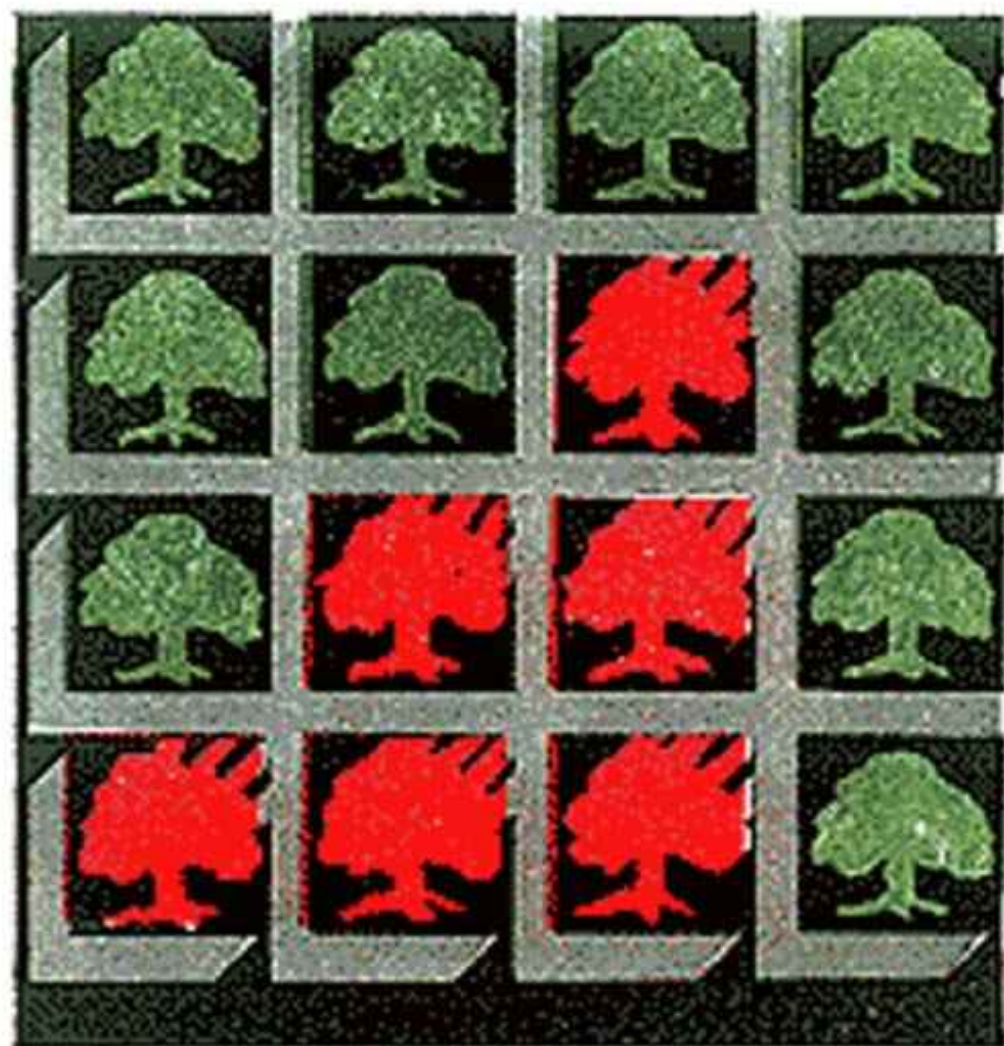


А.Л. ЭФРОС

ФИЗИКА И ГЕОМЕТРИЯ БЕСПОРЯДКА



：
"
" (" " , 19),
" " , .- . , 1982 .

PDF-
1543
, . 1999-2000 .

Предисловие

Наука, которая составляет предмет этой книги, очень молода. Её основные идеи были сформулированы лишь в 1957 г. в работе английских учёных Бродбента и Хаммерсли. Эта работа возникла следующим образом. В середине пятидесятых годов Вродбент занимался разработкой противогазовых масок для шахт по заданию Британского объединения по исследованию применений угля. При этом он столкнулся с интересной проблемой и привлёк к ней внимание математика Хаммерсли.

Основной элемент маски — это уголь, через который должен проходить газ. В угле есть поры, причудливо соединяющиеся друг с другом, так что образуется нечто вроде запутанного лабиринта. Газ может проникать в эти поры, адсорбируясь (осаждаясь) на их поверхности. Оказалось, что если поры достаточно широки и хорошо связаны друг с другом, то газ проникает в глубь угольного фильтра. В противоположном случае газ не проникает дальше поверхности угля. Движение газа по лабиринту представляет собой процесс нового типа, существенно отличающийся от хорошо известного в физике явления диффузии.

Бродбент и Хаммерсли назвали такие явления «*процессами протекания*» (по-английски percolation processes. В буквальном переводе слово percolation означает просачивание, фильтрацию; в русской научной литературе наряду с термином «протекание» можно встретить термин «перколяция», произошедший от английского слова). Теория, изучающая такого рода явления, стала называться «теорией протекания».

За 25 лет, прошедшие с первой работы Бродбента и Хаммерсли, выяснилось, что теория протекания необходима для понимания широчайшего круга явлений, относящихся, главным образом, к физике и химии. Вероятно, наиболее разработанной в настоящее время областью применения теории протекания являются электрические свойства неупорядоченных систем, таких, как аморфные полупроводники, кристаллические полупроводники с примесями или материалы, представляющие собой смесь двух разных веществ — диэлектрика и металла.

Явления, описываемые теорией протекания, относятся к так называемым «критическим явлениям». Эти явления характеризуются «критической точкой», в которой определенные свойства системы резко меняются. К критическим явлениям относятся также фазовые переходы второго рода (например, переход металла из нормального состояния в сверхпроводящее при понижении температуры). Физика всех критических явлений очень своеобразна и имеет общие черты, самая важная из которых состоит в том, что вблизи критической точки система как бы распадается на блоки с отличающимися свойствами, причём размер отдельных блоков неограниченно растёт при приближении к критической точке. Очертания блоков при этом случайны. В некоторых явлениях вся конфигурация хаотически меняется со временем за счёт теплового движения, в других явлениях она заморожена, но меняется при переходе от образца к образцу. Блоки расположены беспорядочно, так что глядя на мгновенную фотографию системы, трудно увидеть какие-либо закономерности. Однако «в среднем» эта геометрия, которую можно назвать «геометрией беспорядка», обладает

вполне определёнными свойствами.

Физические свойства всегда неразрывно связаны с геометрией. Например, физические свойства кристаллов определяются геометрией кристаллических решёток. Точно так же ряд свойств системы, находящейся вблизи критической точки, определяется «геометрией беспорядка». Самое интересное то, что благодаря большим размерам блоков, эта геометрия фактически не зависит от атомной структуры вещества и потому обладает универсальными свойствами, одинаковыми для многих, совершенно разных систем. Отсюда следует универсальность физических свойств, проявляющаяся в окрестности критических точек.

Такого рода связь между физикой и геометрией можно проследить в рамках теории протекания, и в этом — одна из главных задач книги. Теория протекания формулируется с помощью простых геометрических образов таких, как проволочные сетки, шарики, кристаллические решётки. Она не содержит понятия температуры и поэтому даёт возможность получить представление о критических явлениях читателям, не знакомым со статистической физикой.

Как и вся теория критических явлений, теория протекания не превратилась ещё в строгую с математической точки зрения науку. Многие важные утверждения остаются недоказанными, некоторые вопросы не выяснены. В тех случаях, где строгие доказательства существуют, но являются сложными, в книге эти доказательства заменены рассуждениями, скорее, объясняющими результат, чем доказывающими его. Однако автор всюду старался чётко отделить доказанные утверждения от недоказанных.

Книга содержит подробное изложение теории протекания и ее различных применений. Она построена следующим образом. Определение того, что называется теорией протекания и какие процессы она описывает, отложено до самой последней страницы книги. Это определение должно включать столько сложных понятий, что нет смысла давать его вначале. Почти каждая глава содержит какую-либо конкретную проблему, рассмотрение которой приводит к задаче теории протекания. Предполагается, что после нескольких глав читатель должен почувствовать, что общего между разными задачами теории протекания и какое отношение имеет к этому название книги.

Как правило, рассматриваемые проблемы представляют собой важные области применения теории протекания. Однако некоторые из них (проект фруктового сада в гл. 5, распространение слухов в гл. 11) носят иллюстративный и даже несколько иронический характер.

В книге излагаются необходимые для понимания материала сведения из элементарной теории вероятностей. В гл. 1 дается общее представление о вероятностях и случайных величинах. В гл. 2 излагаются правила сложения и умножения вероятностей и вводится функция распределения. Книгу можно прочитать в облегченном варианте, опустив гл. 2 и прочие главы и разделы, помеченные двумя звездочками. Правда, при этом читатель лишается возможности проследить за выводами некоторых количественных результатов, содержащихся в этих разделах, а также во многих упражнениях. Это, однако, не мешает пониманию (может быть, немного упрощенному) прочих разделов книги.

Упражнения, приведённые в книге, должны играть важную роль при освоении новых понятий. Как правило, они очень просты, и их рекомендуется выполнять, не заглядывая предварительно в раздел «Ответы и решения» (за исключением случаев, где это особо оговорено).

Важную роль в создании этой книги сыграл Б. И. Шкловский. Он участвовал в обсуждении её плана, названия и прочитал книгу в рукописи. Я чрезвычайно благодарен ему за это.

Я благодарен также своим коллегам Л. Г. Асламазову, Н. Б. Васильеву, Ю. Ф. Берковской, М. Э. Райху, прочитавшим рукопись книги и сделавшим ряд полезных замечаний.

Я благодарен художнику Б. А. Марской, сделавшей для книги несколько рисунков, в том числе рисунок на обложку. Особую благодарность я хочу выразить своей жене Н. И. Эфрос, взявшей на себя тяжёлый труд оформления рукописи.

А. Эфрос

Часть I

Задача узлов

Глава 1.

Порог протекания

1.1. Два учёных мужа кромсают экранную сетку

Не так уже часто в современных научных журналах появляются отчёты об экспериментах, объектом которых является, например, кусок обыкновенной экранной сетки, купленной с несколько необычной целью [1] (Прямое назначение экранной сетки состоит в том, чтобы защищать различную радиоаппаратуру от электрических помех.) в ближайшем магазине скобяных изделий. И хотя статья американских физиков Ватсона и Лиса, появившаяся в журнале «Физикл Ревью» за 1974 г. была далеко не первой работой в области теории протекания, наш рассказ начнется именно с неё.

Кусок сетки, с которым работали Ватсон и Лис, имел квадратную форму и содержал $137 \times 137 = 18769$ узлов с расстоянием $\frac{1}{4}$ дюйма = 6,35 мм между соседними узлами. Исследователи припаяли к двум противоположным сторонам квадрата медные электроды и включили сетку в электрическую цепь (рис. 1.1, а), чтобы измерить её сопротивление. Затем они стали блокировать отдельные узлы и изучать электрическое сопротивление в зависимости

от доли заблокированных узлов. Как показано на рис. 1.1, б, в, блокировка узла состояла лишь в том, что кусачками перерезались все четыре проволоки, которые связывались этим узлом.

Каждый новый узел, подлежащий блокировке, выбирался среди нетронутых ранее узлов случайно. В принципе, для этого можно было бы написать координаты каждого узла на отдельной бумажке, положить все бумажки в шапку, хорошенько перемешать и вынимать по одиночке. Однако при большом количестве узлов такая процедура (как и другие механические способы жеребьевки) крайне неудобна и потому ученые пользовались случайной последовательностью координат узлов, составленной ЭВМ. Ниже будет подробно описано, каким способом можно «заставить» ЭВМ генерировать случайные числа, а пока без всякого ущерба для понимания можно мысленно заменить ЭВМ шапкой.

Ясно, что по мере увеличения числа заблокированных узлов электропроводность сетки уменьшалась. (Электропроводностью называется величина, обратная сопротивлению. Сопротивление измеряется в омах (Ом), а электропроводность — в обратных омах (Ом^{-1}).) Более того, если обозначить через x отношение числа неблокированных узлов к полному числу узлов (137^2), то при некотором значении x , которое мы будем в дальнейшем называть *пороговым* (критическим) *значением* или *порогом протекания* и обозначать через x_c , электропроводность обращалась в нуль. Это происходило, когда перерезался последний путь, связывающий левый и правый электроды. Определение величины x_c и являлось одной из задач эксперимента. Было найдено, что $x_c = 0,59$.

Наверное, первый вопрос, который требует объяснения, состоит

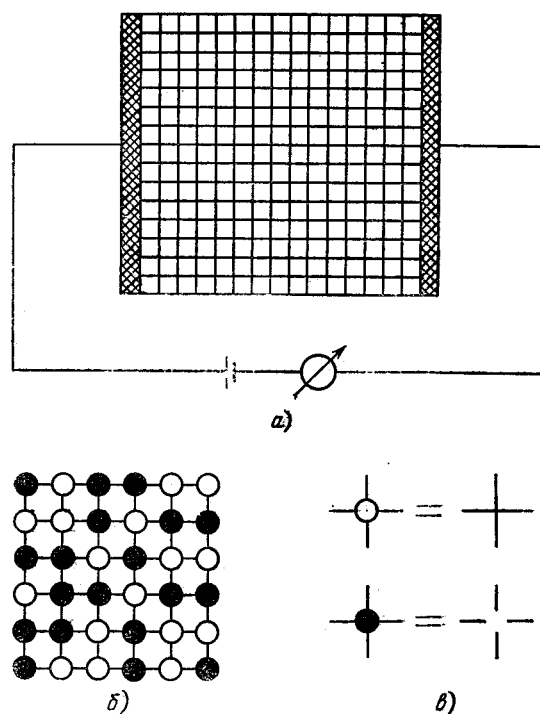


Рис. 1.1. Схема эксперимента Ватсона и Лиса. а) Исходная сетка. Количество узлов на рисунке сильно уменьшено; б) кусок сетки с заблокированными узлами. Блокированные узлы показаны чёрными кружками, а неблокированные — светлыми; в) чёрный узел означает разрыв контакта между четырьмя проволоками, которые связывает узел, светлый узел сохраняет контакт. Через чёрные узлы электрический ток не течёт ни в каком направлении, через светлые узлы ток течёт в любом направлении.

в том, является ли величина x_c случайной и невоспроизводимой от опыта к опыту или вполне определённой? Допустим, что мы повторили эксперимент, взяв другой кусок экранной сетки и воспользовавшись другой случайной последовательностью блокируемых узлов. Здравый смысл подсказывает, что поскольку на каждом этапе вся конфигурация заблокированных и целых узлов во втором эксперименте несколько не похожа на то, что было в первом экс-

перименте, то разрыв последнего пути, соединяющего электроды, тоже должен произойти при другом значении x , так что величина x_c , полученная во втором эксперименте, должна быть иной. Это, конечно, правильно.

Пороговое значение x_c в рассматриваемом эксперименте является *величиной случайной*. Поскольку такого рода величины будут фигурировать всюду на протяжении этой книги, полезно с самого начала пояснить:

1.2. Что такое случайная величина?

В математике так называют величину, о которой известно, какие значения она принимает и как часто она принимает то или иное значение, но неизвестно (и не может быть известно в рамках данной математической задачи), какое именно значение она примет в любом конкретном случае.

Вот классический пример случайной величины: бросают на стол шестигранный кубик с номерами на каждой грани. Случайной величиной является номер грани, оказавшейся сверху. Такая величина называется *дискретной*, потому что она принимает лишь некоторые определенные значения (в данном случае шесть значений — 1, 2, 3, 4, 5, 6). Предсказать заранее, какое именно число получится в данном опыте (т. е. при данном броске) невозможно, но можно предсказать вероятность того, что получится определенное число (например, 4). Допустим, что сделано некоторое число опытов, равное Q , причем число 4 выпало в Q_4 случаях. Отношение $\frac{Q_4}{Q}$ называется *относительной частотой появления данного*

значения случайной величины (числа 4). Если полное число опытов не очень велико, то это отношение колеблется: если сделать еще одну серию из Q опытов, то в новой серии отношение $\frac{Q_4}{Q}$ может быть существенно иным. Однако при увеличении числа опытов Q эти колебания становятся все меньше и меньше. Предел, к которому стремится относительная частота появления данного значения случайной величины, называют *вероятностью* этого значения.

Обозначим вероятность появления числа 4 через $P(4)$. Если кубик «честный», т. е. все его грани одинаковы, то величину $P(4)$ легко предсказать. Любая из шести граней кубика в среднем должна оказываться сверху одинаковое число раз, так что, если число Q велико, $\frac{Q_4}{Q} = \frac{Q_3}{Q} = \dots = \frac{1}{6}$. Таким образом, вероятности появления всех чисел одинаковы и равны $\frac{1}{6}$.

Итак, при очень большом числе бросков случайность уходит на задний план, оставляя место закономерности — симметрии граней кубика.

1.3. Среднее значение и дисперсия

Вернёмся к опыту с экранной сеткой. Было выяснено, что поскольку в опыте использовалась случайная последовательность заблокированных узлов, то критическая концентрация x_c , при которой прерывается ток между левым и правым электродами, также является случайной величиной, и предсказать, чему она равна в каждом конкретном эксперименте, заранее невозможно.

Теоретический подход к вопросу может состоять в том, чтобы изучить «средние» свойства величины x_c , т. е. свойства, про-

являющиеся в достаточно большом количестве экспериментов, выполненных при одинаковых условиях. Этими условиями являются, во-первых, полное число узлов сетки N (137^2 в описанном выше эксперименте), во-вторых, свойства генератора случайных чисел, задающего случайную последовательность блокируемых узлов. То, что свойства генератора не должны меняться от эксперимента к эксперименту, вовсе не означает, что последовательности блокируемых узлов должны быть одинаковыми. (Тогда были бы одинаковыми и все значения x_c !) Нужно лишь, чтобы во всех экспериментах использовался один и тот же способ создания случайной последовательности блокируемых узлов (например, шапка с бумажками).

Выполнив Q экспериментов с сеткой из N узлов, получим Q значений x_i , где i — номер эксперимента. Например, x_{15} обозначает x_c , которое возникло в пятнадцатом по счету эксперименте. Наиболее важной из средних величин является среднее арифметическое \bar{x}_Q , которое получится, если сложить все значения x_i и поделить сумму на число опытов Q :

$$\bar{x}_Q = \frac{x_1 + x_2 + \cdots + x_Q}{Q}. \quad (1.1)$$

Черта над буквой и индекс Q означают, что произведено усреднение по результатам Q опытов. Величина \bar{x}_Q все ещё является случайной. Если сделать еще одну серию из Q опытов в тех же самых условиях и по их результатам снова вычислить значение \bar{x}_Q , то оно окажется несколько иным. Однако чем больше число опытов в серии Q , тем меньше отличаются друг от друга средние значения, взятые по разным сериям. Дело в том, что *случайные колебания*

величин x_i в сумме компенсируют друг друга, так что с увеличением числа Q величина \bar{x}_Q стремится к вполне определённом значению, не зависящему от Q , но зависящему от условий, в которых делались опыты. Это предельное значение называется *средним значением случайной величины*. (В теории вероятностей это предельное значение называют также *математическим ожиданием* случайной величины, но мы не будем пользоваться этим термином.)

Среднее значение порога протекания, сетки из N узлов мы обозначим через $x_c(N)$. Величина $x_c(N)$ является не случайной, а *достоверной*. Её зависимость от N является закономерностью, над которой следует подумать.

Важной характеристикой случайной величины x_c служат также отклонения δ_i значений x_i от среднего значения:

$$\delta_i = x_i - x_c. \quad (1.2)$$

Сами отклонения δ_i меняются от одного эксперимента к другому, и следует выбрать величину, которая характеризовала бы их свойства «в среднем». Взять в качестве такой величины среднее арифметическое нельзя, так как в пределе $Q \rightarrow \infty$ оно стремится к нулю. Действительно,

$$\frac{\delta_1 + \delta_2 + \dots + \delta_Q}{Q} = \frac{x_1 + x_2 + \dots + x_Q}{Q} - x_c(N).$$

Но чем больше значение Q , тем меньше отличается первый член правой части этого равенства от второго, что и доказывает сделанное утверждение. Этот результат связан с тем, что отклонения

от среднего значения обязательно происходят как в ту, так и в другую сторону и в среднем компенсируют друг друга.

Можно было бы взять среднее арифметическое от неотрицательной величины $|\delta_i|$, однако общепринятый способ состоит в вычислении дисперсии $\delta^2(N)$, представляющей среднее арифметическое от квадратов отклонений (при $Q \rightarrow \infty$), которые, естественно, тоже являются неотрицательными числами:

$$\delta^2(N) = \frac{\delta_1^2 + \delta_2^2 + \dots + \delta_Q^2}{Q}. \quad (1.3)$$

Величина $\delta(N) = [\delta^2(N)]^{\frac{1}{2}}$ называется *средним квадратичным отклонением* случайной величины. Она и характеризует типичное отклонение величин x_i от их среднего значения $x_c(N)$. Естественно, что величина $\delta(N)$ также зависит от полного числа узлов сетки N .

Строго говоря, x_c является дискретной величиной, так как она получается делением числа неблокированных узлов на полное число узлов N и потому пробегает лишь те значения, которые превращаются в целые числа после умножения на N . Обозначим через x_k всевозможные различные значения случайной величины x_c .

Среднее значение $x_c(N)$ можно выразить через вероятности $P(x_k)$ того, что случайная величина x_c принимает значение x_k . Вспомним, что в формуле (1.1) суммирование ведется по всем значениям x_i , полученным в результате Q экспериментов. При этом все значения x_i могут встречаться по многу раз. Формулу (1.1)

можно записать в виде

$$\bar{x}_Q = \frac{x_1 Q_1 + x_2 Q_2 + \dots}{Q}, \quad (1.4)$$

где суммирование ведется по всем *различным* значениям x_k , которые может принимать величина x_c (ни одно значение x_k в этой сумме не повторяется дважды!). Число Q_k указывает, сколько раз встретилось значение x_k в серии из Q экспериментов.

