 116, 130
 изоэлектронные, 120
 Принцип Паули, 22
 Расплав, 219
 Распределение примесей при зонной плавке, 209
 Распределение примесей при нормальной направленной кристаллизации, 207
 Раствор, 219
 растворитель, 235
 растворяемое вещество, 235
 Растворимость, 278
 предельная, 278
 электрически активной примеси, 279
 ретроградная, 316
 выделение второй фазы, 318
 температурные зависимости, 313
 Растворимость газов, 312
 Рост эпитаксиальных пленок из газообразной фазы, 323

- механизмы, 323
 - пар (газ) — жидкость — кристалл, 329
 - пар (газ) — кристалл, 323
- молекулярно-кинетический подход, 327
- термодинамический подход, 326
- Сингулярные грани, 183
- Скорость образования центров новой фазы, 176
- Скрытая теплота кристаллизации, 142
- Соединения, 74
- Степень ионности связи, 60
- Твердые растворы, 74
 - внедрения, 75
 - вычитания, 75
 - замещения, 74
 - правило Вегарда, 76
 - упорядочение, 151
- Термодинамическая система
 - гетерогенная, 140
 - гомогенная, 140
 - изменение термодинамического потенциала при смешении, 148
 - компонент системы, 140
 - число степеней свободы, 142
- Точечные дефекты, 87, 89
 - антиструктурные, 91
 - вакансии, 89
 - влияние условий выращивания, 239
 - дивакансии, 91
 - комплексы, 91
 - междоузельные атомы, 89
 - механизмы образования, 92
 - по Френкелю, 89
 - по Шоттки, 89
- Тройные полупроводниковые соединения, 84
- закономерности образования, 85
- закономерности образования соединений с алмазоподобной структурой, 86
- Упорядочение в твердых растворах, 151
- Фазовое равновесие
 - диаграммы, 139
 - понятие термодинамической системы, 139
 - понятие фазы, 140
 - правило Гиббса, 141
- Химическая связь, 22
 - водородная связь, 36
 - ионная связь, 30
 - классификация, 24
 - ковалентная связь, 24
 - металлическая связь, 33
 - молекулярная связь, 34
 - энергия, 23
- Центр кристаллизации, 173
- Центр новой фазы, 173
- Электроотрицательность элементов, 57
 - потенциал ионизации, 58
 - сродство атома к электрону, 58
 - шкала по Малликену, 57
 - шкала по Полингу, 58
- Элементарные полупроводники, 42
- бор, 51

- висмут, 48
- германий, 46
- йод, 51
- кремний, 46
- мышьяк, 48
- олово, 46
- селен, 49
- сера, 49
- сурьма, 48
- теллур, 50
- углерод (алмаз), 43
- фосфор, 47
- Эпитаксия, 320
 - гетероэпитаксия, 320, 333
 - гомоэпитаксия, 320
 - жидкостная, 336
 - из газообразной фазы, 338
 - метод газотранспортных реакций, 341
 - метод химических реакций, 339
 - рост из газовой фазы с использованием металл-органических соединений (MOCVD), 344
 - методы, 336
 - рост из паровой фазы, 351
 - метод горячей стенки, 353
 - молекулярно-лучевая эпитаксия (MBE), 355
 - структурные дефекты в пленках, 335

Оглавление

1. Основные свойства и группы полупроводников	7
1.1. Отличительные свойства полупроводников	7
1.2. Основные группы полупроводниковых материалов	11
2. Основы теории химической связи	16
2.1. Электронная конфигурация свободных атомов	16
2.2. Типы химической связи	22
2.2.1. Ковалентная связь	24
2.2.2. Ионная связь	30
2.2.3. Металлическая связь	33
2.2.4. Молекулярная связь (силы Ван-дер-Ваальса)	34
2.2.5. Водородная связь	36
2.3. Основные свойства ковалентной связи	37
2.4. Элементарные полупроводники	42
2.4.1. Подгруппа IVA: углерод, кремний, германий, α -олово	42
2.4.2. Подгруппа VA: фосфор, мышьяк, сурьма и висмут .	47
2.4.3. Подгруппа VIA: сера, селен, теллур	49
2.4.4. Подгруппа VIIA: йод	51
2.4.5. Подгруппа IIIA: бор	51
2.5. Атомные и ионные радиусы	52
2.6. Полупроводниковые соединения	57
2.6.1. Электроотрицательность элементов	57
2.6.2. Степень ионности связи	60
2.6.3. Химическая связь в полупроводниковых соединениях $A^N B^{8-N}$	62
2.6.4. Металлическая составляющая связи, ионная составляющая связи и ширина запрещенной зоны в полупроводниках	65