Величина $\frac{Q_k}{Q}$ представляет собой относительную частоту появления значения x_k . При очень больших значениях Q она превращается в вероятность $P(x_k)$. По определению при больших Q левая часть формулы (1.4) превращается в $x_c(N)$. Поэтому

$$x_c(N) = x_1 P(x_1) + x_2 P(x_2) + \dots, \quad (1.5)$$

т. е. среднее значение равно сумме всех значений, которые может принимать случайная величина, помноженных на их вероятность. Аналогично

$$\delta^2(N) = (x_1 - x_c)^2 P(x_1) + (x_2 - x_c)^2 P(x_2) + \dots \quad (1.6)$$

Суммирование в формулах (1.5) и (1.6) ведётся по всем возможным значениям, которые может принимать случайная величина x_c , причем каждое значение встречается в сумме лишь один раз.

По определению $Q_1 + Q_2 + \dots = Q$, так что из определения $P(x_k)$ следует

$$P(x_1) + P(x_2) + \dots = 1. \quad (1.7)$$

Сумма вероятностей всех значений, которые может принимать случайная величина, равна единице.

1.4. Зачем нужна большая сетка?

В задаче о «честном» кубике было очень просто вычислить вероятность того, что случайная величина принимает то или иное значение. Свойства случайной величины x_c несравненно более сложные.

В конце следующей главы показано, как решается задача для сетки, представляющей квадрат из 2×2 узлов ($N = 4$). Результат состоит в том, что случайная величина x_c может принимать лишь два значения: $\frac{1}{2}$ и $\frac{1}{4}$. Первое значение она принимает с вероятностью $P(\frac{1}{2}) = \frac{2}{3}$, а второе значение — с вероятностью $P(\frac{1}{4}) = \frac{1}{3}$. В соответствии с формулами (1.5) и (1.6) (где суммы будут включать лишь два члена) $x_c(4) = \frac{5}{12}$, а $\delta(4) = \frac{\sqrt{2}}{12}$.

Набравшись терпения, можно решить и задачу с 3×3 узлами $N = 9$. Опыт показывает, что требуемые усилия чудовищно возрастают с увеличением стороны квадрата даже на один узел. Интерес же представляют в основном свойства сеток, содержащих очень большое число узлов (например, 10^{15}). Такие сетки могут служить моделью пленок, состоящих из атомов. Действительно, как правило, расстояние между атомами в конденсированных веществах (жидкостях, кристаллах) порядка $3 \cdot 10^{-8}$ см. Поэтому пленка площадью 1 см^2 , такая тонкая, что атомы занимают лишь один слой, состоит примерно из 10^{15} атомов.

Задача об определении вероятности того, что порог протекания для сетки из очень большого числа узлов N принимает то или иное значение, является центральной задачей теории протекания. В той или иной форме она обсуждается почти во всей книге.

Сейчас мы отметим, практически без доказательства, самое важное свойство этой задачи, являющееся ключом для понимания проблемы в целом:

Среднее квадратичное уклонение $\delta(N)$ уменьшается по степенному закону с ростом числа узлов N , стремясь к нулю при $N \rightarrow \infty$.

Это свойство выражается формулой

$$\delta(N) = \frac{C}{N^{\frac{1}{2}\nu}}, \quad (1.8)$$

где $C \approx 0,54$, а $\nu \approx 1,3$. (Величина ν называется *индексом радиуса корреляции*. Она подробно обсуждается в третьей части книги.)

Формула (1.8) не может быть результатом эксперимента Ватсона и Лиса. Чтобы ее получить, нужно было использовать сетки с разными значениями N и сделать много экспериментов при одном N . Кроме того, формула (1.8) является результатом теоретических исследований, о которых пойдет речь в третьей части книги.

Из формулы (1.8) следует, что чем больше узлов содержит сетка, тем меньше отличаются друг от друга результаты опытов, использующих разные случайные последовательности заблокированных узлов.

Почему это так? Дело в том, что в достаточно большой сетке встречаются всевозможные конфигурации целых и заблокированных узлов. В разных опытах они как бы меняются местами. Поэтому чем больше N , тем меньшую роль играет случайность. Бесконеч-

ная сетка содержит в себе бесконечное количество больших сеток, поэтому для нее случайность вообще не играет никакой роли, и величина x_c является не случайной, а достоверной, и равна

$$x_c = \lim_{N \rightarrow \infty} x_c(N).$$

Именно эта предельная величина и называется на самом деле порогом протекания. Ради нее и ставился эксперимент Ватсона и Лиса. А иначе зачем было брать сетку, содержащую почти 19000 узлов? Можно было взять сетку 2×2 !

Сформулируем теперь наиболее важный результат этой главы:

В бесконечной системе справедливо представление о четко определённом пороге протекания, не зависящем от того, какая случайная последовательность заблокированных узлов использовалась в эксперименте. В конечной системе чётко определенного порога не существует, а имеется так называемая критическая область с шириной порядка $\delta(N)$, в которую попадают значения x_c , полученные в большинстве экспериментов с различными случайными последовательностями. С увеличением размеров системы эта область стягивается в точку.

Следует, однако, иметь в виду, что зависимость от размера системы важна, только если мы пытаемся искусственно моделировать явление (например, с помощью экранной сетки). Как правило, теория протекания применяется к системам, у которых отдельные элементы имеют очень малые размеры. (Например, ими могут быть атомы (см. гл. 3). Как уже говорилось, 1 см^2 одноатомного слоя содержит $N = 10^{15}$ элементов, а 1 см^3 содержит $N = 10^{22}$! Такую систему с большой точностью можно считать бесконечной

и не обращать внимания на неопределенность порога протекания, связанную с ее размером.)

Задача, которую решали Ватсон и Лис, называется *задачей узлов* (в силу того, что именно узлы являются случайными элементами). К ней сводится целый ряд научных проблем, одна из которых (ферромагнетик с примесями) рассмотрена в гл. 3.

Точное значение порога протекания для этой задачи до сих пор неизвестно. Величину $x_c(N)$ определяют при больших значениях N с помощью ЭВМ или с помощью так называемых аналоговых экспериментов, похожих на эксперимент Ватсона и Лиса. (Их техника может быть весьма разнообразной.)

По изменению $x_c(N)$ с изменением N можно оценить, насколько близок полученный результат к искомому предельному значению. Сопоставление результатов, полученных разными методами, позволяет думать, что число 0,59 с написанной точностью (два знака после запятой) верно (хотя заранее совсем не очевидно, что число узлов $N = 137^2$ является для этого достаточным), но, разумеется, x_c можно бесконечно уточнять за счёт следующих знаков после запятой. (См. упражнение 4.)

Упражнения

1. Дискретную случайную величину a определим как номер грани шестигранного кубика, оказавшейся сверху после броска. Найдите среднее значение величины a .
2. Определим порог протекания как то́ значение x , при котором возникает протекание сверху вниз, а не слева направо. Изме-

няться ли при этом результаты отдельных опытов, $x_c(N)$, x_c ?
Сетку считать квадратной.

3. Тот же вопрос при условии, что порогом протекания названо минимальное значение x , при котором существует протекание и слева направо, и сверху вниз.
4. Тот же вопрос при условии, что порогом протекания названо максимальное значение x , при котором отсутствует протекание и слева направо, и сверху вниз.
5. Пользуясь формулой (1.8), вычислите среднее квадратичное уклонение, соответствующее условиям опыта Ватсона и Лиса ($N = 137^2$). На какую точность можно рассчитывать, если сделан только один опыт?

У к а з а н и е : В принципе результат одного опыта может отличаться от среднего значения $x_c(N)$ очень сильно. Однако, используя функцию распределения порогов протекания, приведённую ниже (формула (2.6)), можно доказать: вероятность того, что результат случайно выбранного опыта лежит в интервале от $x_c(N) - \delta$ до $x_c(N) + \delta$, равна примерно 0,7. Чем больше N , тем меньше δ и тем меньше «типичное отклонение» от среднего значения.

Глава 2.

Основные правила расчёта вероятностей и непрерывные случайные величины**

В этой книге, посвящённой закономерностям беспорядка, такие понятия, как вероятности и случайные величины, используются очень широко. Отчасти они уже были введены в предыдущей главе, так что читатель, не имеющий желания вдаваться в математическую сторону вопроса, может этим ограничиться и пропустить гл. 2, а также все последующие главы и разделы, помеченные двумя звёздочками. Читателю, который хочет проследить за решением ряда красивых математических задач, приведённых в книге, и составить более глубокое представление о теории протекания, нужно знать правила сложения и умножения вероятностей, изложенные в этой главе.

2.1. События и их вероятности

Понятие вероятности используется не только, когда речь идёт о численных значениях, принимаемых случайной величиной. Мож-

но обсуждать любые опыты со случайным результатом. Разные результаты опытов называют событиями. Относительной частотой появления события называют отношение числа опытов, которые привели к данному событию, к полному числу опытов. Вероятностью события называют предел, к которому стремится относительная частота появления события, при бесконечном увеличении числа опытов.

Пример. В ящик сложили одинаковое число красных, зелёных и синих шариков. Затем шарики перемешали и вынули наугад один из них. Какова вероятность события, состоящего в том, что вынутый шарик — красный? В противоположность опыту с кубиком в данном случае события отличаются не количеством, а качеством (цветом шарика). Однако рассуждать нужно по той же схеме. Так как число шариков каждого из трёх цветов одинаково, то красный шарик будет вынут в $\frac{1}{3}$ части всех опытов. Поэтому искомая вероятность равна $\frac{1}{3}$. Вероятности вынуть синий шарик и вынуть зелёный шарик также равны $\frac{1}{3}$.

Вероятность есть по определению величина, изменяющаяся от нуля до единицы. Нулевой вероятностью обладает, например, событие, состоящее в том, что из ящика, в котором имеются только красные шарики, вынули синий шарик. Вероятность вынуть из того же ящика красный шарик равна единице. Событие, вероятность которого равна единице, называется уже не случайным, а **достоверным**.

Понятие вероятности играет огромную роль при выяснении закономерностей мира случайных процессов. Очень часто закономерность оказывается буквально погребённой под случайно-

стью. Представьте, что по сведениям какого-то родильного дома вы пытаетесь установить закономерности в рождении мальчиков и девочек. Перед вами случайная последовательность типа МДММММДМДД . . . Иногда вам кажется, что мальчики рождаются гораздо чаще, иногда — наоборот. Ваш приятель заверит вас, что «сейчас рождаются только девочки». Это может быть связано с тем, что девочки родились в трёх знакомых ему семьях.

Однако закономерность есть. Вероятность рождения мальчика относится к вероятности рождения девочки как 51,5 к 48,5. В таких больших странах, как СССР или США, эта закономерность неплохо выполняется, даже если взять данные лишь за один год.

В отличие от задачи с разноцветными шариками, задачу о рождении девочек и мальчиков очень трудно решить теоретически. Однако статистические данные, о которых идёт речь, отражают совершенно определённые и хорошо изученные свойства человеческой физиологии.

2.2. Сложение вероятностей

События называются *несовместимыми*, если они не могут наблюдаться в одном и том же единичном опыте. Например, событие, состоящее в том, что вынут красный шарик, несовместимо с событием, состоящим в том, что вынут синий шарик, потому что по условию задачи в одном опыте можно вынуть лишь один шарик: красный, синий или зелёный. Несовместимы события, состоящие в том, что бросок кубика даст число 5 и число 2.

Докажем два важных свойства вероятностей.

1. *Правило сложения. Вероятность того, что произойдёт какое-либо одно (безразлично какое) из нескольких несовместимых событий, равна сумме вероятностей этих событий.* Пусть нужно найти вероятность того, что бросок кубика даст либо число 3, либо число 4. Число опытов, в которых возникли интересующие нас значения, равно сумме числа опытов, давших значение 3, и числа опытов, давших значение 4. По определению, для нахождения искомой вероятности эту сумму следует поделить на полное число опытов Q и перейти к пределу при $Q \rightarrow \infty$. Поскольку предел каждого члена суммы, поделённого на Q , равен вероятности получения одного из интересующих нас чисел, то искомая вероятность действительно равна сумме вероятностей получения каждого из чисел. Таким образом, вероятность того, что выпадет либо 3, либо 4, равна $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$. Вероятность того, что выпадет либо 1, либо 2, либо 3, либо 4, равна $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{2}{3}$. Вероятность же того, что выпадет либо 1, либо 2, либо 3, либо 4, либо 5, либо 6 равна $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$. Этот результат является частным случаем второго свойства.

2. *Назовём полной системой событий такую совокупность несовместимых событий, которая исчерпывает возможные результаты данного опыта.* Например, в опыте с кубиком полной системой событий являются события, состоявшие в том, что получатся числа 1, 2, 3, 4, 5, 6. Второе свойство состоит в следующем:

Сумма вероятностей событий, образующих полную систему,

равна единице. Согласно первому свойству эта сумма равна вероятности того, что произойдёт какое-то одно из событий, образующих полную систему. Но по определению полной системы одно из них обязательно произойдёт. (В примере с кубиком это означает, что какое-нибудь число из шести возможных обязательно выпадет.) Событие, которое обязательно происходит, является достоверным, и его вероятность равна единице. Это и доказывает второе свойство. (В случае кубика оно гласит, что сумма вероятностей шести возможных значений равна единице.)

В применении к вероятностям различных значений, которые может принимать любая случайная величина, это свойство было сформулировано в виде формулы (1.7).

В некоторых случаях оговорка относительно несовместимости событий может оказаться очень существенной при применении правила сложения вероятностей. Рассмотрим следующий пример.

Пример. Пять стрелков стреляют в цель одновременно. Квалификация стрелков одинакова: каждый из них поражает цель с вероятностью $\frac{1}{3}$. Какова вероятность того, что в цель попадёт хотя бы один из них?

Нужно найти вероятность того, что произойдёт одно из пяти событий (всё равно какое): в цель попадёт первый стрелок, второй стрелок и т. д. Возникает мысль воспользоваться правилом сложения вероятностей. Согласно этому правилу вероятность того, что

попадёт один из стрелков, равна сумме вероятностей:

$$P = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{5}{3}.$$

Результат получился явно нелепый. Вероятность оказалась больше единицы, что совершенно бессмысленно. Где же ошибка? Вспомним, что правило сложения формулируется лишь для несовместимых событий. А разве не могут сразу несколько стрелков попасть в цель? Конечно могут. Это типичный пример совместимых событий. Поэтому и нельзя пользоваться правилом сложения.

Для того чтобы решить задачу о стрелках, нужно воспользоваться правилом умножения вероятностей, сформулированным в следующем разделе.

2.3. Умножение вероятностей

Мы снова бросаем шестигранный кубик. Вопрос, на который нужно теперь ответить, состоит в следующем. Бросаем кубик два раза подряд и получаем два числа. Какова вероятность того, что этими числами окажутся 6 и 4, причём именно в такой последовательности: сначала 6, а потом 4?

Схема решения обычная. Делается Q опытов (каждый — из двух бросков) и определяется число опытов, давших нужный результат. Сначала отберём те опыты, в которых первый бросок дал 6, а второй — что угодно. С этой задачей мы уже знакомы. Поскольку все грани кубика — одинаковые, то при первом броске число 6 (как и любое другое от 1 до 6) появилось в $\frac{1}{6}$ части опытов, т. е. на первом этапе мы отобрали $Q_1 = \frac{Q}{6}$ опытов. (Имеется

в виду, что число Q очень велико, так что случайные отклонения от величины Q_1 малы.) Теперь надо отобрать те из них, в которых второй бросок дал 4. Во втором броске появление всех чисел тоже равновероятно. Поэтому число 4 появилось в $\frac{1}{6}$ части опытов. Итак, число опытов, в которых после числа 6 появилось число 4, равно $Q_2 = \frac{1}{6} \cdot \frac{1}{6} \cdot Q$, а вероятность этого события равна

$$\frac{Q_2}{Q} = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}.$$

Усложним задачу. Пусть опыт состоит из трёх бросков и нужно найти вероятность того, что этот опыт дал три числа в определённой последовательности, например, 4, 5, 1 или 6, 6, 6. Рассуждая подобным образом, найдём, что число опытов, дающих искомый результат, равно $Q_3 = \frac{1}{6}Q_2 = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}Q$, а вероятность этого результата есть $\frac{Q_3}{Q} = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$.

Разберём ещё один пример. Допустим, что на каждые десять тысяч велосипедов, выпускаемых некоторым заводом, один велосипед имеет дефект передней втулки и два — дефект задней втулки. То есть вероятность того, что выбранный наугад велосипед имеет бракованную переднюю втулку, равна $\frac{1}{10000}$, а задней — $\frac{2}{10000}$. Допустим, что передние и задние втулки изготавливают в разных цехах и наличие дефекта в одной из них не увеличивает и не уменьшает вероятность дефекта в другой. Нужно найти вероятность того, что выбранный наугад велосипед имеет дефекты обеих втулок. Рассуждать нужно, как и в предыдущих случаях. Из Q велосипедов отберём имеющие дефект передней втулки. Их число равно $\frac{Q}{10000}$. Из них выберем велосипеды с дефектом и задней втулки. Получим $\frac{Q}{10000} \cdot \frac{2}{10000}$. Искомая вероятность равна $\frac{1}{10000} \cdot \frac{2}{10000} = 2 \cdot 10^{-8}$.

В обоих примерах были заданы вероятности нескольких событий и требовалось найти вероятность того, что эти события наступят совместно, т. е. в одном опыте. Полученные результаты можно сформулировать в общем виде:

Вероятность совместного наступления нескольких событий равна произведению вероятностей этих событий.

К этому правилу нужно сделать важное дополнение. Во всех примерах фигурировали *независимые* события. Два события называются независимыми, если осуществление одного из них не влияет на вероятность осуществления другого. Например, то обстоятельство, что при первом броске кубика выпало число 6, никак не влияет на вероятность того, что во втором броске выпадет 4, существование брака передней втулки не сказывается на вероятности брака задней втулки.

Легко понять, что независимость событий очень важна для вывода правила умножения вероятностей.

Рассмотрим снова пример с велосипедами и допустим, что независимость событий нарушена следующим образом. Передние и задние втулки каждого велосипеда собираются одновременно, причём в определённые дни вероятность брака больше, чем в другие дни. Тогда наличие брака одной из втулок увеличивает вероятность брака второй, поскольку увеличивается вероятность того, что данный велосипед изготовлялся в неудачные дни. Поэтому возрастает вероятность брака обеих втулок.

Чтобы лучше это понять, рассмотрим крайний случай: допустим, весь брак делается в определённые дни. Все велосипеды, выпущенные в эти дни, имеют брак задней втулки, а половина их

имеет к тому же брак передней втулки. Тогда вероятность того, что велосипед, выбранный наугад среди продукции целого года, имеет брак обеих втулок, равна вероятности брака передней втулки, т. е. равна $\frac{1}{10000}$, а вовсе не $2 \cdot 10^{-8}$. Итак, правило умножения вероятностей справедливо только для независимых событий.

Правило умножения позволяет легко решить задачу о пяти стрелках, сформулированную в предыдущем разделе. Напомним условие: пять стрелков одновременно стреляют в цель, причём вероятность попасть для каждого из них равна $\frac{1}{3}$. Найти вероятность того, что в цель попадёт хотя бы один из них.

Проще всего можно решить задачу, найдя вероятность того, что ни один из стрелков не попадёт в цель (обозначим эту вероятность через P_0). Так как попадания различных стрелков в цель следует считать независимыми событиями, вероятность P_0 равна произведению вероятностей того, что промажет каждый из стрелков. Событие, состоящее в том, что некоторый стрелок попадёт в цель, и событие, состоящее в том, что он промажет, составляют полную систему событий. Сумма вероятностей двух этих событий равна единице. Если вероятность того, что стрелок попадёт в цель, равна $\frac{1}{3}$, то вероятность того, что он не попадёт, равна $1 - \frac{1}{3} = \frac{2}{3}$. Вероятность того, что промажут все пять стрелков, равна

$$P_0 = \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} = \left(\frac{2}{3}\right)^5.$$

Событие, состоящее в том, что не попадёт в цель ни один из стрелков, и событие, состоящее в том, что попадёт хотя бы один, образуют полную систему событий. Сумма вероятностей этих событий равна единице. Поэтому искомая вероятность P удовлетворяет

уравнению

$$P + P_0 = 1,$$

откуда следует, что

$$P = 1 - P_0 = 1 - \left(\frac{2}{3}\right)^5 \approx 0,87.$$

Упражнения

1. В сетке, содержащей N узлов, заблокированы N' узлов. Какова вероятность того, что случайно выбранный узел окажется заблокированным? Неблокированным?
2. Найдите вероятность того, что три броска кубика, произведённых подряд, дали три числа 1, 2, 3 в любой последовательности. Дали числа 1, 2, 2 в любой последовательности.
3. Первый цех выпускает хорошие детали с вероятностью 0,8, а второй цех — с вероятностью 0,9. Взяли наугад 3 детали первого цеха и 4 детали второго. Найдите вероятность того, что все 7 деталей хорошие.

2.4. Порог протекания в сетке 2×2

Сообщённых выше сведений из теории вероятности вполне достаточно, чтобы исследовать задачу о протекании в квадратной сетке из 4 узлов ($N = 4$).

Схема опыта с сеткой 2×2 изображена на рис. 2.1. Номера четырёх узлов написаны на отдельных бумажках, бумажки положены в шапку и перемешаны. Допустим, что в первый раз вытащена бумажка с номером 1 и блокирован первый узел (рис. 2.1, б). (Схема рассуждений и конечные результаты несколько не меняются, если в первый раз блокирован узел с другим номером. Дело в том, что в сетке из четырёх узлов все узлы занимают равноправные положения.) Если во второй раз будет блокирован узел 2, то ток не прекратится (рис. 2.1, в) — он будет течь по нижнему проводу. После того, как будет блокирован третий узел (3 или 4), ток, конечно, прекратится и будет зафиксировано, что критическая доля неблокированных узлов равна $\frac{1}{4}$. Если же во второй раз будет блокирован узел 3 или 4, то ток прекратится и критическая доля окажется равной $\frac{1}{2}$ (рис. 2.1, г, д). Таким образом, порог протекания x_c является дискретной случайной величиной, принимающей значения $\frac{1}{4}$ и $\frac{1}{2}$. Вычислим вероятность P того, что она принимает каждое из этих значений: $P(\frac{1}{4})$ и $P(\frac{1}{2})$.

Всё зависит от того, какой узел будет блокирован вторым. Если узел 2, то $x_c = \frac{1}{4}$, если 3 или 4, то $x_c = \frac{1}{2}$.

Таким образом, вероятность $P(\frac{1}{4})$ равна вероятности того, что вторым окажется узел 2, а $P(\frac{1}{2})$ равна вероятности того, что вторым будет либо узел 3, либо узел 4. После того, как блокирован узел 1, все три оставшихся узла имеют одинаковую вероятность быть блокированными в следующий раз. Сумма трёх вероятностей равна единице, поскольку эти три события образуют полную систему. Следовательно, каждая из вероятностей равна $\frac{1}{3}$.

Итак, вероятность того, что следующим блокированным ока-

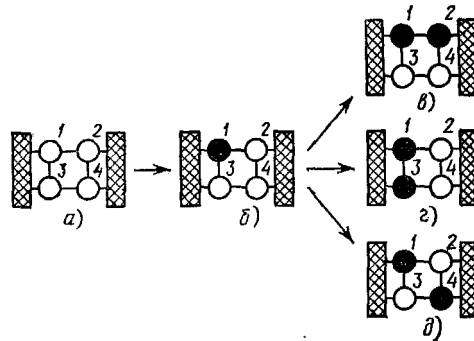


Рис. 2.1. Расчёт сетки 2×2 . а) Исходная сетка; б) блокирован один узел; в, г, д) блокированы два узла. В случае в ток прерывается лишь после того, как блокирован третий узел, так что $x_c = \frac{1}{4}$. В случаях г и д ток прерывается после того, как блокирован второй узел, так что $x_c = \frac{1}{2}$. Три случая в, г и д равновероятны.

жется узел 2, равна $\frac{1}{3}$. Но если узел 2 — следующий, то $x_c = \frac{1}{4}$. Следовательно, вероятность того, что $x_c = \frac{1}{4}$, равна $\frac{1}{3}$, т. е. $P(\frac{1}{4}) = \frac{1}{3}$. Теперь надо найти вероятность того, что во второй раз будет блокирован либо узел 3, либо узел 4. Согласно правилу сложения вероятностей она равна сумме вероятностей: $\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$. Это и есть вероятность того, что x_c принимает значение $\frac{1}{2}$, таким образом, $P(\frac{1}{2}) = \frac{2}{3}$. Так как возможны всего два значения x_c , то должно выполняться равенство $P(\frac{1}{2}) + P(\frac{1}{4}) = 1$. Действительно, $P(\frac{1}{2}) + P(\frac{1}{4}) = \frac{2}{3} + \frac{1}{3} = 1$.

Легко предсказать среднее значение порога протекания $x_c(4)$. Согласно формуле (1.5), $x_c(4) = \frac{1}{2} \cdot P(\frac{1}{2}) + \frac{1}{4} \cdot P(\frac{1}{4}) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3} = \frac{5}{12}$. Это число довольно сильно отличается от порогового значения $x_c = \lim_{N \rightarrow \infty} x_c(N)$, которое, как уже говорилось, равно $\approx 0,59$.

Легко вычислить дисперсию порога протекания. Согласно фор-

муле (1.6),

$$\delta^2(4) = \left(\frac{1}{2} - \frac{5}{12}\right)^2 \cdot \frac{2}{3} + \left(\frac{1}{4} - \frac{5}{12}\right)^2 \cdot \frac{1}{3} = \frac{1}{72}.$$

Среднее квадратичное уклонение

$$\delta(4) = \frac{\sqrt{2}}{12}$$

Упражнение

4. Повторите рассуждения, предположив, что в первый раз был блокирован узел 3.

2.5. Непрерывная случайная величина

До сих пор обсуждались дискретные случайные величины. Однако бывают и *непрерывные* случайные величины, которые могут принимать любое значение на некотором отрезке числовой оси.

Допустим, что случайная величина a может принимать все значения y , лежащие в области от A до B ($A \leq y \leq B$), однако, одни значения она принимает часто, а другие редко. Чтобы описать это математически, вводят функцию распределения $f(y)$ случайной величины a .

Основное свойство функции распределения состоит в следующем: если точки A_1 и B_1 лежат внутри интервала (A, B) , причём $A_1 < B_1$, то вероятность того, что значение случайной величины окажется в интервале $A_1 \leq y \leq B_1$, равно площади, ограниченной графиком функции $f(y)$, осью абсцисс и перпендикулярами,

восстановленными в точках A_1 и B_1 (на рис. 2.2 эта площадь заштрихована). Те, кто знаком с интегральным исчислением, поймут, что эта вероятность (обозначим её через $P(A_1, B_1)$) выражается формулой

$$P(A_1, B_1) = \int_{A_1}^{B_1} f(y) dy.$$

Так как все значения случайной величины находятся в интервале (A, B) и какое-то из них она обязательно принимает, то искомая площадь равна единице. Иными словами,

$$P(A, B) = \int_A^B f(y) dy = 1. \quad (2.1)$$

Это равенство иногда называют условием нормировки функции распределения.

Фигура, площадь которой выражается интегралом (2.1), называется криволинейной трапецией (см. рис. 2.2). Если интервал (A_1, B_1) столь мал, что функция распределения внутри этого интервала практически не успеваает измениться, то криволинейную трапецию можно с успехом заменить прямоугольником с высотой $f(y_1)$, где y_1 — любая точка из интервала (A_1, B_1) . Тогда

$$P(A_1, B_1) = f(y_1) \Delta, \quad (2.2)$$

где $\Delta = B_1 - A_1$ представляет собой ширину интервала (A_1, B_1) .

В математической литературе функцию $f(y)$ называют *плотностью вероятности*. Как видно из формулы (2.2), при малой ширине

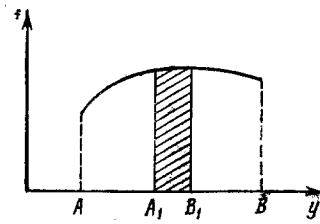


Рис. 2.2.

интервала (только тогда и применима эта формула!) вероятность того, что случайная величина окажется внутри интервала, прямо пропорциональна ширине интервала. Функция $f(y)$ представляет собой вероятность, поделённую на ширину интервала, или вероятность, отнесённую к единице длины интервала, или, иными словами, плотность вероятности. Однако физики часто предпочитают термин «функция распределения».

Формулы (1.5) и (1.6) для среднего значения и для дисперсии в случае непрерывной случайной величины перепишем в виде

$$\bar{a} = \int_A^B y f(y) dy, \quad (2.3)$$

$$\delta^2 = \int_A^B (y - \bar{a})^2 f(y) dy, \quad (2.4)$$

где \bar{a} — среднее значение непрерывной случайной величины a .

Приведём пример функции распределения.

Равномерное распределение. Непрерывная случайная величина с одинаковой вероятностью принимает все значения от нуля до единицы и не может принимать никаких других значений.

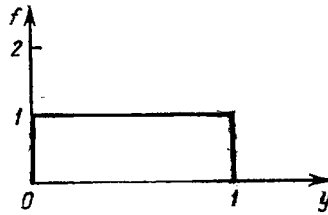


Рис. 2.3.

Очевидно, что функция $f(y)$ не зависит от y внутри интервала $(0, 1)$ и равна нулю вне этого интервала (рис. 2.3). Её значение внутри интервала легко найти из условия нормировки (2.1). В данном случае $A = 0$, $B = 1$, а криволинейная трапеция превратилась в прямоугольник, площадь которого равна $f_0 \cdot 1$, где f_0 — значение функции внутри интервала, а ширина интервала равна 1. Из условия нормировки следует, что $f_0 \cdot 1 = 1$, т. е. $f_0 = 1$. Итак,

$$f(y) = \begin{cases} 1 & \text{при } 0 \leq y \leq 1, \\ 0 & \text{при } y > 1 \text{ и } y < 0. \end{cases} \quad (2.5)$$

Упражнение

5. Непрерывная случайная величина a с равной вероятностью принимает все значения от -1 до $+1$. Найдите вероятность того, что она попадёт в интервал от $-\frac{3}{4}$ до $-\frac{1}{4}$.

2.6. Порог протекания как непрерывная случайная величина

Порог протекания является, строго говоря, дискретной случай-

ной величиной, так как все значения, которые он может принимать, превращаются в целые числа после умножения на полное число узлов N . Однако при больших значениях N разность ближайших возможных значений этой случайной величины очень мала (она равна N^{-1}). Поэтому в наиболее важном случае, когда число узлов велико, порог протекания x_c можно с хорошей точностью рассматривать как непрерывную случайную величину, принимающую всевозможные значения, расположенные внутри некоторого интервала на числовой оси. Тогда величину x_c следует характеризовать функцией распределения $f(y)$. В этом разделе описано, как выглядит функция $f(y)$ при больших значениях N .

Функция распределения пороговых значений x_c должна зависеть от числа узлов сетки N , с которой ведутся опыты. Поэтому будет правильнее обозначать функцию распределения через $f_N(y)$. Под величиной y удобно понимать не само пороговое значение, а его отклонение от среднего значения $x_c(N)$. Тогда $f_N(y)\Delta$ представляет вероятность того, что пороговое значение, полученное в некотором опыте, отличается от среднего значения $x_c(N)$ на величину, лежащую в малом интервале Δ около значения y . По определению среднее значение, вычисленное с помощью функции $f_N(y)$ по формуле (2.3), равно нулю.

На рис. 2.4 представлена функция $f_N(y)$ при трёх различных значениях N . Как видно из рисунка, с увеличением числа узлов N функция распределения становится всё более острой. Это означает, что отклонения от среднего значения (напомним, что оно принято равным нулю!) становятся с ростом N всё менее и менее вероятными. Согласно формуле (2.1) предыдущего раздела пло-

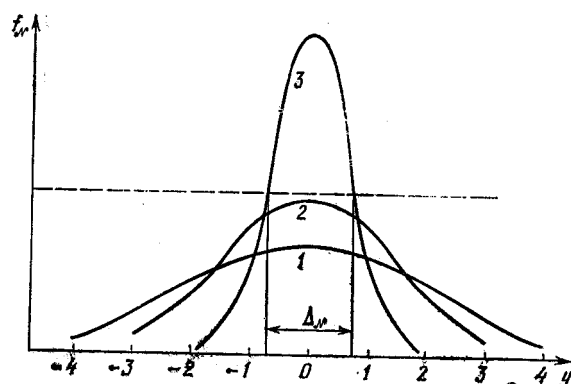


Рис. 2.4. Функции $f_N(y)$. Число узлов N возрастает с номером кривой. Штриховая линия показывает половину высоты кривой 3, Δ_N — полуширина этой кривой.

щади под всеми тремя колоколообразными кривыми должны быть одинаковыми. С увеличением N увеличивается максимальная высота кривых и уменьшается ширина. Ширину колоколообразующей кривой можно определить как расстояние между точками пересечения этой кривой с горизонтальной прямой, проведённой на расстоянии от оси абсцисс, равном половине максимальной высоты кривой (рис. 2.4). Обозначим эту ширину через Δ_N — обычно её называют полушириной.

Значения порога протекания, лежащие вне полуширины кривой, имеют вероятность по крайней мере в два раза меньшую, чем наиболее вероятное значение порога. Таким образом, полуширина характеризует типичный разброс порогов протекания, включающий отклонения, вероятность которых в два раза меньше, чем вероятность в максимуме кривой $f_N(y)$.

Вспомним, что по существу такую же информацию несёт и среднее квадратичное уклонение (гл. 1). Оно не определяет уклоне-

ние, вероятность которого ровно в два раза меньше максимальной, но тоже определяет типичный разброс значений порогов протекания.

Для любой колоколообразной кривой величины Δ_N и δ_N пропорциональны друг другу, но коэффициент пропорциональности зависит от вида кривой. Расчёты на ЭВМ показали, что функция распределения порогов протекания является гауссовой функцией (названной по имени великого математика К. Гаусса). Эта функция имеет вид

$$f_N(y) = \frac{1}{\delta_N \sqrt{2\pi}} \exp\left(-\frac{y^2}{2\delta_N^2}\right), \quad (2.6)$$

где $\exp a = e^a$; $e \approx 2,72$ — основание натурального логарифма. Она и изображена на рис. 2.4 при различных значениях δ_N . Поскольку функция симметрична относительно точки $y = 0$, в которой она достигает максимума, полуширину Δ_N можно найти из соотношения (см. рис. 2.4)

$$f_N\left(\frac{\Delta_N}{2}\right) = \frac{1}{2}f_N(0).$$

Используя формулу (2.6), получаем

$$\Delta_N = 2(2 \ln 2)^{\frac{1}{2}}\delta_N.$$

Согласно формуле (1.8), величина δ_N обращается в нуль по степенному закону, если $N \rightarrow \infty$. Это значит, что при безграничном увеличении числа узлов полуширина функции распределения порогов протекания стремится к нулю, т. е. сама функция превращается в резкий пик. Все значения порога протекания, кроме одного,

имеют нулевую вероятность. В связи с этим мы ещё раз повторим самое главное утверждение предыдущей главы: при $N \rightarrow \infty$ порог протекания превращается из величины случайной в величину достоверную.

Упражнение

6. (Для владеющих интегральным исчислением!) Подставьте функцию $f_N(y)$, определяемую формулой (2.6), в формулы (2.3) и (2.4) и докажите, что среднее значение \bar{a} , вычисленное с помощью этой функции, равно нулю, а дисперсия δ^2 равна δ_N^2 .

Глава 3.

Бесконечный кластер

В этой главе речь по-прежнему будет идти о задаче узлов теории протекания, однако на этот раз она будет сформулирована на другом языке — на языке кластеров. Кроме того, будет обсуждаться другой объект — вместо сетки с заблокированными узлами мы будем говорить о ферромагнетике с примесными атомами. Это значительно более сложный объект, о нём следует хотя бы вкратце рассказать.

3.1. Постоянный магнит

Наверное, почти все знают, отчего железо, никель, кобальт и некоторые другие материалы могут быть постоянными магнитами. Это явление объясняется тем, что атомы, из которых составлены такие вещества, сами являются элементарными магнитами. Они обладают *магнитными моментами*.

Хорошо известной системой, обладающей магнитным моментом, является стрелка компаса. Магнитный момент — вектор. Стрелка компаса имеет южный и северный полюсы, а её магнитный момент

направлен от южного полюса к северному. Внешнее магнитное поле заставляет стрелку компаса поворачиваться, так чтобы она была ориентирована вдоль магнитных силовых линий. Точно так же поворачивается во внешнем поле всякий магнитный момент. Стрелка компаса создает внешнее магнитное поле. Точно такое же магнитное поле создает любой магнитный момент.

Еще в начале XIX в. было выяснено, что источником магнетизма является движение электрических зарядов, т. е. электрический ток. Магнитный момент создается током. Для плоской петли с током, изображённой на рис. 3.1, магнитный момент μ определяется формулой $\mu = \frac{1}{c}IS$, где I — сила тока, S — площадь петли, c — скорость света (в системе единиц СГС). Вектор направлен перпендикулярно плоскости петли, причём так, что ток течет против часовой стрелки, если смотреть с той стороны, куда показывает стрелка вектора.

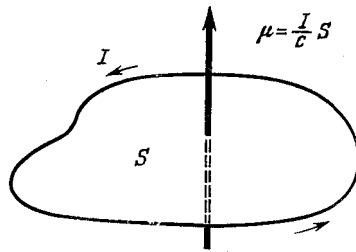


Рис. 3.1. Петля с током и её магнитный момент.

Если система состоит из нескольких петель с током, то можно, пользуясь правилом параллелограмма, сложить магнитные моменты петель и найти суммарный магнитный момент системы.

Как возникает магнитный момент у атомов? Как известно, атом состоит из тяжелого ядра и электронной оболочки. Магнетизм

твёрдых тел связан именно с магнитным моментом оболочки (ядро атома тоже может обладать магнитным моментом, но примерно в тысячу раз меньшим, чем момент оболочки).

Момент оболочки связан, во-первых, с движением электронов вокруг тяжелого ядра. Этому движению можно сопоставить некоторую силу тока I и эффективную площадь S . Кроме того, квантовая механика приписывает каждому электрону дополнительный магнитный момент, который называется *спиновым*. Он никак не связан с характером движения электрона, а является его внутренним свойством. Однако спиновый момент создаёт магнитное поле так же, как и обычный момент. Чаще всего суммарный магнитный момент электронных оболочек атомов, составляющих твердое тело, равен нулю. Однако у некоторых веществ, таких как железо, никель, кобальт и т. д., электронные оболочки обладают магнитным моментом.

В твёрдом теле магнитные моменты соседних атомов взаимодействуют друг с другом. В принципе, это взаимодействие похоже на взаимодействие стрелок двух компасов, поставленных рядом. Каждая стрелка создает магнитное поле, действующее на другую стрелку. Однако дело значительно усложняется тем, что взаимодействие происходит не в пустоте. Внешние электронные оболочки атомов существенно влияют на характер взаимодействия, изменяя даже направления действующих сил.

Эксперимент показывает, что в некоторых веществах взаимодействие между магнитными моментами таково, что силы, действующие между ними, заставляют их ориентироваться в одном направлении. Эти вещества называются ферромагнитными

(рис. 3.2).

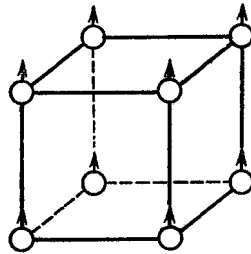


Рис. 3.2. Фрагмент кристаллической решетки ферромагнетика. Стрелки показывают направления магнитных моментов.

Если магнитные моменты всех атомов ориентированы в одном направлении, то суммарный магнитный момент M равен арифметической сумме отдельных моментов: $M = \mu N$, где N — число атомов в твёрдом теле, а μ — магнитный момент на один атом.

При увеличении размеров тела магнитный момент растёт пропорционально его объёму (число атомов N пропорционально объёму). Удельной характеристикой магнитных свойств, т. е. величиной, не зависящей от размеров тел, а зависящей лишь от свойств атомов, которые составляют это тело, является так называемая *спонтанная намагниченность* K . Она определяется как магнитный момент единицы объёма, т. е. равна полному моменту, поделённому на объём тела V :

$$K = \frac{M}{V} = \mu \frac{N}{V} = \frac{\mu}{v_0},$$

где $v_0 = \frac{V}{N}$ — объём, приходящийся на один атом.

Слово *спонтанная* означает самопроизвольная. Оно подчёркивает, что намагниченность K возникает не под действием внешнего магнитного поля, а за счёт внутренних сил.

Постоянный магнит — это и есть такое тело, внутри которого отлична от нуля спонтанная намагниченность. Благодаря намагниченности в среде, окружающей магнит (или в пустоте), создаётся магнитное поле.

Спонтанная намагниченность в системе единиц СГС измеряется в гауссах. Например, в железе при очень низких температурах $K = 1740$ Гс. Отсюда можно найти магнитный момент μ , приходящийся на один атом. Он равен примерно 2,2 от спинового магнитного момента электрона. То, что момент μ оказывается порядка спинового момента, подтверждает правильность наших представлений о природе спонтанной намагниченности.

Тепловое движение разрушает магнитный порядок, и потому существует критическая температура, называемая температурой Кюри, выше которой спонтанная намагниченность равна нулю. Для железа, например, температура Кюри составляет 770°C . При более высоких температурах железо не может быть постоянным магнитом.

3.2. Ферромагнетик с примесями

Рассмотрим теперь вещество, представляющее собой твёрдый раствор (смесь) магнитных и немагнитных (не имеющих магнитного момента) атомов. Это кристалл, в узлах которого располагаются магнитные или немагнитные атомы, причём их расположение оказывается не упорядоченным, а совершенно случайным. Допустим, что взаимодействие между магнитными моментами атомов убывает с расстоянием так быстро, что учитывать нужно лишь взаимо-

действие ближайших соседей. Это значит, что если два магнитных атома стоят рядом, то их моменты обязательно параллельны, но если между ними оказался хотя бы один немагнитный атом, то моменты могут быть направлены произвольно: они уже «ничего не знают» друг о друге.

Вопрос, который будет теперь поставлен, состоит в том, существует ли спонтанная намагниченность, если есть немагнитные атомы, и сколько немагнитных атомов требуется, чтобы разрушить спонтанную намагниченность. Ниже показано, что ответ на этот вопрос сводится к решению задачи узлов, которая была сформулирована в гл. 1.