2.6.5. Ионная составляющая связи и подвижность носителей заряда	68
2.6.6. Химические связи в полупроводниках, производных от $A^N B^{8-N}$	71
2.6.7. Полупроводниковые соединения других групп	81
2.6.8. Тройные полупроводниковые соединения	84
3. Дефекты в полупроводниковых материалах	87
3.1. Собственные дефекты	89
3.1.1. Точечные дефекты	89
3.1.2. Дислокации	95
3.1.3. Двумерные дефекты	111
3.1.4. Объемные нарушения	114
3.2. Примеси	116
3.2.1. Примеси в элементарных полупроводниках	120
3.2.2. Примеси в полупроводниковых соединениях	132
4. Фазовые равновесия и элементы теории образования фаз	138
4.1. Фазовые равновесия в полупроводниковых системах	139
4.1.1. Диаграммы фазового равновесия	139
4.1.2. $T - X$ диаграммы состояния бинарных систем	145
4.1.3. Кристаллизация в неравновесных условиях	162
4.1.4. $P - T - X$ диаграммы состояния	167
4.2. Элементы общей теории образования фаз	171
4.2.1. Образование центров новой фазы	173
4.2.2. Рост центров новой фазы	183
5. Получение чистых полупроводниковых материалов	190
5.1. Кристаллизационные методы очистки	190
5.2. Коэффициент разделения примесей	195
5.2.1. Равновесный коэффициент разделения	196
5.2.2. Эффективный коэффициент разделения примесей	201
5.3. Методы кристаллизационной очистки	204
5.3.1. Распределение примесей при нормальной направленной кристаллизации	207
5.3.2. Распределение примесей при зонной плавке	209
5.3.3. Модификация метода очистки зонной плавкой. Контроль чистоты материала и оценка содержания в нем примесей	213

6. Выращивание объемных кристаллов полупроводников	219
<i>Получение кристаллов из жидкой фазы</i>	219
6.1. Выращивание кристаллов из расплава	221
6.1.1. Методы нормальной направленной кристаллизации	222
6.1.2. Методы вытягивания кристаллов из расплава	225
6.1.3. Выращивание кристаллов методами зонной плавки	232
6.2. Выращивание монокристаллов из раствора	235
6.3. Условия выращивания и дефектность монокристаллов	239
<i>Выращивание кристаллов из газообразной фазы</i>	250
6.4. Метод сублимации–конденсации	257
6.5. Метод химических реакций	259
6.6. Метод химического транспорта (реакций переноса)	261
7. Легирование полупроводниковых материалов	264
7.1. Легирование объемных кристаллов из жидкой фазы	266
7.2. Методы выравнивания состава вдоль кристалла	268
7.2.1. Пассивные методы выравнивания состава	269
7.2.2. Активные методы выравнивания состава кристаллов	271
7.3. Растворимость примесей	278
8. Диффузия примесей в кристаллах	283
8.1. Возможные механизмы диффузии в твердых телах	286
8.1.1. Междоузельный механизм	287
8.1.2. Вакансионный механизм	289
8.1.3. Обменный механизм	290
8.2. Основные параметры диффузии	292
8.3. Распределение диффундирующих атомов по глубине	294
8.3.1. Диффузия из бесконечно тонкого слоя (из ограниченного источника)	294
8.3.2. Диффузия из постоянного источника	296
8.4. Экспериментальное определение коэффициентов диффузии	297
8.5. Диффузия в германии и кремнии	301
8.6. Зарядовое состояние диффундирующих примесей	304
8.6.1. Диффузия Li в Ge	304
8.6.2. Диффузия Cu в Ge. Диссоциативный механизм диффузии	305
8.7. Скорость и параметры диффузии	309
8.7.1. Самодиффузия	309
8.7.2. Гетеродиффузия	310
8.8. Диффузия и растворимость газов	312

8.9. Температурные зависимости растворимости примесей	313

9. Монокристаллические пленки	320

9.1. Эпитаксия	320
9.2. Рост эпитаксиальных пленок из газообразной фазы	323
9.3. Процессы зародышеобразования и структура пленок	330
9.4. Особенности гетероэпитаксии	333
9.5. Причины образования структурных дефектов	335
9.6. Методы эпитаксии	336
9.6.1. Жидкостная эпитаксия	336
9.6.2. Эпитаксия из газообразной фазы	338
9.6.3. Конденсация из паровой фазы	351