Введём некоторые определения. Будем называть два магнитных атома связанными друг с другом, либо если они стоят рядом, либо если они соединены друг с другом цепочкой из стоящих рядом магнитных атомов (рис. 3.3). Выражение «стоят рядом» означает, что атомы являются ближайшими соседями. В квадратной решётке, изображенной на рис. 3.3, ближайшими соседями являются соседи по горизонтали и по вертикали, но не по диагонали. Совокупность связанных атомов принято называть кластером (в переводе с английского языка слово кластер (cluster) означает гроздь или кисть). Смысл этого определения состоит в следующем. Благодаря магнитному взаимодействию связанные атомы ориентируют свои магнитные моменты в одну сторону. Таким образом, каждый кластер обладает результирующим магнитным моментом, пропорциональным числу атомов, из которых он состоит. Кроме того, мы договорились, что магнитные атомы, не являющиеся ближайшими соседями, не взаимодействуют друг с другом. Поэтому

не взаимодействуют друг с другом атомы, принадлежащие разным кластерам. Вследствие этого взаимная ориентация магнитных моментов, принадлежащих разным кластерам, оказывается произвольной (рис. 3.3).

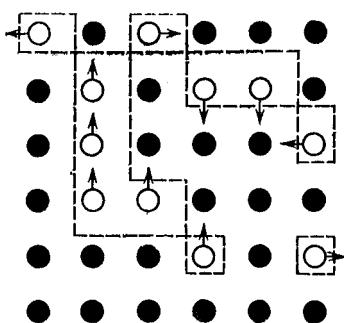


Рис. 3.3. Кусок плоской решётки с магнитными (светлые) и немагнитными (тёмные) атомами. Магнитные атомы образуют один кластер из четырёх атомов, один кластер из двух атомов и пять кластеров из одного атома. Границы кластеров указаны штриховыми линиями. Моменты разных кластеров могут быть направлены в разные стороны.

Обозначим через x долю магнитных атомов, т. е. отношение числа магнитных атомов к полному числу узлов решётки. По определению величина x меняется в интервале от 0 до 1.

Предположим сначала, что магнитных атомов очень мало ($x \ll 1$). Естественно, что тогда они в основном располагаются по одиночке (как изюминки в булке). Кластер из двух магнитных атомов представляет собой редкое событие, из трёх — ещё более редкое и т. д. Это утверждение крайне важно для дальнейшего, и чуть ниже оно будет доказано математически. А пока пусть те, кому это утверждение не кажется очевидным, примут его «на веру».

Таким образом, при $x \ll 1$ число кластеров примерно равно числу магнитных атомов N и, следовательно, оно растёт при уве-

личении полного числа узлов пропорционально N . Однако магнитные моменты этих кластеров «ничего не знают» друг о друге, и, следовательно, ориентированы относительно друг друга хаотически (рис. 3.3). Чтобы получить полный магнитный момент системы M , нужно сложить моменты отдельных атомов по правилу параллелограмма. Благодаря случайному направлению эти моменты компенсируют друг друга, так что спонтанная намагниченность оказывается равной нулю. *Итак, мы получили, что при малых концентрациях магнитных атомов спонтанная намагниченность отсутствует.*

3.3. Появление бесконечного кластера

Теперь рассмотрим случай, когда почти все атомы — магнитные. Очевидно, что небольшая примесь немагнитных атомов не уничтожает спонтанной намагниченности, а только уменьшает её. Обсудим этот вопрос на языке кластеров. При $x = 1$ все N атомов принадлежат единому кластеру. Если x немного отличается от единицы, то часть атомов выпадает из этого кластера. Это происходит, во-первых, потому что некоторые атомы замещаются немагнитными (атомы A на рис. 3.4), а во-вторых, потому что некоторые магнитные атомы образуют изолированные кластеры (атом B на рис. 3.4) со своим направлением магнитного момента. Тем не менее при значениях x , близких к единице, сохраняется единый кластер, пронизывающий всю решётку, сколь бы велика она ни

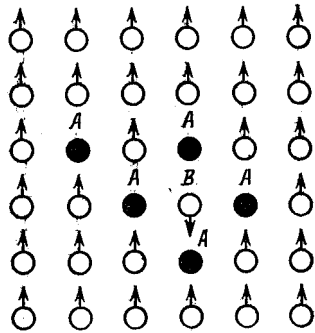


Рис. 3.4. Фрагмент плоской решётки с магнитными (светлые) и немагнитными (тёмные) атомами в случае большой концентрации магнитных атомов. Все магнитные атомы за исключением атома B принадлежат одному кластеру и имеют одинаковое направление магнитных моментов.

была. Этот кластер принято называть *бесконечным кластером*.

Разумеется, строгий смысл это понятие приобретает лишь для бесконечной системы. Возьмём большую серию образцов с заданными значениями числа магнитных атомов и полного числа атомов и выберем в каждом из них кластер с максимальным числом магнитных атомов. Усредним число магнитных атомов, принадлежащее максимальному кластеру, по всем образцам серии и обозначим результат усреднения через N_{max} . Таким образом, N_{max} — среднее число атомов в самом большом кластере. Величина N_{max} зависит от N и от x . Существование бесконечного кластера проявляется в том, что при заданном значении x отношение $\frac{N_{max}}{N}$ при безграничном увеличении N стремится к отличному от нуля пределу

$$\lim_{N \rightarrow \infty} \frac{N_{max}}{N} = P(x).$$

Доля атомов $P(x)$, принадлежащих самому большому кластеру, не зависит от числа атомов N , если оно достаточно велико, но зави-

сит от x . Сама же величина N_{max} стремится к бесконечности при бесконечном увеличении N . Поэтому и говорят о существовании бесконечного кластера.

В системе может существовать только один бесконечный кластер. Допустим, что при заданных значениях N и x определено не только среднее число атомов в самом большом кластере, но и среднее число атомов в следующем по величине кластере. Обозначим эту последнюю величину через N'_{max} . По определению $N'_{max} < N_{max}$. Утверждение о том, что в системе может быть только один бесконечный кластер, означает, что

$$\lim_{N \rightarrow \infty} \frac{N'_{max}}{N} = 0$$

при всех значениях x . Это означает, что два кластера, пронизывающих всю систему, неизбежно должны где-то связаться друг с другом и превратиться в один кластер².

Итак, мы получили, что при достаточной концентрации магнитных атомов x определенная доля этих атомов принадлежит одному кластеру и имеет одинаковое направление магнитных моментов. Это означает, что существует спонтанная намагниченность

$$K = \frac{\mu}{v_0} = P(x).$$

Вспомним теперь, что при малой концентрации магнитных атомов x есть только небольшие кластеры. В этом случае увеличение числа узлов N приводит лишь к росту числа малых кластеров, но

²Строго говоря, утверждение о существовании лишь одного бесконечного кластера не является доказанным. Существуют доводы в его пользу той или иной степени достоверности, но правильнее сказать, что специалисты просто принимают «на веру» это утверждение.

не к увеличению числа частиц в каждом из них. Тогда

$$\lim_{N \rightarrow \infty} \frac{N_{max}}{N} = 0, \quad \text{т. е.} \quad P(x) = 0.$$

Таким образом, мы приходим к выводу о существовании кри-

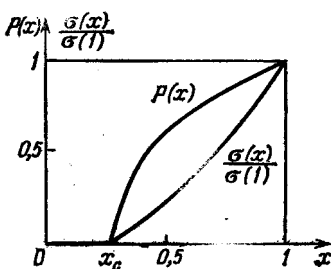


Рис. 3.5. Графики функций $P(x)$ и $\frac{\sigma(x)}{\sigma(1)}$. Обе функции обращаются в нуль в одной точке, но по причинам, о которых пойдет речь в третьей части книги, вид этих функций вблизи критического значения x_c резко различается.

тической концентрации x_c , при которой возникает бесконечный кластер, причём x_c удовлетворяет неравенствам $0 < x_c < 1$. При этой же концентрации x_c появляется спонтанная намагниченность и становится отличной от нуля функция $P(x)$ (рис. 3.5). Итак, если доля немагнитных атомов оказывается большей, чем $1 - x_c$ (доля магнитных атомов меньше, чем x_c), то вещество не может быть постоянным магнитом.

Упражнение

1. Найдите вид функции $P(x)$ при значениях x , близких к единице.

3.4. Снова задача узлов

Теперь нам осталось только сказать, что с точки зрения критической концентрации x_c задача об электропроводности сетки и задача о примесном ферромагнетике — это одна и та же задача.

Задачу об электропроводности также легко сформулировать на языке кластеров. Нужно только во всех определениях заменить понятие «немагнитный атом» на «блокированный узел». На рис. 3.3 изображена некоторая конфигурация магнитных (светлые кружки) и немагнитных (чёрные кружки) атомов. Произведем для этой конфигурации указанную выше замену и перейдем от ферромагнетика с примесями к экранной сетке с вырезанными узлами. Для этого на рис. 3.3 нужно убрать стрелки, показывающие направления магнитных моментов, и изобразить проволочки, связывающие узлы друг с другом (рис. 3.6).

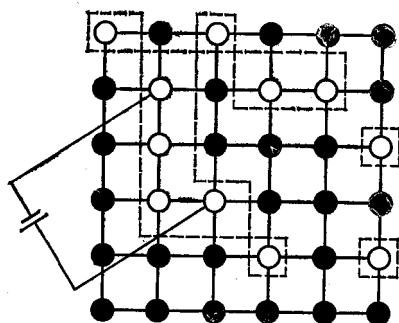


Рис. 3.6. Та же конфигурация, что на рис. 3.3, но магнитные атомы заменены неблокированными узлами.

На рис. 3.6 хорошо видно основное свойство кластеров в применении к задаче о сетке. Если приложить к любой паре узлов одного кластера разность потенциалов, возникает замкнутая цепь, по которой течёт электрический ток. (Разумеется, это свойство

имеет смысл только для кластеров, содержащих не менее двух узлов.) При приложении разности потенциалов к любой паре узлов, принадлежащих разным кластерам, цепь не замыкается и ток не течёт. Если $x < x_c$, то в системе есть только кластеры из конечного числа узлов, и поэтому при увеличении размеров системы ток через боковые электроды рано или поздно обязательно прервется. Если же $x > x_c$, то в очень большой системе на боковых гранях обязательно окажутся узлы, принадлежащие бесконечному кластеру. Этот бесконечный кластер и обеспечит отличную от нуля и не зависящую от размеров системы удельную электропроводность $\sigma(x)$.

Вернёмся к рис. 3.5, где изображены функции $P(x)$ (доля узлов, принадлежащих бесконечному кластеру) и $\frac{\sigma(x)}{\sigma(1)}$ ($\sigma(1)$ — электропроводность при $x = 1$, т. е. без заблокированных узлов). Обе функции обращаются в нуль в одной и той же точке, которая сначала была названа порогом протекания, а потом точкой, где возникает бесконечный кластер.

Таким образом, речь всё время шла о задаче теории протекания, которая называется задачей узлов. Если бы нас интересовало значение x_c для «плоского ферромагнетика», то можно было бы сказать, воспользовавшись результатом опыта с сеткой, что оно равно 0,59. Однако реальные ферромагнетики кристаллизуются не в плоские, а в объёмные (трёхмерные) решётки. Примером трёхмерной решётки является простая кубическая решётка, одна ячейка которой была изображена на рис. 3.2.

Задача об электропроводности экранной сетки легко обобщается на трёхмерный случай. Представим себе куб, содержащий

много ячеек и спаянный из проволоки, как показано на рис. 3.7. К двум противоположным граням этого куба можно припаять металлические пластинки, составить так же, как на рис. 1.1, электрическую цепь и изучить электропроводность в зависимости от числа заблокированных узлов. При блокировании каждого узла разрывается контакт между шестью проволоками, входящими в этот узел. Как и в двумерном случае, существует критическая концентрация x_c неблокированных узлов, ниже которой электропроводность равна нулю.

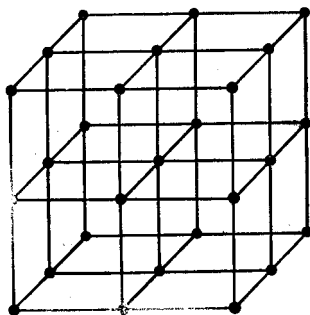


Рис. 3.7. Простая кубическая решётка.

Задача о примесном ферромагнетике и связанное с ней представление о бесконечном кластере относились в равной мере к плоским и к трёхмерным решёткам. Критическая концентрация магнитных атомов x_c , при которой возникает бесконечный кластер, одновременно является порогом протекания с грани на грань в кубе достаточно большого размера. Нужно иметь в виду, что сама величина x_c сильно зависит от типа решётки. Если для квадратной решётки она равнялась 0,59, что для простой кубической решётки $x_c = 0,31$. (Подробнее см. гл. 6.)

3.5. Кластеры при низкой концентрации магнитных атомов**

Выводы, сделанные в предыдущих разделах, были в значительной мере основаны на утверждении о том, что при малой концентрации магнитных атомов x эти атомы в основном расположены по одиночке, кластеры из двух атомов редки, из трёх — еще более редки и т. д. Докажем это утверждение.

Введём функцию $P_M(x)$ — вероятность того, что выбранный наугад атом принадлежит кластеру не менее чем из M атомов. Это значит, что выбранный наугад атом: а) магнитный, б) связан с не менее, чем $M - 1$, другими магнитными атомами. Вычислим функцию $P_M(x)$ при $M = 1$ и $M = 2$.

Функция $P_1(x)$ равна вероятности того, что выбранный наугад атом окажется магнитным. Эта вероятность равна x (см. упр. 1 к гл. 2, где нужно заменить слово «неблокированный» словом «магнитный» а слово «блокированный» словом «немагнитный»). Итак,

$$P_1(x) = x. \quad (3.1)$$

Функция $P_2(x)$ равна вероятности того, что выбранный наугад атом является магнитным и при этом среди его ближайших соседей есть по крайней мере ещё один магнитный атом. Оба указанных события являются, очевидно, независимыми, и потому исходную вероятность можно представить в виде произведения вероятностей этих событий. Так как первая из них (вероятность, что атом магнитный) равна x , то

$$P_2(x) = xW(x), \quad (3.2)$$

где $W(x)$ — вероятность того, что среди ближайших соседей некоторого атома есть по крайней мере один магнитный атом. Функция $W(x)$ зависит от того, какая решётка рассматривается. Ограничимся квадратной решёткой, в которой каждый атом имеет четырёх ближайших соседей (рис. 3.8). Нужно найти вероятность того, что хотя бы один из атомов 1, 2, 3, 4 является магнитным.

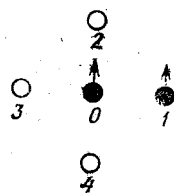


Рис. 3.8.

Эту задачу проще всего решить следующим образом. Событие, состоящее в том, что все четыре атома магнитные, и событие, состоящее в том, что по крайней мере один из четырёх атомов магнитный, образуют полную систему событий. Сумма вероятностей двух этих событий равна единице. Вероятность первого события обозначим через W_0 , а вероятность второго события и есть искомая величина W . Итак, $W + W_0 = 1$. Вероятность того, что атом 1 — немагнитный, равна $1 - x$. Вероятность того, что атом 2 или атом 3, или атом 4 — немагнитные, тоже равна $1 - x$. События, состоящие в том, что разные атомы оказались немагнитными, являются независимыми. Поэтому вероятность того, что все четыре атома — немагнитные, равна произведению вероятностей: $W_0 = (1 - x)^4$. Отсюда следует, что $W = 1 - W_0 = 1 - (1 - x)^4$.

Согласно формуле (3.2)

$$P_2(x) = x [1 - (1 - x)^4]. \quad (3.3)$$

Если $x \ll 1$, выражение для $P_2(x)$ можно упростить, отбросив члены с высокими степенями x . Пользуясь формулой бинома, получаем

$$1 - (1 - x)^4 = 4x - 6x^2 + 4x^3 - 4x^4.$$

Теперь заметим, что если $x \ll 1$, отношение каждого последующего члена к предыдущему мало:

$$\frac{6x^2}{4x} = \frac{3}{2}x \ll 1, \quad \frac{4x^3}{6x^2} = \frac{2}{3}x \ll 1 \quad \text{и} \quad \frac{4x^4}{4x^3} = x \ll 1.$$

Поэтому при $x \ll 1$ с хорошей точностью можно написать

$$1 - (1 - x)^4 \approx 4x,$$

откуда следует

$$P_2(x) \approx 4x^2. \quad (3.4)$$

Сравнивая формулы (3.1) и (3.4), видим, что при $x \ll 1$ отношение

$$\frac{P_2(x)}{P_1(x)} \approx 4x \ll 1, \quad (3.5)$$

т. е. вероятность того, что выбранный наугад атом принадлежит кластеру из двух и более атомов, гораздо меньше, чем вероятность того, что он образует кластер из одного атома.

Существует более простой вывод формулы (3.4) для $P_2(x)$, в котором сразу же учитывается условие $x \ll 1$. С помощью этого вывода можно сравнительно просто вычислить функции $P_M(x)$ при

$M > 2$ (см. упр. 3). Вывод состоит в следующем: кластер более чем из двух атомов, включающий атом 0, обязательно включает либо атом 1, либо 2, либо 3, либо 4. Вероятность того, что атомы 0 и 1 принадлежат одному кластеру, равна вероятности того, что оба они магнитные, и равна произведению вероятностей того, что каждый из этих атомов магнитный, т. е. $x \cdot x = x^2$. То же самое можно сказать про вероятность того, что кластер образован атомами 02, 03 или 04. Все эти вероятности равны x^2 . Вероятность того, что осуществится хотя бы одно из этих событий, равна сумме вероятностей, т. е. $4x^2$, что и приводит к формуле (3.4).

Этот вывод справедлив лишь при $x \ll 1$. Только при этом условии можно пользоваться правилом сложения вероятностей. Действительно, правило сложения справедливо для несовместимых событий. Но событие, состоящее в том, что атомы 0 и 1 — магнитные, совместимо с тем, что атомы 0 и 2 — магнитные. Совмещение событий означает, что все три атома 0, 1, 2 — магнитные и, следовательно, образуют кластер из трёх атомов. Вероятность совмещенного события равна произведению вероятностей того, что все три атома — магнитные, т. е. равна $x \cdot x \cdot x = x^3$.

При $x \ll 1$ эта вероятность много меньше, чем вычисленная вероятность образования кластера из двух атомов. Поэтому вероятностью совмещения событий можно пренебречь и рассматривать события как несовместимые. Это и оправдывает сделанный выше вывод при условии $x \ll 1$.

Фактически это означает, что если $x \ll 1$, то при расчёте $P_2(x)$ можно пренебречь вероятностью образования кластера из трёх атомов.

Таким образом, при $x \ll 1$ функция $P_2(x)$ фактически совпадает с вероятностью того, что выбранный наугад узел принадлежит кластеру из двух (и не более!) атомов. Соответственно, функция $P_3(x)$ описывает кластер из трёх атомов. Она пропорциональна x^3 и мала по сравнению с $P_2(x)$. Общий результат состоит в том, что функция $P_M(x)$ содержит степени x , не меньшие, чем x^M и при $x \ll 1$ будет $P_M(x) \ll P_{M-1}(x)$.

Итак, если при $x \ll 1$ выбранный наугад узел оказался магнитным, то он с подавляющей вероятностью образует кластер из одного узла. Вероятность того, что он принадлежит кластеру из M узлов, резко убывает с ростом M .

Упражнения

2. Найдите $P_2(x)$ для простой кубической решётки, изображённой на рис. 3.7. Для любой решётки, в которой каждый атом имеет z ближайших соседей.
3. Найдите $P_3(x)$ для квадратной решётки, пользуясь условием $x \ll 1$.
4. Найдите $P_3(x)$ для квадратной решётки, не пользуясь условием $x \ll 1$.

Глава 4.

Решение задачи узлов методом Монте-Карло на ЭВМ

Метод Монте-Карло — наиболее распространённый способ решения задач теории протекания. Цель этой главы — дать общее представление об этом методе, объяснить подробно, как работает главный элемент метода — генератор случайных чисел и в заключение привести конкретную программу для ЭВМ, позволяющую находить порог протекания задачи узлов. Наверное, самый первый вопрос, который возникает, это:

4.1. Почему Монте-Карло?

«— А что такое зéго? Вот этот крупер курчавый, главный-то, крикнул зéго? И почему он все загрёб, что ни было на столе? Эдакую кучу, всё себе взял? Это что такое?»

— А зéго, бабушка, выгода банка. Если шарик упадёт на зéго, то всё, что ни поставлено на столе, принадлежит банку без расчёта

...

— Вот-те на! А я ничего не получаю?

— Нет, бабушка, если вы перед этим ставили на зéго, вам платят в тридцать пять раз больше.

— Как, в тридцать пять раз, и часто выходит? Что ж они, дураки, не ставят?

— Тридцать шесть шансов против, бабушка.

— Вот вздор! Потапыч! Потапыч! Постой, и со мной есть деньги, — вот!

Она вынула из кармана туго набитый кошелёк и взяла из него фридрихсдор.

— На, поставь сейчас на зéго.

— Бабушка, зéго только что вышел, — сказал я, — стало быть, теперь долго не выйдет. Вы много проставите; подождите хоть немного.

— Ну врёшь, ставь!

— Извольте, но он до вечера, может быть, не выйдет, вы до тысячи проставите, это случалось.

— Ну, вздор, вздор! Волка бояться — в лес не ходить. Что? проиграл? ставь ещё!»³

Этот фрагмент из романа Ф. М. Достоевского «Игрок» описывает самую азартную игру минувшего века — рулетку. Город Монте-Карло, находящийся в княжестве Монако, снискал себе славу всемирной столицы рулетки. В честь этого города и получил свое название один из наиболее мощных современных математических

³ Следует заметить, что с точки зрения теории вероятности всеопытная экзальтированная бабушка проявляет больше здравого смысла, чем консультирующий ее Игрок. Вероятность выхода зéго несколько не уменьшается от того, что он вышел перед этим. Ждать, как советует Игрок, нет никакого смысла. Это часто встречающееся недоразумение, видимо, основано на неправильном понимании того факта, что вероятность два раза подряд получить зéго мала. Но отсюда совсем не следует, что если один раз вышел зéго, то вероятность его выхода в следующий раз меньше, чем в первый. Вероятность, конечно, остается точно такой же.

методов.

Что же общего между этим методом и рулеткой? — То, что основным элементом метода Монте-Карло является тот самый вращающийся шарик, который в многочисленных игорных залах Монте-Карло вершит судьбы людей, повергая одних и возвышая других. Правда, математики сильно усовершенствовали его. Это уже совсем не шарик, а стандартная программа ЭВМ, которая называется «генератор случайных чисел». Но суть дела от этого не меняется. Шарик рулетки с математической точки зрения тоже не что иное, как генератор случайных чисел.

4.2. Что такое метод Монте-Карло?

Методом Монте-Карло, как правило, называют любой математический метод, существенно использующий генератор случайных чисел.

Обычно современная ЭВМ имеет стандартную программу, генерирующую случайные числа, равномерно распределённые в интервале от нуля до единицы, т. е. «разыгрывающую» значения непрерывной случайной величины, принимающей с равной вероятностью все значения в интервале $(0, 1)$.

Каждое обращение к этой программе даёт одно такое число с определённым числом знаков после запятой, зависящим от класса ЭВМ.

Простейшее применение метода Монте-Карло состоит, например, в вычислении интегралов. Допустим что надо сосчитать объём, ограниченный замкнутой поверхностью, имеющей сложную

форму. Выберем куб, заведомо включающий в себя всю поверхность (рис. 4.1). С помощью генератора случайных чисел получим набор точек, равномерно распределённых внутри куба. Это делается следующим образом. Допустим, что длина ребра куба равна L и все три координаты заключённых внутри него точек меняются от нуля до L (рис. 4.1). Трижды обратившись к генератору случайных чисел, получим три числа, y_1, y_2, y_3 , лежащие в интервале $(0, 1)$. Построим из них координаты первой точки, лежащей внутри куба, по формулам $X_1 = Ly_1, Y_1 = Ly_2, Z_1 = Ly_3$. Проведя эту процедуру Q раз, получим Q точек, в среднем равномерно заполняющих куб. Пусть Q_1 — число точек, оказавшихся внутри поверхности. Так как точки распределяются равномерно, то число Q_1 характеризует объём, ограниченный поверхностью. Именно, если число Q достаточно велико, то искомый объём равен $L^3 \frac{Q_1}{Q}$.

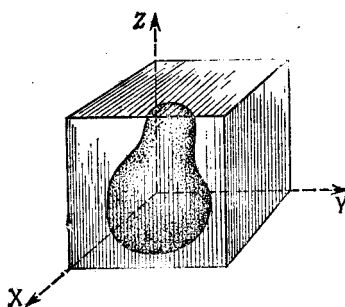


Рис. 4.1. Определение объёма груши методом Монте-Карло

Существует теория, позволяющая сказать, какое число точек Q требуется для получения результата с заданной точностью. Можно подойти к этому вопросу эмпирически и просто несколько раз повторить эксперимент, используя другие наборы случайных чисел, а затем сравнить ответы. Если в пределах заданной точности

ничего не меняется — значит, всё в порядке, ответ правильный. В случае пространств большого числа измерений (многократных интегралов) метод Монте-Карло имеет очень существенные преимущества над обычными методами интегрирования.

В целом ряде случаев метод Монте-Карло является единственно возможным. Представьте себе, что изучается поведение системы, состоящей из огромного числа частиц, например поведение газа. В принципе такая задача должна решаться методами статистической физики, однако, если взаимодействие между частицами сильное (так бывает при большой плотности и низкой температуре), эти методы оказываются неэффективными. Тогда свойства газа изучаются с помощью моделирования на ЭВМ. Число частиц газа, участвующих в моделировании, определяется объёмом памяти ЭВМ. В памяти должна храниться информация о координатах всех частиц. Моделирование состоит в том, что случайным образом выбирается одна из частиц, которая затем перемещается на случайное расстояние. (Это значит, что в памяти ЭВМ меняются координаты этой частицы.) Затем случайно выбирается другая частица и т. д. Потенциальная энергия взаимодействия частиц газа друг с другом зависит от их взаимного расположения. Она вычисляется с самого начала, а затем пересчитывается после каждого перемещения. Вероятности перемещений частиц на то или иное расстояние выбираются в соответствии с потенциальной энергией таким образом, чтобы модельная система в среднем «жила» так же, как реальная.

В результате в памяти ЭВМ оказываются как бы «мгновенные фотографии» газа, сделанные в последовательные моменты вре-

мени. Фотографии включают координаты всех частиц газа, и по ним можно вычислить средние термодинамические характеристики, такие как давление, теплоемкость и т. д.

Сама процедура моделирования очень напоминает игру, которая ведётся по строго установленным правилам, исключающим обращения к рулетке, т. е. к генератору случайных чисел. Малейшие отклонения от правил или «жульническая» рулетка приводят к тому, что одни конфигурации атомов газа появляются чаще других. Это сказывается на результате усреднения и делает ответ неправильным.

Генератор случайных чисел используется не только в методе Монте-Карло, но и в так называемых аналоговых экспериментах, примером которых является опыт с экранной сеткой, описанный в первой главе. Как уже говорилось, случайная последовательность блокируемых узлов, нужная для этого опыта, составлялась на ЭВМ. Чтобы выбрать очередной узел, нужно обратиться к программе и получить случайное число y . Его следует умножить на полное число узлов N и прибавить к произведению единицу. Затем от $Ny + 1$ нужно взять целую часть. При этом возникнет целое число, лежащее в требуемом интервале от 1 до N . Правда, такие числа могут повторяться. Но это не страшно. Если окажется, что узел с полученным номером был блокирован ранее, нужно запросить у ЭВМ новое случайное число и превратить его в номер узла.

Несколько позднее будет описана программа, с помощью которой вычисляются пороги протекания методом Монте-Карло, а сейчас речь пойдёт о самом кардинальном элементе этого метода — генераторе случайных чисел.

4.3. Как придумать случайное число?

Итак, вам нужны случайные числа, равномерно распределённые в интервале от нуля до единицы. Этим задание ещё не сформулировано. Необходимо знать, сколько знаков после запятой требуется у каждого числа. Допустим, что нужно всего два знака. Тогда простейший рецепт состоит в следующем. Возьмите телефонную книгу, откройте её на любой странице и списывайте подряд две последние цифры каждого телефонного номера, поставив перед ними «0,». Вы получите неплохую двузначную таблицу случайных чисел. А что делать, если требуются десятизначные числа? Пожалуй, тогда вам не обойтись без ЭВМ.

Если немного подумать, то сама мысль о том, что ЭВМ может генерировать случайные числа, покажется странной. Ведь ЭВМ работает по предложенному ей алгоритму, т. е. совершает в точности те действия, которые требует от неё человек. Как же внести в эти действия элемент случайности?

На самом деле никакого элемента случайности в программе генератора случайных чисел нет. Принцип её работы состоит в следующем. При первом обращении к программе необходимо задать некоторое число y_0 . С помощью совершенно определенной последовательности действий это число преобразуется в новое число

$$y_1 = \Phi(y_0), \quad (4.1)$$

где Φ — выбранная определённым образом функция или последовательность операций, преобразующих y_0 в y_1 . Ею и определяется алгоритм генерации случайных чисел. В свою очередь число y_1 служит основой для получения следующего числа y_2 по тому же

самому рецепту:

$$y_2 = \Phi(y_1). \quad (4.2)$$

Разумеется, функция Φ устроена так, что все числа y_1, y_2, \dots, y_n удовлетворяют неравенствам $0 \leq y_n \leq 1$. Они и являются искомыми случайными числами.

Легко убедиться в том, что последовательность чисел, полученная таким образом, не может быть бесконечной. Действительно, ЭВМ оперирует лишь числами, содержащими определенное количество знаков (разрядов). Количество таких чисел ограничено. (Существует всего 10^2 двузначных чисел и 10^n n -значных чисел.) Поэтому рано или поздно очередное число y_n совпадает с тем, которое уже было раньше, например, с y_{n-L} . После этого всё начнёт повторяться: y_{n+1} совпадёт с y_{n-L+1} и т. д.

Таким образом, последовательность чисел, получаемая с помощью формул (4.1), (4.2), неизбежно оказывается периодической. Поэтому такие числа называют не истинно случайными, а псевдослучайными (т. е. как бы случайными или похожими на случайные).

Однако ими можно пользоваться как случайными, если количество чисел, которое требуется для решения данной задачи, меньше, чем период последовательности L .

В свою очередь, период L определяется количеством десятичных знаков, с которым оперирует ЭВМ (т. е. количеством ячеек памяти, предоставляемых каждому числу), а также качеством алгоритма (т. е. свойствами функции Φ , фигурирующей в формулах (4.1) и (4.2)).

Разработать хороший генератор случайных чисел — задача очень сложная. Генераторы, составленные наугад, как правило, оказываются плохими. Ниже рассмотрены некоторые конкретные генераторы.

4.4. Метод середины квадрата

Исторически это первый метод генерации псевдослучайных чисел на ЭВМ. Он был предложен в 1946 г. известным математиком фон Нейманом. Метод позволяет генерировать числа с любым числом знаков, соответствующим возможности ЭВМ. Метод очень прост. Допустим, что нужны четырёхзначные числа. Выбираем первое число X_0 произвольно. Например, $X_0 = 8219$. Возводим его в квадрат. Получится восьмизначное число 67551961. Извлекаем средние цифры: 5519. Следующим числом последовательности является $X_1=5519$. Возводим в квадрат 5519, получаем 30459361. Следующее случайное число $X_2=4593$. Если первые из средних цифр оказываются нулями, то получается число с меньшим количеством знаков. Например, $X_2^2=21095649$, $X_3=956$. Возводя его в квадрат, нужно получить восьмизначное число, дописав спереди нули $X_3^2=00913936$, так что $X_4=9139$ и т. д.

Случайные числа y_n , равномерно распределённые в интервале от нуля до единицы, получаются из чисел X_n по формуле $y_n = \frac{X_n}{1000}$, где $n = 0, 1, 2, 3, \dots$, так что $y_0=0,8219$; $y_1=0,5519$; $y_2=0,4593$ и т. д.

На первый взгляд метод кажется хорошим. Однако тщательное исследование показало, что это далеко не так. Главный недостаток метода состоит в том, что при некоторых начальных числах последовательность «зацикливается». Выяснилось, например, что в классе четырёхзначных чисел последовательности часто завершаются циклом 6100, 2100, 4100, 8100, 6100. Период цикла равен всего-навсего четырём, что, конечно, никуда не годится.

Существует число, которое сразу же воспроизводит самое себя. Это 3792 ($3792^2 = 14379264$). Воспроизводит себя также нуль, и очень часто последовательности, полученные методом середины квадрата, вырождаются в нуль. Поэтому метод середины квадрата представляет в наше время лишь исторический интерес.

Упражнение

1. Составьте последовательность четырёхзначных чисел, начав с 0085, 0067, 0032. Покажите, что все они монотонно убывающие (каждое следующее число меньше предыдущего) и довольно быстро вырождаются в нуль.
2. Теперь докажите, что это общий недостаток метода середины квадрата: если используются $2n$ -значные числа X_i и в последовательности появилось число b , у которого n старших цифр равны нулю, то с этого момента последовательность становится монотонно убывающей и в конце концов вырождается в нуль.

4.5. Линейный конгруэнтный метод

Этот метод получения случайных чисел в настоящее время считается наилучшим. Его суть в следующем. Выбирают четыре целых положительных числа:

1. Множитель k ;
2. Сдвиг c ;
3. Модуль m ;
4. Первое число последовательности X_0 .

Последовательность случайных чисел определяется формулой

$$X_{n+1} = (kX_n + c) \bmod m, \quad (4.3)$$

где индекс n пробегает значения $0, 1, 2, 3, \dots$. Символ $b \bmod m$ означает остаток от деления числа b на m . Например:

b	25	6	30	3	147
m	10	10	10	12	12
$b \bmod m$	5	6	0	3	3

Очевидно, что $b \bmod m < m$. Поэтому все числа последовательности X_n удовлетворяют неравенству $X_n < m$. Последовательность чисел y_n , равномерно распределённых в интервале от нуля до единицы, получается по формуле

$$y_n = \frac{X_n}{m}, \quad n = 0, 1, 2, \dots \quad (4.4)$$

Далеко не любой выбор четырёх исходных чисел ведёт к хорошим результатам. Заметим прежде всего, что последовательность

чисел X_n неизбежно должна быть периодической, причём период не может быть больше, чем m . Действительно, так как все X_n — целые числа, причём $X_n < m$, количество различных чисел не может превышать m . Поэтому, начиная по крайней мере с $n = m$, появится число, которое уже встречалось, и всё повторится заново.

Однако получить последовательность с максимально возможным периодом $L = m$ далеко не просто. Если выбирать исходные числа не задумавшись, то, как правило, будут получаться последовательности с маленьким периодом.

Упражнение

3. Напишите последовательность чисел X_n , следующую из формулы (4.3) с $k = 3$, $c = 0$, $X_0 = 5$, $m = 20$.
4. Напишите последовательность чисел X_n , следующую из формулы (4.3) с $k = 3$, $c = 1$, $X_0 = 5$, $m = 20$.
5. Напишите последовательность чисел X_n , следующую из формулы (4.3) с $k = 3$, $c = 2$, $X_0 = 5$, $m = 20$. Убедитесь в том, что во всех трёх случаях период последовательности существенно меньше, чем 20. Разберите другие примеры.

Доказана следующая теорема. Если последовательность определяется формулой (4.3) с $c \neq 0$, то её период равен m тогда и только тогда, когда выполнены следующие условия:

- а) c и m — взаимно простые числа (не имеют общих делителей, кроме единицы);

- б) $b = k - 1$ кратно p для любого простого числа p , являющегося делителем m ;
- в) b кратно 4, если m кратно 4. К сожалению, доказательство этой теоремы слишком сложно, чтобы его здесь приводить.
6. Убедитесь в том, что во всех примерах, приведённых в упражнениях 3–5, условия, диктуемые сформулированной выше теоремой, не выполнялись.
7. Убедитесь в том, что набор чисел $k = 11$, $c = 3$, $m = 5$ удовлетворяет условиям теоремы и при всех X_0 даёт период $L = 5$.

Итак, чтобы получить генератор с максимально возможным периодом L , нужно взять в качестве m самое большое число, с которым может оперировать данная ЭВМ, и выбрать остальные числа в соответствии с приведённой выше теоремой.

Однако период — не единственный показатель качества случайной последовательности. Рассмотрим, например, последовательность, соответствующую $k = c = 1$. Она имеет вид $0, 1, 2, 3, \dots, m - 1, 0, 1, 2, 3, \dots, m - 1, 0, \dots$. Её период равен m , но это никуда не годная случайная последовательность.

Разработана сложная система тестов, которая позволяет определить качества генератора случайных чисел. Поэтому рекомендуется использовать только проверенные генераторы.

При выборе генератора свойства ЭВМ важны не только для того, чтобы выбрать максимально возможный период. От выбора исходных чисел зависит и скорость генерации случайных чисел.

При этом оказывается, что для ЭВМ разных конструкций оптимальными являются разные генераторы.

Программы, работающие по методу Монте-Карло, часто включают в себя огромное количество обращений к генератору случайных чисел (десятки и сотни миллионов раз). Поэтому быстроедействие — одно из важнейших качеств генератора.

Для ЭВМ марки БЭСМ-6 рекомендуется генератор с $k = 5^{17}$, $c = 0$, $m = 2^{40}$ и нечётными значениями X_0 . Этот набор чисел не удовлетворяет требованиям приведённой выше теоремы ($c = 0$), и период такого генератора меньше, чем m . Однако для генераторов с $c = 0$ доказана другая теорема, согласно которой период рекомендованного генератора равен $2^{38} \approx 2,75 \times 10^{11}$.

4.6. Определение порога протекания методом Монте-Карло на ЭВМ. Распределение заблокированных и неблокированных узлов

Теперь будет подробно описана программа, по которой работает ЭВМ при определении порога протекания по методу Монте-Карло. Заметим, что эта программа далеко не единственная. Более того, каждая группа исследователей, занимающаяся этими вопросами, предпочитает использовать собственную программу, в той или иной мере отличающуюся от других. Это связано отчасти с особенностями различных ЭВМ, а отчасти просто с индивидуальным опытом программистов.

Речь идёт о задаче узлов, причём для простоты рассматривается только плоская квадратная решётка. Впрочем, как будет видно

из дальнейшего, обобщение метода на любую решётку любой размерности не составляет труда.

Допустим, что нужно изучить протекание в квадрате, одна сторона которого содержит L узлов, так что полное число узлов $N = L^2$. Расстояние между узлами принимаем равным единице, а узлы характеризуем координатами X и Y . Например, узел с координатами $X = 9, Y = 25$ — это узел, находящийся в девятом слева столбце и в двадцать пятом ряду снизу.

Чтобы изучить протекание, нужно задать, какой узел заблокирован, а какой нет, и иметь возможность изменять число заблокированных узлов, чтобы проходить через порог протекания. Для этого сначала каждому узлу приписывается некоторое число V . Так как узел характеризуется двумя координатами X и Y , то это равносильно тому, что вводится функция двух переменных $V(X, Y)$, аргументы которой X и Y пробегает не все значения, а могут быть лишь целыми числами в интервале от 1 до L . В программировании такую функцию называют двумерным массивом, а значения, которые она принимает, называют элементами этого массива. Например, элемент массива $V(31, 97)$ есть некоторое число, относящееся к узлу с координатами $X = 31, Y = 97$. Массив V имеет всего $L \times L = N$ элементов, и в памяти ЭВМ нужно зарезервировать место, необходимое для записи N чисел.

Программа начинается с выработки этого массива. Его элементами являются случайные числа, равномерно распределённые от нуля до единицы. Генератор случайных чисел даёт число y , и это число присваивается элементу массива $V(1, 1)$. Это значит, что оно записывается в соответствующей ячейке памяти ЭВМ и, начи-

ная с этого момента, ЭВМ «помнит», что $V(1, 1) = y$. Следующее число, которое даёт генератор, присваивается элементу $V(1, 2)$ и т. д. Так вырабатываются все элементы массива V .

Далее вырабатывается второй двумерный массив, который будет обозначен через K . Элементы этого массива являются нулями и единицами, причём если, например, $K(25, 16) = 0$, то это значит, что узел с координатами $X = 25, Y = 16$ заблокирован, а если $K(25, 16) = 1$, то узел неблокирован. Для выработки массива K используется массив V и некоторое число t , лежащее в интервале от нуля до единицы. Изменяя число t , можно менять число заблокированных узлов.

Массив K получается по следующему правилу. Возьмём узел с координатами X и Y . Если $V(X, Y) \leq t$, то $K(X, Y) = 1$, если $V(X, Y) > t$, то $K(X, Y) = 0$. В первом случае узел с координатами X и Y считается неблокированным, во втором — заблокированным. Так как величины V равномерно распределены в интервале от нуля до единицы, то, полагая t близким к нулю, получим, что почти все узлы заблокированы. И наоборот, если t близко к единице, то почти все узлы неблокированы. При $t = \frac{1}{2}$ число заблокированных и неблокированных узлов должно быть примерно одинаковым.

Пользуясь функцией распределения случайных чисел, которые даёт генератор, можно связать число t со средней долей неблокированных узлов x , которая получается в результате описанной выше процедуры. Оказывается (см. упр. 8), что $t = x$. Это равенство, однако, справедливо, если взять очень большое число узлов N или если выработать много массивов K с одним и тем же t ,

а затем усреднить полученные в каждом массиве доли неблокированных узлов. Для каждого конкретного массива это равенство может слегка нарушаться в ту или иную сторону, причём чем больше N , тем точнее оно выполняется.

Итак, в памяти ЭВМ записан массив V , а из него можно получить массив K , описывающий, какой узел заблокирован, а какой нет. Вид массива K регулируется числом t , которое примерно совпадает с долей неблокированных узлов, получаемых в этом массиве. Плавно меняя t , можно получить распределения заблокированных и неблокированных узлов с плавно меняющейся концентрацией неблокированных узлов x .

Упражнение

8. Докажите, что средняя доля x неблокированных узлов в массиве K равна t .

4.7. Поиск путей протекания

Допустим, что составлен массив V , задано число t и найден массив K , содержащий определенную долю неблокированных узлов.

Теперь ЭВМ точно знает, какой узел заблокирован, а какой нет, и начинается второй этап программы — поиск путей протекания. Допустим, что изучается протекание слева направо. Прежде всего все единицы, находящиеся в самом левом столбце ($X = 1$), переименовываются в двойки. Переименование состоит в том, что в ячейке памяти, соответствующей данному элементу массива K ,

стирается единица и записывается двойка. В памяти ЭВМ составляется список координат узлов, переименованных в двойки. Затем ЭВМ изучает каждый узел этого списка. Она вычисляет, какие узлы являются ближайшими соседями изучаемого узла, и запрашивает у массива K сведения относительно этих соседей. Если ближайший сосед оказался единицей, то он переименовывается в двойку, а его координаты заносятся в новый список. По окончании изучения первого списка в памяти ЭВМ оказывается список доек «второго поколения», т. е. список единиц, переименованных в двойки, благодаря тому, что они находились в контакте с двойками первого поколения.

В целях экономии памяти ЭВМ на этом этапе первый список стирается — он больше не нужен, а соответствующие ячейки памяти освобождаются. Машина переходит к изучению второго списка и образованию списка доек третьего поколения. По окончании стирается второй список и начинается изучение третьего. Оно сопровождается составлением четвёртого списка и т. д.

В ходе этого процесса число доек в массиве K увеличивается. Двойки — это неблокированные узлы, связанные путём протекания с каким-либо неблокированным узлом крайнего левого столбца, т. е. двойками отмечаются пути протекания.

Процесс поиска путей протекания прекращается в двух случаях:

1. На правой стороне квадрата появилась хотя бы одна двойка. ЭВМ фиксирует, что при данном значении t протекание существует.

2. На правой стороне квадрата двоек нет, и изучение очередного списка не привело к образованию ни одной новой двойки. Это значит, что все пути порвались, и при данном t протекания нет.

4.8. Определение порога

Допустим, что при данном t протекание существует. Тогда ЭВМ уменьшает t и, используя тот же самый массив V , находит новый массив K с уменьшенным числом неблокированных узлов. Снова происходит поиск путей протекания. Если опять фиксируется протекание, то число t ещё уменьшается и так происходит до тех пор, пока при некотором t не обнаруживается отсутствие протекания. Тогда интервал между этим значением t и минимальным значением, при котором протекание ещё было, делится пополам, и при этом промежуточном значении t производится поиск путей протекания. Если теперь оказывается, что протекания нет, то интервал между этим последним значением и минимальным значением, при котором протекание есть, снова делится пополам. Если же протекание есть, то пополам делится интервал между последним значением t и тем значением, при котором протекания не было.

Таким образом, порог протекания берётся «в вилку», которую можно сколь угодно сужать. Если при первом выбранном значении t протекания не было, то нужно увеличивать t до тех пор, пока оно не возникнет, а потом снова делать «вилку». Такой метод позволяет найти значение t , соответствующее порогу протекания с любой степенью точности. При этом значении t вычисляется доля

неблокированных узлов x , которая, как уже говорилось, близка к значению t , но не обязательно ему равна. Это значение x и объявляется порогом протекания, полученным в данном опыте. На рис. 4.2 показан первый путь протекания слева направо, появившийся в квадрате 30×30 .

```

200000011100222220101000001110
002201001000022200111011000100
022220000000222020000000001111
222000011110220000220011110110
022010010000222000222200000110
2220100020202022222200001100
00222000022202220220200011110
20222220000202202202222000010
2022222222020222222222220000
0022220000101002220002222011
2022222200002022220000222200
0222222222222222222222002001
022020222200220222002222000001
022000022222022222220201111
2000110002000000222220000000
0000010000000002222200100102
01000111111000000222220000002
200000000010001022222200002
0000111010100000222222200102
00000010010102222200020010022
00022220000110022220000000002
222222222200100022000111001002
0222222001001002200000000002
02222022200100022202222000002
22222200001000220002222202002
2022200100100222222222222222
20022001000222200022200022222
00020000100222202222222202022
0000010000102222000022202022
0100011100000220202222220002

```

Рис. 4.2. Картина распределения нулей, единиц и двоек в момент возникновения протекания. Показан путь, по которому двойки "просочились" с левой стороны квадрата на правую. В данном случае ЭВМ не прекратила работу при появлении первой двойки на правой стороне квадрата, а продолжала её, пока не перестали появляться новые двойки.

Затем производится много идентичных опытов, использующих разные наборы случайных чисел в массиве V . Это соответствует изменению случайной последовательности заблокированных узлов в эксперименте с экранной сеткой. Результаты этих опытов позволяют найти среднее значение порога протекания $x_c(N)$ при задан-

ном числе узлов N . (Для этого нужно просто сложить все полученные значения порогов и поделить на число опытов.) Для нахождения истинного порога протекания $x_c = \lim_{N \rightarrow \infty} x_c(N)$ нужно менять число узлов в квадрате N и получать зависимость $x_c(N)$. Для этой зависимости нужно подобрать аналитическое выражение вида⁴

$$x_c(N) = x_c(\infty) + \frac{D}{N^\gamma}, \quad (4.5)$$

т. е. выбрать три величины, $x_c(\infty)$, D и γ так, чтобы выражение (4.5) наилучшим образом описывало полученные с помощью ЭВМ результаты. Если это удаётся сделать таким образом, что $\gamma > 0$, то можно сказать, что величина $x_c(\infty)$ и равна предельному значению x_c . Действительно, согласно выражению (4.5) $\lim_{N \rightarrow \infty} x_c(N) = x_c(\infty)$. Точность этой процедуры будет тем лучше, чем больше получено данных, нужных для установления зависимости $x_c(N)$. Это, в свою очередь, упирается в быстроедействие и объём памяти используемой ЭВМ.

Упражнение

1. Рассмотрите внимательно рис. 4.2 и разберитесь в том, каким способом появлялись отдельные группы двоек.

⁴На вопрос о том, почему это выражение зависит от N именно степенным образом, строгого ответа нет, но результаты многочисленных расчётов указывают на то, что это так для всех задач теории протекания.

Часть II

Различные задачи теории протекания и их применения

Глава 5.

Задачи на плоских решётках

5.1. Мы сажаем фруктовый сад (задача связей)

Проектируется фруктовый сад огромного размера. Деревья в нём должны расти не как-нибудь, а регулярно. Они будут посажены в узлах какой-либо периодической решётки, нанесённой на поверхности земли. Таких решёток можно придумать много, но мы ограничимся тремя: квадратной, треугольной и шестиугольной (последняя называется ещё «медовые соты»). Все они изображены на рис. 5.1.

Естественно, что хотелось бы посадить деревья поближе друг к другу — земля дорога, — но по ряду причин сажать деревья близко друг к другу нельзя. Одна из причин состоит в том, что проектировщики опасаются заразных заболеваний деревьев. Допустим, что специалисты по болезням фруктовых деревьев сообщили следующие сведения ¹:

¹Автор не несёт ответственности за эти сведения, и потому призывает не относиться слишком серьёзно к практическим выводам, вытекающим из решения поставленной здесь задачи. Она дана лишь для того, чтобы показать возможности теории протекания.

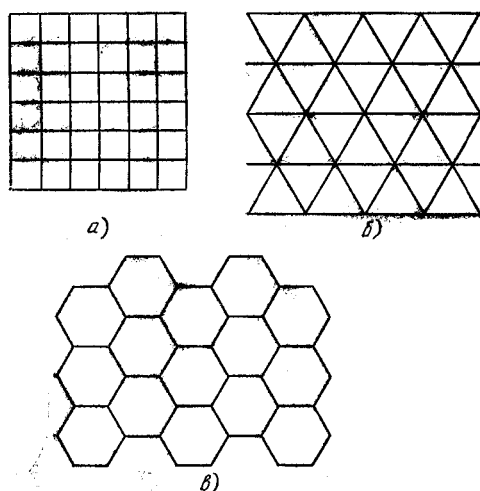


Рис. 5.1. Плоские решётки: а) квадратная; б) треугольная; в) шестиугольные («медовые соты»).

1. Заболевшее дерево может заразить только своих ближайших соседей.
2. Некоторые деревья, являющиеся ближайшими соседями, обязательно заражают друг друга, если одно из них заболело. В других случаях заражения не происходит. (Это может зависеть, например, от расстояния между ветвями данной пары деревьев.) Пару деревьев, внутри которой заражение обязательно происходит, назовём *связанной парой*.
3. Специалисты также предоставили в наше распоряжение вид функции $x(a)$ — вероятность того, что выбранная наугад пара ближайших соседей является связанной. Эта вероятность зависит от расстояния a между ближайшими соседями в данной решётке. Естественно, что функция $x(a)$ возрастает с уменьшением аргумента a : чем ближе друг к другу растут деревья,

тем легче заражают они друг друга.

Мы должны ответить на следующий вопрос: какое количество деревьев может заразить одно заболевшее дерево? Ответ на этот вопрос может быть только вероятностным. Если данное дерево образует с кем-либо из своих соседей связанную пару, то происходит заражение. Заболевшие деревья в свою очередь заражают своих соседей и т. д. Поэтому можно спросить лишь: какова вероятность того, что заболевшее дерево заразит определённое количество деревьев сада?

Для дальнейшего удобно перейти к кластерной терминологии, введённой в предыдущей главе. Будем считать, что между двумя соседними узлами, в которых находятся деревья, образующие связанную пару, имеется связь, которую представим в виде проволоки, соединяющей два узла. Если деревья, являющиеся ближайшими соседями, не образуют связанную пару, то связь между ними (проволока) разорвана (рис. 5.2).

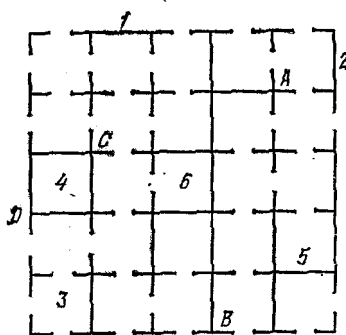


Рис. 5.2. Фрагмент квадратной решётки с разорванными связями. Изображены три кластера из двух узлов (1, 2, 3), один кластер из четырёх узлов (4), один кластер из шести узлов (5) и один кластер из десяти узлов (6)

Назовём два узла *связанными*, если между ними имеется целая связь или если они соединены непрерывной цепочкой узлов, являющихся ближайшими соседями и имеющими между собой целые связи (например, связаны друг с другом узлы A и B , а также C и D на рис. 5.2).

Совокупность связанных узлов назовём кластером. В применении к данной задаче важнейшее свойство кластера состоит в том, что заболевшее дерево заражает все деревья своего кластера и ни одного дерева вне его.

По определению, доля неразорванных связей равна x . Дальнейшие рассуждения такие же, как в гл. 3. При малых значениях x неразорванные связи располагаются по одиночке, почти все кластеры состоят из двух узлов, кластеры из трёх узлов редки, из четырёх — ещё более редки. При больших значениях x существует бесконечный кластер из связанных узлов. При $x = 1$ ему принадлежат все узлы системы. С уменьшением x часть узлов из него выпадает и, наконец, при некотором критическом значении x_c бесконечный кластер прекращает своё существование.

Бесконечный кластер — это и есть то бедствие, от которого мы призваны оградить сад. Пусть $P^{\text{CB}}(x)$ — вероятность того, что выбранный наугад узел принадлежит бесконечному кластеру. Если $x < x_c$, так что $P^{\text{CB}}(x) = 0$, то одно заболевшее дерево может заразить только несколько других деревьев. Если же $x > x_c$, то одно заболевшее дерево с вероятностью $P^{\text{CB}}(x)$ заражает в бесконечном саду бесконечное количество деревьев. Таким образом, в случае $x < x_c$ очаг болезни, случайно занесённый в сад, остаётся локализованным вблизи того места, куда он попал, а в случае $x > x_c$

болезнь может распространиться по всему саду.

Чтобы дать практические рекомендации, нужно найти значение x_c и приравнять к нему функцию $x(a)$, представленную специалистами. Отсюда определяется расстояние a_c , являющееся решением уравнения $x(a_c) = x_c$. Очаг болезни, возникший в саду, остаётся локализованным, если расстояние между деревьями больше a_c , и распространяется по всему саду в противоположном случае.

Сформулированная выше задача о нахождении x_c называется в теории протекания *задачей связей* в знак того, что случайным элементом являются здесь связи, которые с заданной вероятностью могут быть целыми или разорванными. На первый взгляд задача связей похожа на задачу узлов, обсуждавшуюся в предыдущих главах. Однако для заданной решётки эти задачи друг к другу не сводятся и имеют разные ответы.

В этой и последующих главах придётся несколько усложнить обозначения. Обозначим порог протекания задачи узлов через x_u , а порог протекания задачи связей через $x_{св}$. Эти пороги зависят от вида решётки. Примем следующие сокращённые обозначения для названий плоских решёток: К — квадратная, Т — треугольная, Ш — шестиугольная. Тогда $x_u(Ш)$ будет обозначать порог протекания задачи узлов на шестиугольной решётке, $x_{св}(Т)$ — порог задачи связей на треугольной решётке и т. д. Функцию $P^{св}(x)$, введённую в этом разделе для задачи связей, следует отличать от функции $P(x)$, определённой ранее для задачи узлов.

Задачу связей можно сформулировать не только на кластерном языке, но и как задачу о протекании с одной стороны квадрата на другую. Вспомним опыт с экранной сеткой, которым начи-

налась книга. Возможно, у некоторых читателей возник вопрос, почему нужно было блокировать узлы, т. е. разрезать сразу четыре проволоки, входящие в каждый узел, вместо того, чтобы разрезать по одиночке случайно выбранные проволоки (связи). Легко понять, что, разрезая связи, исследователи определили бы порог $x_{\text{св}}(K)$ вместо порога $x_{\text{у}}(K)$, который был найден в их эксперименте. Теперь можно объяснить, почему была выбрана задача узлов: как будет показано ниже, задача связей на квадратной решётке имеет точное аналитическое решение, из которого следует, что $x_{\text{св}}(K) = 0,5$. Поэтому ради $x_{\text{св}}(K)$ не было смысла ставить столь трудоёмкий эксперимент, а $x_{\text{у}}(K)$ известен лишь из приближённых вычислений.

Упражнение

1. Найдите функцию $P^{\text{св}}(x)$ при $1 - x \ll 1$ для трёх решёток, изображённых на рис. 5.1.

5.2. Неравенство, связывающее $x_{\text{св}}$ и $x_{\text{у}}$

При обсуждении задачи узлов вместо функции $P(x)$ часто вводят функцию $P^{\text{у}}(x)$, связанную с $P(x)$ соотношением

$$P(x) = xP^{\text{у}}(x). \quad (5.1)$$

По определению $P(x)$ есть вероятность того, что выбранный наугад узел принадлежит бесконечному кластеру. Её можно представить как произведение вероятностей двух независимых событий. Первое из них на языке задачи о ферромагнетике состоит в

том, что выбранный наугад узел оказался магнитным. Вероятность этого события равна x (см. упр. 1 к гл. 1). Второе событие состоит в том, что этот узел связан с бесконечным кластером из магнитных узлов. Таким образом, функция $P^Y(x)$, определённая формулой (5.1), есть вероятность того, что выбранный наугад магнитный узел связан с бесконечным кластером. Иными словами, $P^Y(x)$ представляет собой долю магнитных узлов, принадлежащих бесконечному кластеру, т. е. отношение числа узлов, принадлежащих бесконечному кластеру, к числу магнитных узлов. Напомним, что $P(x)$ представляет собой отношение числа узлов, принадлежащих бесконечному кластеру, к полному числу узлов. Естественно, что функция $P^Y(x)$ монотонно возрастает с ростом x , равна единице при $x = 1$ и равна нулю при $x \leq x_y$.

Английский математик Хаммерсли, тот, кто первым заговорил о теории протекания, доказал теорему, согласно которой

$$P^Y(x) \leq P^{CB}(x). \quad (5.2)$$

Обе функции, P^Y и P^{CB} , монотонно возрастают с увеличением аргумента x . Поэтому (рис. 5.3) из формулы (5.2) следует, что

$$x_{CB} \leq x_y, \quad (5.3)$$

т. е. для любой решётки (не обязательно плоской) пороговое значение для задачи связей не больше, чем для задачи узлов. Этот результат можно переписать в виде другого неравенства:

$$1 - x_{CB} \geq 1 - x_y, \quad (5.4)$$

и трактовать следующим образом. Допустим, что нужно заблокировать протекание электрического тока по сетке из проводов или

протекание жидкости по сетке труб, и сделать это можно, либо блокируя узлы, либо разрезая связи (провода или трубы), соединяющие эти узлы. Неравенство (5.4) означает, что, вырезая узлы, заблокировать систему легче, чем разрезая связи. Доля заблокированных узлов, при которой прекращается ток, меньше, чем доля разорванных связей. Результат этот представляется совершенно естественным, так как при блокировании одного узла разрывается не одна связь, а все связи, входящие в этот узел.

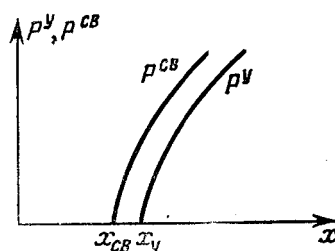


Рис. 5.3.

Упражнение

2. Найдите функцию $P_y(x)$ при $1 - x \ll 1$ для трёх решёток, изображённых на рис. 5.1. Сравните с результатом упр. 1 и проверьте справедливость неравенства (5.2).

Пояснение. В упражнении 1 к гл. 3 мы рекомендовали найти функцию $P(x)$ для задачи узлов при $1 - x \ll 1$. Однако, в этом месте речь шла лишь о главном члене функции, т. е. $P(x) = x$. Если подставить это выражение в формулу (5.1), получим $P^y =$

1. Это правильный результат в том смысле, что $\lim_{x \rightarrow 1} P^Y(x) = 1$. Теперь читателю предлагается найти малые члены, отличающие функцию $P^Y(x)$ от единицы. Разумеется, предел этих членов при $x \rightarrow 1$ равен нулю. Поэтому результат может быть представлен в виде $P^Y(x) = 1 - A(1 - x)^n$, где A и n — положительные численные коэффициенты, определяемые типом решётки.

5.3. Покрывающие и включающие решётки

Задача узлов является более общей, чем задача связей. Задача связей сводится к задаче узлов, но на другой решётке, которая называется покрывающей. Покрывающая решётка строится из исходной по следующим правилам:

1. Посередине каждой связи исходной решётки нужно поставить узел покрывающей решётки.
2. Два узла покрывающей решётки нужно связать друг с другом в том и только в том случае, если связи исходной решётки, на которых поставлены эти два узла, сходятся в узле исходной решётки.

Результатом этого построения является новая периодическая решётка, которая называется покрывающей по отношению к исходной.

На рис. 5.4 изображена покрывающая решётка для случая, когда исходной является квадратная. Тонкими линиями показана исходная квадратная решётка. На том месте, где исходная решётка имеет узлы, нарисованы дуги. Покрывающая решётка состоит из

тонких и толстых линий, но в пересечениях тонких линий нет узлов покрывающей решётки. Они находятся только в пересечениях толстых линий и отмечены на рисунке точками.

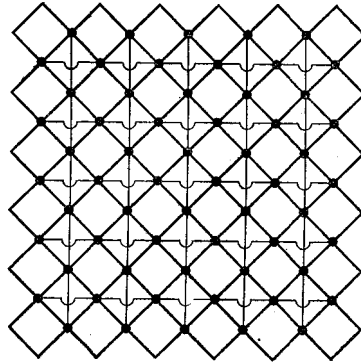


Рис. 5.4. Покрывающая решётка для квадратной решётки.

Каждая связь исходной решётки стыкуется с тремя другими связями на одном своём конце и ещё с тремя связями на другом конце. Поэтому каждый узел покрывающей решётки должен быть связан с шестью другими узлами. Это показано на рис. 5.4. Каждый узел связан с четырьмя другими узлами толстыми линиями и ещё с двумя узлами — тонкими линиями.

Допустим теперь, что на исходной решётке сформулирована задача связей, т. е. определённая доля случайно выбранных связей разорвана.

Предположим, что если связь исходной решётки разорвана, то узел покрывающей решётки, который стоит на данной связи, блокирован. Теперь возникла задача узлов на покрывающей решётке. Узлы на ней оказались блокированными случайно, а доля блокированных узлов равна доле разорванных связей на исходной решётке.

Заметим, что в задаче связей существование бесконечного кластера из связанных друг с другом узлов всегда означает существование бесконечного кластера из стыкующихся друг с другом неразорванных связей. И наоборот, отсутствие бесконечного кластера из узлов означает отсутствие кластера из связей.

По построению покрывающей решётки видно, что существование бесконечного кластера из неразорванных связей на исходной решётке означает существование бесконечного кластера из неблокированных узлов на покрывающей решётке, и, наоборот, отсутствие бесконечного кластера связей на исходной решётке означает отсутствие бесконечного кластера узлов на покрывающей решётке.

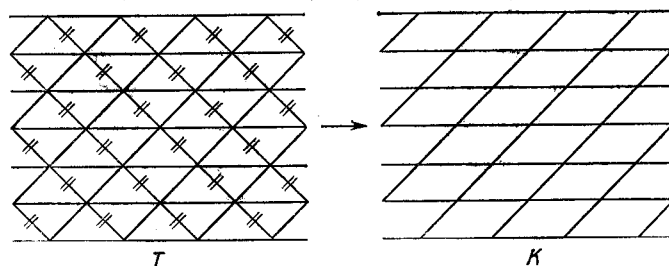


Рис. 5.5. Треугольная решётка включает квадратную.

Отсюда следует, что порог протекания задачи связей на исходной решётке равен порогу протекания задачи узлов на покрывающей решётке. Если обозначить исходную решётку буквой L , а покрывающую — L_{Π} , то сказанное можно записать в виде формулы

$$x_{\text{св}}(L) = x_{\text{у}}(L_{\Pi}). \quad (5.5)$$

Ряд неравенств между порогами протекания для различных решёток можно получить, используя понятие *включающей* решётки. Допустим, что решётка L получается из решётки $L_{\text{вкл}}$ с помощью вычеркивания определённого количества связей. Тогда говорят, что *решётка $L_{\text{вкл}}$ включает решётку L* .

Обратимся, например, к треугольной решётке. Если уничтожить в ней все связи, отмеченные на рис. 5.5 двумя чёрточками, то она перейдёт в решётку, показанную в правой части рисунка. Легко увидеть, что с точки зрения задачи узлов или связей эта новая решётка эквивалентна квадратной. Действительно, то обстоятельство, что углы между связями на новой решётке не равны 90° , не играет никакой роли, если обсуждается вопрос о связи различных узлов друг с другом (новую решётку можно просто мысленно «распрямить»). Порог протекания задачи связей (и задачи узлов!) на этой решётке точно равен порогу протекания на квадратной решётке. Поэтому говорят, что треугольная решётка включает в себя квадратную.

Допустим теперь, что определённая доля связей включающей решётки разорвана. Связи включающей решётки можно разделить на те связи, которые являются общими для неё и для включённой решётки и те, которые являются специфическими для включающей решётки (последние помечены двумя чёрточками на левой части рис. 5.5). Поскольку связи разрываются совершенно случайно, то доля разорванных связей и в той, и в другой категории связей будет одинакова и равна общей по всей решётке доле разорванных связей. Поэтому, чтобы получить включённую решётку с той же долей разорванных связей, необходимо дополнительно

разорвать на включающей решётке те связи, которые остались у этой решётки целыми, но являются для неё специфическими, т. е. вовсе отсутствуют на включённой решётке.

Отсюда видно, что число целых связей, выходящих из каждого узла включающей решётки, не меньше (больше или равно) числа целых связей, выходящих из того же узла включённой решётки. Поэтому вероятность того, что выбранный наугад узел принадлежит бесконечному кластеру для включающей решётки, не меньше, чем для включённой. Итак, получено неравенство

$$P_L^{\text{СВ}}(x) \leq P_{L_{\text{ВКЛ}}}^{\text{СВ}}(x). \quad (5.6)$$

В левой части неравенства (5.6) стоит функция $P^{\text{СВ}}(x)$, вычисленная для включённой решётки, а справа — для включающей. Подобно тому, как из неравенства (5.2) следует неравенство (5.3), из неравенства (5.6) следует, что

$$x_{\text{СВ}}(L_{\text{ВКЛ}}) \leq x_{\text{СВ}}(L), \quad (5.7)$$

т. е. $x_{\text{СВ}}$ у включающей решётки ниже, чем у включённой.

Как уже говорилось, треугольная решётка включает квадратную. Поэтому

$$x_{\text{СВ}}(T) \leq x_{\text{СВ}}(K). \quad (5.8)$$

Допустим теперь, что из квадратной решётки вычеркнуты некоторые связи, как показано на рис. 5.6. При этом получится решётка, показанная в правой части рисунка. Посмотрите на неё внимательно. Она эквивалентна шестиугольной. Нужно немного

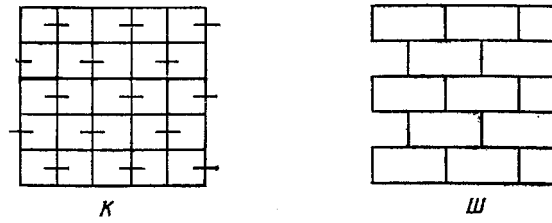


Рис. 5.6. Квадратная решётка включает шестиугольную.

потянуть её вверх, немного деформировать связи, и её ячейки (рис. 5.7) превратятся в «медовые соты», показанные на рис. 5.1. Таким образом, квадратная решётка *включает* шестиугольную и, следовательно,

$$x_{\text{св}}(K) \leq x_{\text{св}}(\overline{\text{III}}) \quad (5.9)$$

Из неравенств (5.8) и (5.9) следует, что

$$x_{\text{св}}(T) \leq x_{\text{св}}(\overline{\text{III}}) \quad (5.10)$$

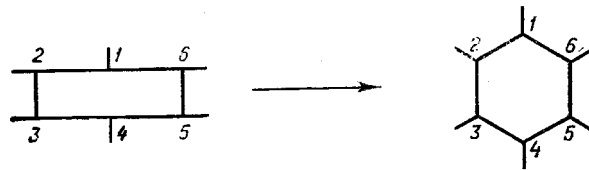


Рис. 5.7. Преобразование одной ячейки, изображённой в правой части рис. 5.6, в «медовые соты».

Обратимся теперь к задаче узлов. Допустим, что на включающей и включённой решётках блокированы одни и те же узлы (соответственно доля заблокированных узлов здесь и там одинакова). Пусть на включённой решётке существует бесконечный кластер

из целых узлов. Это значит, что он существует и на включающей решётке, поскольку дополнительные связи могут только способствовать его появлению. Если же известно, что бесконечный кластер существует при заданной доле узлов на включающей решётке, то отсюда нельзя сделать никаких выводов по поводу существования бесконечного кластера на включённой решётке. Уничтожение ряда связей между узлами при переходе от включающей решётки к включённой может оказаться для бесконечного кластера роковым. Таким образом, порог протекания задачи узлов на включённой решётке не может быть меньше, чем на включающей:

$$x_y(L_{\text{вкл}}) \leq x_y(L). \quad (5.11)$$

Отсюда следует цепочка неравенств для квадратной, треугольной и шестиугольной решёток:

$$x_y(T) \leq x_y(K) \leq x_y(\overline{\text{III}}), \quad (5.12)$$

точно такая же, как для задачи связей.

Обратимся ещё раз к рис. 5.4, где показана покрывающая решётка для квадратной решётки. Допустим, что связи, изображённые тонкими линиями, уничтожены. Легко понять, что в результате получилась квадратная решётка, только повернутая на 45° , что, разумеется, совершенно не существенно для задач теории протекания.

Итак, покрывающая решётка для квадратной решётки включает в себя квадратную решётку. Согласно формуле (5.5)

$$x_y(L_{\text{п}}) \leq x_{\text{св}}(K), \quad (5.13)$$

где под L_{Π} понимается покрывающая решётка, изображённая на рис. 5.4. Но из неравенства (5.11) и из того, что данная покрывающая решётка включает квадратную, следует

$$x_{\text{У}}(L_{\Pi}) \leq x_{\text{У}}(\text{К}). \quad (5.14)$$

Из неравенств (5.13) и (5.14) вытекает, что

$$x_{\text{СВ}}(\text{К}) \leq x_{\text{У}}(\text{К}). \quad (5.15)$$

Таким образом, мы получили, что для квадратной решётки порог задачи связей выше, чем порог задачи узлов. Неравенство (5.15) является частным случаем общей теоремы Хаммерсли, которая выражается формулой (5.3) и была приведена нами без доказательства. При выводе (5.15) эта теорема не использовалась, и потому можно сказать, что приведённые выше соображения доказывают её для случая квадратной решётки.

5.4. Белое протекание и чёрное протекание

Взглянем теперь на задачу связей с несколько иной точки зрения. До этого момента говорилось, что существуют целые и разорванные связи, случайно распределённые по решётке, а кластером называлась совокупность узлов, соединённых целыми связями.

Задачу можно сформулировать более симметрично. Переименуем разорванные связи в «чёрные связи», а целые связи — в «белые связи». Совокупность узлов, соединённых белыми связями, будем называть белым кластером, а совокупность узлов, соединённых чёрными связями, — чёрным кластером (по старой терминологии

кластером назывался «белый» кластер). Пусть доля белых связей по-прежнему обозначается через x . Долю чёрных связей обозначим через q . Так как каждая связь либо чёрная, либо белая, то $q = 1 - x$.

В такой формулировке можно говорить и о протекании по белым, и о протекании по чёрным связям.

При малой концентрации x белых связей не существует бесконечного белого кластера, но существует бесконечный чёрный кластер, т. е. бесконечный кластер из узлов, связанных чёрными связями. Наоборот, при малой концентрации q чёрных связей (т. е. при значениях x , близких к единице) существует бесконечный белый кластер и нет бесконечного чёрного кластера.

При изменении x от нуля до единицы происходят два события: исчезает чёрный бесконечный кластер и появляется белый, или, что то же самое, исчезает протекание по чёрным связям и возникает протекание по белым связям. Но в какой последовательности происходят эти события?

Белые и чёрные связи ничем не отличаются друг от друга, кроме названия. Поэтому очевидно, что критическая концентрация $q_{\text{св}}$, при которой возникает протекание по чёрным связям, равна концентрации $x_{\text{св}}$, при которой возникает протекание по белым связям.

Таким образом, при увеличении x протекание по белым связям возникает, когда $x = x_{\text{св}}$, а протекание по чёрным связям исчезает, когда $x = 1 - q_{\text{св}} = 1 - x_{\text{св}}$. Последовательность, в которой происходят эти события, зависит от знака разности $x_{\text{св}} - 0,5$.

Если для осуществления протекания по белым связям необходимо, чтобы белых связей было больше половины (это значит, что в точке протекания белых связей должно быть больше, чем чёрных), то при увеличении x сначала исчезает протекание по чёрным связям, а потом появляется протекание по белым (рис. 5.8, а). В области I на рис. 5.8, а существует только протекание по чёрным связям, в области III — только по белым связям, а в области II отсутствует протекание и по белым, и по чёрным связям.

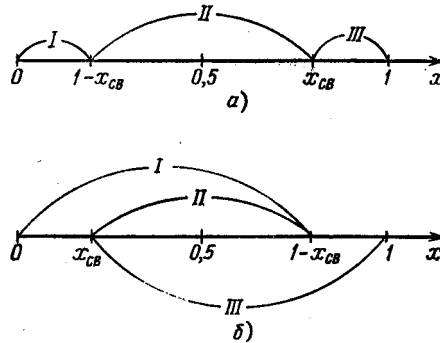


Рис. 5.8.

Если $x_{CB} < 0,5$, то сначала появляется протекание по белым связям и лишь потом исчезает протекание по чёрным. В области I на рис. 5.8, б существует протекание по чёрным связям, в области III — по белым связям, а в области II — и по чёрным, и по белым.

Такая же симметричная формулировка существует и для задачи узлов. Напомним, что в этом случае все связи — целые, зато узлы бывают двух сортов. В задаче о сетке они назывались целыми и заблокированными. В задаче о ферромагнетике — магнитными и немагнитными. Теперь, так же как и в задаче связей, вводится

универсальное обозначение: целые, или магнитные, узлы называются белыми, а заблокированные, или немагнитные, — чёрными. Белые узлы называются связанными, если они являются соседями друг друга или соединены цепочкой белых узлов, являющихся ближайшими соседями. Точно так же могут быть связанными друг с другом и чёрные узлы.

Можно говорить о протекании по белым и чёрным узлам. Если $x_y > 0,5$, то существует область значений x , в которой нет протекания ни по белым, ни по чёрным узлам ($1 - x_y < x < x_y$). Если $x_y < 0,5$, то в области $x_y < x < 1 - x_y$ существует протекание и по белым, и по чёрным узлам.

Симметричный подход оказывается конструктивным, потому что иногда, рассматривая решётку, удаётся выяснить, что на ней не может быть ни белого, ни чёрного протеканий или что на ней обязательно должно быть одно из этих протеканий. Исходя из этого, легко сделать некоторые заключения о пороге протекания.

Рассмотрим, например, задачу узлов на треугольной решётке. Допустим, что по белым узлам есть протекание. Легко увидеть, что при этом не может быть протекания по чёрным узлам. Допустим, что изучается протекание тока слева направо на сетке очень большого размера, как это делалось в эксперименте, с описания которого начиналась эта книга. Только теперь сетка сделана не в виде квадратной, а в виде треугольной решётки. Легко понять, что треугольная решётка устроена так, что наличие пути протекания по белым узлам слева направо исключает возможность протекания по чёрным узлам сверху вниз. Действительно, чёрные узлы не могут «пробиться» через ломаную линию, проходящую через всю

сетку слева направо и соединяющую белые узлы.

Как объяснялось в гл. 3, существование бесконечного кластера обеспечивает протекание в любом направлении, если размер системы достаточно велик. Поэтому следует сделать вывод, что в треугольной решётке при одном и том же значении x не может существовать бесконечного кластера и из белых, и из чёрных узлов, т. е. не может быть протекания и по белым, и по чёрным узлам. Отсюда следует, что

$$x_{\text{Y}}(\text{T}) \geq 0,5.$$

Такой же вывод можно сделать и для квадратной решётки: $x_{\text{Y}}(\text{K}) \geq 0,5$.

Для треугольной решётки доказана теорема, согласно которой $x_{\text{Y}}(\text{T}) = 0,5$. Мы не можем привести здесь её доказательство, но суть дела сравнительно легко понять, используя представление о белом и чёрном протекании. Изобразив различные конфигурации чёрных и белых узлов, можно увидеть, что отсутствие протекания по белым узлам слева направо обязательно влечёт за собой протекание по чёрным узлам сверху вниз (квадратная решётка не обладает таким свойством!).

Итак, на треугольной решётке *не может быть протекания и по белым, и по чёрным узлам сразу и не может не быть протекания хотя бы по одному сорту узлов*. Отсюда следует, что область II на рис. 5.8 (на этом рисунке следует теперь заменить $x_{\text{СВ}}$ на x_{Y}), вырождается в точку, т. е. $x_{\text{Y}}(\text{T}) = 0,5$.

В случае задачи связей такого рода исследования удобно делать с помощью понятия дуальной решётки.

5.5. Дуальные решётки

Дуальными могут быть только плоские решётки. Под плоскими решётками понимают такие решётки, которые могут быть размещены в плоскости, причем так, что связи решётки пересекаются только в тех точках, где находятся узлы решётки. Например, все решётки, изображённые на рис. 5.1, являются плоскими, а покрывающая решётка на рис. 5.4 — не плоская, потому что в местах, показанных дугами, её связи пересекаются, а узлов в точках пересечения нет. (Дуги представляют собой как бы мосты, обеспечивающие развязку путей, идущих слева направо и сверху вниз.)

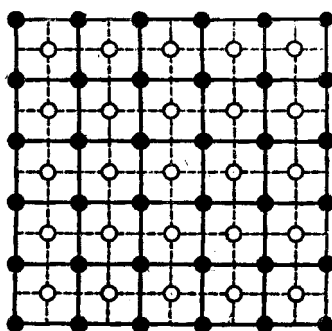


Рис. 5.9. Построение, показывающее, что квадратная решётка дуальна сама по себе. Сплошными линиями и чёрными кружками показаны связи и узлы исходной решётки, штриховыми линиями и светлыми кружками — связи и узлы дуальной решётки.

Каждая плоская решётка разбивает плоскость на ячейки. Решётка L^D называется дуальной к решётке L , если каждая связь L^D пересекает одну и только одну связь решётки L , и, наоборот, каждая связь L пересекает одну и только одну связь L^D .

Кроме того, в каждой ячейке решётки L должен находиться только один узел решётки L^D (и наоборот).

Как видно из определения, свойство дуальности является взаимным: если L^D дуальна к L , то L дуальна к L^D . На рис. 5.9, 5.10 показано, что квадратная решётка является дуальной квадратной решётке, а треугольная и шестиугольная решётки дуальны друг другу.

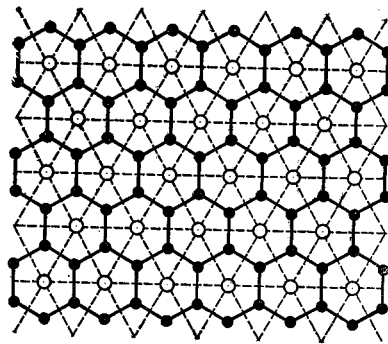


Рис. 5.10. Построение, показывающее, что треугольная решётка дуальна к шестиугольной (и наоборот). Обозначения те же, что и на рис. 5.9.

Вернёмся теперь к задаче связей. Примем, что если некоторая связь на исходной решётке белая (неразорванная), то пересекающая её связь на дуальной решётке чёрная (разорванная). Поэтому если концентрация белых связей на исходной решётке равна x , то концентрация белых связей на дуальной решётке $q = 1 - x$.

Ниже будет удобно говорить о пороге протекания как о таком значении концентрации белых связей, при котором впервые возникает (или исчезает) проводимость слева направо в очень большой сетке с припаянными к ней электродами (см. рис. 1.1). Это та же постановка задачи, которая обсуждалась в самом начале книги,

только теперь нарушаются не узлы сетки, а связи.

Допустим, что на исходной решётке существует путь протекания по белым связям слева направо. Как легко понять, это означает отсутствие пути протекания по белым связям на дуальной решётке сверху вниз. Действительно, белая связь исходной решётки по определению пересекается только чёрной связью дуальной решётки. Поэтому, если на исходной решётке существует нигде не обрывающаяся ломаная линия из белых связей, пересекающая всю сетку слева направо, то это значит, что белые связи дуальной решётки ни в каком месте не могут «пробиться» через неё, чтобы пройти сверху вниз.

Если решётка достаточно велика, то наличие пути слева направо по исходной решётке означает, что доля белых связей больше пороговой:

$$x > x_{\text{св}}(L). \quad (5.16)$$

Отсутствие же пути сверху вниз по белым связям дуальной решётки означает, что доля белых связей дуальной решётки $q = 1 - x$ меньше пороговой доли для дуальной решётки, т. е. $1 - x < x_{\text{св}}(L^{\mathcal{D}})$, или

$$x > 1 - x_{\text{св}}(L^{\mathcal{D}}) \quad (5.17)$$

Согласно сказанному выше все значения x , удовлетворяющие неравенству (5.16), удовлетворяют также и неравенству (5.17). Отсюда следует, что $x_{\text{св}}(L) \geq 1 - x_{\text{св}}(L^{\mathcal{D}})$, или

$$x_{\text{св}}(L) + x_{\text{св}}(L^{\mathcal{D}}) \geq 1. \quad (5.18)$$

Для квадратной решётки $L = L^{\mathcal{A}}$, и из неравенства (5.18) следует, что

$$x_{\text{св}}(\text{К}) \geq 0,5, \quad (5.19)$$

а если использовать ещё неравенство (5.15), то получим

$$x_{\text{у}}(\text{К}) \geq 0,5, \quad (5.20)$$

т. е. тот же вывод, что и в предыдущем разделе. Вспомните, что эксперимент с квадратной сеткой, описанный в гл. 1, дал значение $x_{\text{у}}(\text{К}) = 0,59$, что не противоречит неравенству (5.20).

Для двух решёток — квадратной и треугольной — строго доказано, что знак неравенства в формуле (5.18) следует заменить знаком равенства, т. е.

$$x_{\text{св}}(L) + x_{\text{св}}(L^{\mathcal{A}}) = 1. \quad (5.21)$$

Отсюда сразу вытекают два новых результата:

$$x_{\text{св}}(\text{К}) = 0,5 \quad (5.22)$$

и

$$x_{\text{св}}(\text{Т}) + x_{\text{св}}(\text{Ш}) = 1. \quad (5.23)$$

Строгое доказательство формулы (5.21) здесь не приводится: оно требует введения ряда новых понятий, которые не понадобятся в дальнейшем. Однако этой формуле можно дать достаточно ясную интерпретацию. Рисуя различные конфигурации белых и чёрных связей, можно увидеть, что для квадратных и треугольных решёток отсутствие протекания по белым связям слева направо

на исходной решётке всегда означает наличие протекания сверху вниз по белым связям дуальной решётки. Примем это утверждение на веру. Допустим, что доля белых связей такова, что на исходной решётке нет протекания слева направо. Для сетки достаточно большого размера это означает, что

$$x < x_{\text{св}}(L). \quad (5.24)$$

При этом по дуальной решётке есть протекание сверху вниз. Доля белых связей на ней равна $1 - x$. Следовательно, $1 - x > x_{\text{св}}(L^{\text{Д}})$, или

$$x < 1 - x_{\text{св}}(L^{\text{Д}}). \quad (5.25)$$

Согласно сказанному выше все значения x , удовлетворяющие неравенству (5.24), должны удовлетворять и неравенству (5.25). Это возможно, если $1 - x_{\text{св}}(L^{\text{Д}}) \geq x_{\text{св}}(L)$, или

$$x_{\text{св}}(L) + x_{\text{св}}(L^{\text{Д}}) \leq 1. \quad (5.26)$$

Из неравенств (5.18) и (5.26) следует равенство (5.21).

Формула (5.23) не позволяет найти по отдельности $x_{\text{св}}(\text{T})$ и $x_{\text{св}}(\text{Ш})$. Однако, воспользовавшись известным в теории электрических цепей преобразованием «звезда — треугольник», можно получить ещё одно соотношение, связывающее пороги протекания задачи связей на треугольной и шестиугольной решётках. В результате становится известным каждый из порогов

$$\begin{aligned} x_{\text{св}}(\text{T}) &= 2 \sin\left(\frac{\pi}{18}\right) \approx 0,347296, \\ x_{\text{св}}(\text{Ш}) &= 1 - x_{\text{св}}(\text{T}) = 1 - 2 \sin\left(\frac{\pi}{18}\right) \approx 0,652704. \end{aligned} \quad (5.27)$$

Упражнение

3. Посмотрите на рис. О.4 (с. 257), изображающий преобразование «звезда — треугольник». Если, глядя на него, вы поймёте, как написать соотношение, о котором только что шла речь, и получите формулы (5.27), — то это прекрасно! Если не удастся, то не огорчайтесь — задача эта не простая. Вывод формул (5.27), сделанный впервые в 1963 г. английскими математиками Сайксом и Эссамом, явился событием в теории протекания. Прочитайте внимательно текст, помещённый рядом с рисунком, и вы, наверное, получите удовольствие от красивого применения теории вероятностей, которое привело к столь нетривиальному результату.

5.6. Результаты для плоских решёток

Нам осталось только привести сводную таблицу порогов протекания для плоских решёток (табл. 5.1).

Таблица 5.1.

Тип решётки	$x_{св}$	x_y
Треугольная	0,3473	0,5
Квадратная	0,5	0,59
Шестиугольная	0,6527	0,70

Лишь два числа в этой таблице, а именно, $x_y(K)$ и $x_y(Ш)$, получены приближёнными методами. Все остальные представляют

собой результаты точных решений. Как будет показано в следующей главе, с трёхмерными решётками дело обстоит существенно хуже. Для них не получено *ни одного* точного решения. Это не должно казаться странным. Для аналитического решения задач теории протекания не существует никакого метода. Каждое точное решение, о котором шла речь, производит впечатление чуда. Поэтому на наш взгляд следует скорее удивляться тому, как много таких решений уже известно.

Упражнение

4. Вернёмся к фруктовому саду, о котором шла речь в начале этой главы. Допустим, что расстояние между деревьями выбирается из условия, что доля связанных пар равна пороговому значению. Пусть нам известна функция $a(x)$, представляющая расстояние между соседними деревьями в зависимости от доли связанных пар x . Естественно, что чем больше x , тем меньше расстояние a , так как чем ближе друг к другу растут деревья, тем легче они заражают друг друга. Если $x = x_{\text{св}}$, то расстояние между деревьями равно $a(x_{\text{св}})$. Найдите площадь, приходящуюся на одно дерево при таком выборе расстояния, для трёх разных решёток. Решётка с наименьшей площадью на дерево является самой выгодной. Можно ли сказать, какая решётка приводит к меньшей площади на дерево, основываясь только на том, что функция $a(x)$ монотонно убывает с ростом x ?

5.7. Ориентированное протекание

Допустим теперь, что в узлах плоской решётки мы посадили не сад, а лес, и наш лес загорелся. Некоторые деревья, являющиеся ближайшими соседями, сильно перепутаны ветвями и передают огонь друг на друга. В соответствии с общей терминологией будем говорить, что узлы, в которых находятся такие деревья, соединены белыми связями. Другие деревья, которые тоже являются ближайшими соседями, не зажигают друг друга. Будем говорить, что соответствующие узлы соединены чёрными связями. Белые и чёрные связи беспорядочно разбросаны по решётке, причём доля белых связей равна x .

Задача состоит в том, чтобы найти критическое значение $x_{\text{св}}$ такое, что при $x < x_{\text{св}}$ возникший очаг пожара остаётся локализованным, а при $x > x_{\text{св}}$ огонь распространяется по всему лесу.

Легко понять, что это просто ещё один пример задачи связей. Значение $x_{\text{св}}$ можно найти, заглянув в табл. 5.1 (с. 112).

Допустим теперь, что во время пожара дует сильный ветер, так что огонь распространяется только по ветру. Это приводит к новой интересной задаче, которая называется задачей об *ориентированном (или направленном) протекании*.

Ориентированные белые связи показаны толстыми линиями со стрелкой, чёрные связи показаны тонкими линиями. Огонь может распространяться слева направо по пути 1 и не может распространяться по пути 2. Двумя чёрточками показаны два участка пути 2, на которых огню пришлось бы двигаться против ветра.

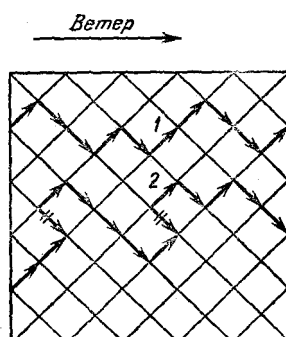


Рис. 5.11. Ориентированные белые связи показаны толстыми линиями со стрелкой, чёрные связи показаны тонкими линиями. Огонь может распространяться слева направо по пути 1 и не может распространяться по пути 2. Двумя чёрточками показаны два участка пути 2, на которых огню пришлось бы двигаться против ветра.

Будем считать, что лес посажен в узлах квадратной решётки, а ветер дует в направлении диагонали квадратов. На рис. 5.11 направление ветра показано стрелкой, а решётка повёрнута на 45° по отношению к обычному способу её изображения. Условия задачи формулируются теперь следующим образом. Каждая белая связь превращается в вектор, стрелка которого поставлена так, что проекция вектора на направление ветра положительна. Чёрные связи по-прежнему не переносят огонь ни в каком направлении, а белые переносят его только в направлении стрелки. Требуется определить критическую долю белых связей, начиная с которой одно загоревшееся дерево в бесконечно большом лесу может вызвать пожар, уходящий на бесконечно большое расстояние.

На рис. 5.11 толстыми линиями со стрелками изображены белые связи и показаны два пути протекания: 1 и 2. Путь 1 может изображать движение огня, а путь 2 не соответствует усло-

виям ориентированного протекания: в двух местах, показанных чёрточками, движение происходит против направления белой связи, т. е. против ветра.

Итак, если при неориентированном протекании белые связи использовались в обоих направлениях, то при ориентированном протекании они могут использоваться только в одном направлении. Отсюда следует, что критическая доля белых связей при ориентированном протекании $x_{\text{СВ}}^{\text{О}}$ не может быть меньше, чем при обычном протекании, т. е. $x_{\text{СВ}}^{\text{О}} \geq x_{\text{СВ}}$.

В настоящее время для ряда задач с ориентированным протеканием получено приближённое решение. В частности, для описанной выше задачи связей на квадратной решётке $x_{\text{СВ}}^{\text{О}} = 0,63$ или $0,64$ (результаты, полученные разными методами, несколько отличаются друг от друга). Напомним, что для неориентированной задачи связей на квадратной решётке $x_{\text{СВ}} = 0,5$.

Многие физические задачи сводятся к ориентированному протеканию. В качестве примера можно назвать движение электрона в сильном электрическом поле в случайно-неоднородной среде. В такой среде, где свойства меняются случайным образом от точки к точке, на пути движения электрона возникают препятствия, которые он должен огибать. В то же время однородное электрическое поле играет роль ветра, который гонит электрон в одну сторону.

Ориентированное протекание возникает и в задаче об электропроводности проволочной сетки (гл. 1), если принять, что в каждую связь между узлами сетки включён диод (выпрямитель), пропускающий ток только в одном направлении. Тогда доля разорванных связей, при которой прекращается ток через сетку, соот-

ветсвует порогу ориентированного протекания. Рассматривалась и смешанная задача, когда диоды включены не во все связи сетки.

Глава 6.

Объёмные решётки и приближённые оценки порогов протекания

Как было показано в гл. 5, порог протекания существенно зависит от типа рассматриваемой решётки. Цель настоящей главы состоит в том, чтобы качественно объяснить, какие именно свойства решёток важны для порогов протекания. Поняв это, можно научиться, не решая задачи, предсказывать (с точностью порядка 10%) пороги протекания. Так как решёток существует великое множество, а расчёт одного порога требует (кроме хорошей квалификации!) около часа работы ЭВМ самого высшего класса, то умение сразу предсказать результат (хотя бы и не очень точно) представляется ценным.

Кроме того, задачи протекания не исчерпываются решёточными задачами. Как мы увидим ниже, для различных приложений важны большей частью задачи, заданные не на решётках. Оказывается, что идеи приближённой оценки, которые изложены в этой главе, весьма плодотворны для нерешёточных задач. На основании этих идей ряд порогов протекания для нерешёточных задач был предсказан с большой точностью задолго

до того, как эти задачи были решены на ЭВМ.

Чтобы набрать необходимый опыт, нужно выйти за рамки плоских решёток, рассмотренных в предыдущей главе, и обратиться к объёмным решёткам.

6.1. Объёмные решётки

Простейшей из объёмных решёток является *простая кубическая* (сокращённо ПК). Она изображена на рис. 3.7. Её основу составляет элементарный кубик, изображённый на рис. 6.1. Векторы a_1 , a_2 , a_3 называются *векторами трансляции*. Удлиняя каждый из векторов трансляции в целое число раз (n_1, n_2, n_3) , а затем складывая получившиеся векторы, можно получить вектор R_{n_1, n_2, n_3} , выходящий из начала координат и упирающийся в любой узел простой кубической решётки:

$$R_{n_1, n_2, n_3} = n_1 a_1 + n_2 a_2 + n_3 a_3.$$

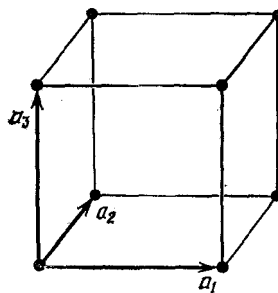


Рис. 6.1. Элементарный кубик простой кубической решётки

В кубические решётках все векторы a_1 , a_2 , a_3 имеют одинаковую длину, равную a , и потому числа n_1 , n_2 , n_3 являются просто

три декартовыми координатами узлов решётки, выраженными в единицах a . Говорят, что простая кубическая решётка получается трансляцией (параллельным переносом) элементарного кубика на векторы, кратные векторам a_1, a_2, a_3 .

Важнейшей характеристикой решётки является число ближайших соседей (его ещё называют *координационным числом*), которое обозначается через z . Для простой кубической решётки $z = 6$.

В форме простой кубической решётки кристаллизуются щёлочно-галлоидные соединения, такие как NaCl (поваренная соль), KCl (каменная соль), LiF , NaI и др. При этом ионы щёлочного металла (например, Na^+) чередуются в узлах решётки с ионами галогена (например Cl^-).

Объёмноцентрированной кубической решётки (сокращённо ОЦК). Эту решётку можно получить из двух простых кубических решёток, если сместить их относительно друг друга на половину диагонали элементарного кубика.

(Диагональю куба называют линию, соединяющую противоположные углы куба и проходящую через центр куба.)

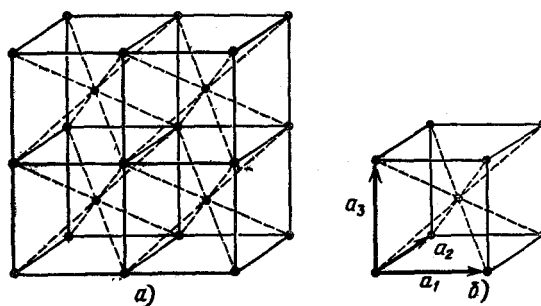


Рис. 6.2. а) Объёмноцентрированная кубическая решётка; б) элементарный кубик ОЦК-решётки.

ОЦК-решётка изображена на рис. 6.2, а. Её основу составляет элементарный кубик, показанный на рис. 6.2, б. ОЦК-решётку можно получить трансляцией этого кубика на векторы a_1 , a_2 , a_3 , однако в этом случае, для того чтобы перебрать все узлы решётки, недостаточно поместить вектор — R_{n_1, n_2, n_3} в начале координат. Надо, кроме того, поместить его начало на центральном атоме. Только тогда будут транслироваться обе простые кубические решётки, составляющие ОЦК-решётку.

Ближайшие соседи каждого узла ОЦК-решётки расположены в направлении диагонали элементарного кубика (рис. 6.2, б). Таких соседей у каждого узла 8, так что $z = 8$. Расстояние до ближайшего соседа равно половине диагонали куба, т. е. $\frac{\sqrt{3}a}{2}$, где буквой a по-прежнему обозначена длина ребра элементарного кубика.

В форме ОЦК-решётки кристаллизуются одновалентные щелочные металлы **Li**, **Na**, **K**, **Rb**, **Cs**, двухвалентный **Ba** и ряд других веществ.

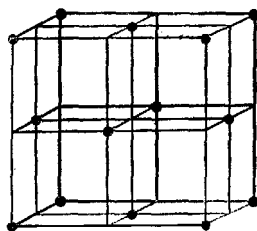


Рис. 6.3. Элементарный кубик гранецентрированной кубической решётки

Гранецентрированная кубическая решётка (сокращённо ГЦК). Элементарный кубик ГЦК-решётки показан на рис. 6.3. От элементарного кубика для ПК-решётки он отличается дополнительными узлами, находящимися в центре каждой грани. Чтобы транс-

лизовать этот кубик, нужно поместить начало вектора R_{n_1, n_2, n_3} , в начало координат, а также в центры трёх непровотивоположных друг другу граней куба. Ближайшие соседи каждого узла ГЦК-решётки расположены в направлении диагоналей грани куба. Расстояние до ближайшего соседа равно $\frac{a}{\sqrt{2}}$. В каждой из трёх взаимно перпендикулярных плоскостей, пересекающихся в данном узле, имеется 4 ближайших соседа этого узла, так что число ближайших соседей z равно 12.

В форме ГЦК-решётки кристаллизуются такие металлы, как медь, серебро, золото, алюминий, свинец.

Последней решёткой, которую мы введём в рассмотрение, является *решётка типа алмаза*. Она изображена на рис. 6.4, а. Её можно представить в виде двух гранецентрированных решёток, сдвинутых относительно друг друга на четверть диагонали куба.

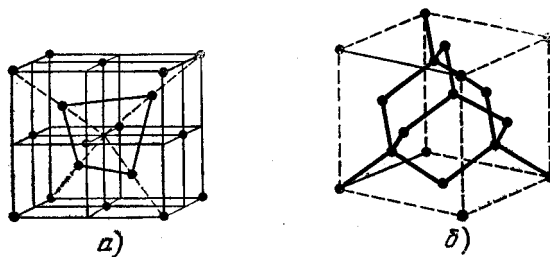


Рис. 6.4. а) Решётка типа алмаза; б) тетраэдрическая структура связей в решётке типа алмаза.

В форме решётки типа алмаза кристаллизуются элементы четвертой группы таблицы Менделеева: углерод (алмаз), а также важнейшие полупроводниковые элементы германий и кремний. Все эти элементы четырёхвалентны, и атомы связаны в решётке ковалентными силами. Упрощённо можно представить, что каждый атом имеет четыре «руки», соответствующие четырём валент-

ным электронам. Этими четырьмя руками атом берёт за руку своих четырёх ближайших соседей. Решётка типа алмаза прекрасно приспособлена для связи такого типа. Каждый узел этой решётки находится в центре правильного тетраэдра, образованного другими узлами (рис. 6.4, б). Число ближайших соседей $z = 4$.

6.2. Пороги протекания для объёмных решёток

Задачи связей и узлов ставятся для объёмных решёток точно так же, как для плоских. По-прежнему предполагается, что связи имеются только между узлами, являющимися ближайшими соседями.

В табл. 6.1 дана сводка порогов протекания задач узлов и связей для описанных выше объёмных решёток. Как уже говорилось, в трёхмерном случае не существует ни одного точного решения. Все результаты, приведённые в табл. 6.1, получены различными приближёнными методами, как правило, использующими ЭВМ. Естественно, что среди имеющихся в научной литературе результатов возникают небольшие противоречия. В табл. 6.1 отображены наиболее достоверные с нашей точки зрения результаты.

Таблица 6.1. Пороги протекания для объёмных решёток

Тип решётки	$x_{св}$	x_y
Простая кубическая	0,25	0,31
Объёмноцентрированная кубическая	0,18	0,25
Гранецентрированная кубическая	0,12	0,20
Типа алмаза	0,39	0,43

Теперь задача состоит в том, чтобы постараться понять, глядя на эту таблицу, а также на табл. 5.1 (с. 112), где даны результаты для плоских решёток, почему для одних решёток пороги протекания сравнительно большие, а для других — маленькие. Начнем с задачи связей.

6.3. От чего зависит порог протекания задачи связей?

Если все связи целые, то каждый узел связан с другими узлами, где число ближайших соседей z сильно меняется от решётки к решётке. При заданной доле целых связей x каждый узел в среднем связан с zx другими узлами. Попробуем проверить следующую гипотезу: может величина zx , представляющая среднее число узлов, с которыми связан каждый узел, содержать информацию, достаточную для того, чтобы сказать, есть в решётке протекание или нет? Может быть, никакой другой информации о свойствах решётки, кроме числа z , и не надо, и протекание возникает у всех решёток при одном и том же значении величины zx ? Довольно ясно, что эта гипотеза не может быть точной. Но может быть она справедлива приближённо?

Проверить это очень просто. Нужно для всех решёток с известными порогами протекания задачи связей вычислить произведение zx_{cv} . Если оно окажется универсальным, т. е. одинаковым для всех решёток или хотя бы приближённо одинаковым, значит, высказанная гипотеза верна или верна приближённо.

Соответствующие данные собраны в табл. 6.2. Видно, что с погрешностью меньше чем 10%, для плоских решёток справедлива

формула

$$zx_{св} = 2 \quad (6.1)$$

а для объёмных решёток формула

$$zx_{св} = 1,5. \quad (6.2)$$

Таблица 6.2. Произведение $zx_{св}$ для разных решёток

Тип решётки	z	$x_{св}$	$zx_{св}$
Плоские решётки			
Квадратная	4	0,5	2,0
Треугольная	6	0,35	2,1
Шестиугольная	3	0,65	2,0
Объёмные решётки			
Простая кубическая	6	0,25	1,5
Объёмноцентрированная кубическая	8	0,18	1,4
Гранецентрированная кубическая	12	0,12	1,4
Типа алмаза	4	0,39	1,6

Таким образом, гипотеза об универсальности среднего числа связей на узел, требуемого для возникновения протекания, не является точной, но приближенно выполняется. Если принять во внимание, что как в группе плоских решёток, так и в группе объёмных решёток каждая из величин z и $x_{св}$ меняется по крайней мере в два раза, то точность, с которой в каждой группе величина $zx_{св}$ постоянна, следует признать высокой.

Итак, чтобы приближенно оценить порог протекания задачи связей, достаточно знать число ближайших соседей и воспользоваться формулой (6.1) в случае плоских решёток и формулой (6.2)

в случае объёмных решёток. Порог протекания задачи связей наиболее чувствителен к числу ближайших соседей и значительно менее чувствителен ко всем прочим свойствам решёток (например, к числу вторых соседей, т. е. соседей, следующих по удаленности от данного узла).

Таким образом, мы получили очень простой и в то же время довольно точный способ оценки порогов протекания задачи связей, пригодный для любой решётки.

6.4. Как оценить порог протекания задачи узлов?

Разберём теперь схему такого же типа для задачи узлов. Естественно сначала испробовать предыдущий вариант, т. е. посмотреть, как меняется от решётки к решётке величина zx_y . Легко убедиться, что она меняется почти так же, как каждая из величин z и x_y по отдельности. Этому не следует удивляться; в случае задачи связей произведение $zx_{св}$ имеет четкий физический смысл — среднее число целых связей, приходящееся на один узел. В случае задачи узлов связь работает, если она соединяет два белых узла и не работает во всех прочих случаях. Поэтому произведение zx_y никакого особого смысла не имеет.

В 1970 г. американские физики Шер и Заллен предложили иной метод оценки порога протекания задачи узлов. Их идея состояла в том, чтобы сопоставить каждому узлу определённую долю пространства. После этого говорилось, что протекание по белым узлам возникает, когда доля пространства, занятая этими узлами, превышает некоторое критическое значение, слабозависящее от

типа решётки.

Вообразим вокруг каждого узла решётки шар (или круг в случае плоской решётки) с радиусом, равным половине расстояния до ближайшего соседа. При этом шары (круги), построенные вокруг соседних узлов, касаются друг друга (рис. 6.5). Белому узлу припишем белый шар, а чёрному узлу — чёрный. Если два белых узла связаны друг с другом, то между ними есть путь по касающимся друг друга белым шарам (рис. 6.5). Поэтому возникновение протекания означает появление путей бесконечной длины по касающимся друг друга белым шарам.

Предположим теперь, что протекание возникает, когда доля полного объёма (площади), занимаемая белыми шарами (в плоском случае кругами), превысит некоторое критическое значение, не зависящее от типа решётки. Чтобы проверить это предположение, нужно вычислить доли объёма, занимаемые белыми шарами при $x = x_y$, для различных решёток с известным значением x_y и сравнить их друг с другом.

Сначала необходимо сосчитать долю объёма, занимаемого белыми шарами при $x = 1$, т. е. в случае, когда все шары - белые. Эту величину обозначают буквой f и называют коэффициентом заполнения. Коэффициент заполнения равен доле объёма, занятой шарами, построенными вокруг каждого узла решётки и имеющими радиус, равный половине расстояния до ближайшего соседа. Коэффициент заполнения существенно зависит от типа решётки, и для каждой решётки его нужно считать отдельно.

Чтобы узнать долю объёма, заполненного белыми шарами при $x < 1$, нужно умножить коэффициент заполнения на долю белых

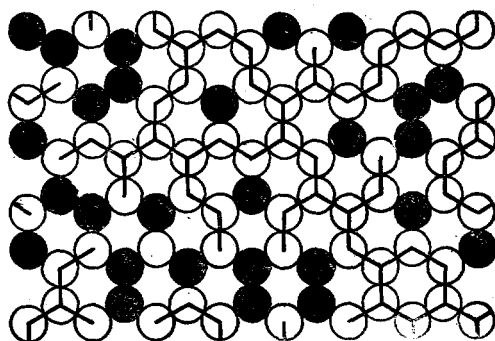


Рис. 6.5. Построение касающихся друг друга окружностей в случае шестиугольной решётки. Сама решётка изображена на рис. 5.1, в. Окружности имеют радиус, равный половине расстояния до ближайшего соседа. Белым узлам соответствует белая, а чёрным — чёрная окружности. Пути протекания по белым окружностям показаны жирными линиями.

шаров, т. е. на x . Таким образом, доля объёма, заполненного белыми шарами, равна fx . На пороге протекания она равна fx_y . Если предположение о универсальности доли объёма, при которой возникает протекание, правильно, то величина fx_y должна быть одинаковой для всех решёток.

Коэффициенты заполнения для различных решёток даны во втором столбце табл. 6.3. Чтобы представить, как они получены, вычислим величину для шестиугольной решётки, изображённой на рис. 6.5. В упражнениях 4 и 5 было показано, что на долю одного узла шестиугольной решётки приходится площадь $\frac{3\sqrt{3}}{4}a^2$, где a — сторона шестиугольника. Этот результат имеет следующий смысл: нарисуем на плоскости, на которую нанесена решётка, квадрат, прямоугольник, круг или любую другую геометрическую фигуру, но обязательно такую, чтобы её размеры были во много раз больше расстояния между соседними узлами решётки. Поделим её площадь на число узлов решётки, оказавшихся внутри этой

фигуры. Площадь, приходящаяся на узел, есть предел этого отношения при неограниченном увеличении размеров фигуры.

Таблица 6.3. Произведения fx_y для разных решёток

Тип решётки	f	x_y	fx_y
Плоские решётки			
Квадратная	0,79	0,59	0,47
Треугольная	0,91	0,5	0,46
Шестиугольная	0,61	0,7	0,43
Объёмные решётки			
Простая кубическая	0,52	0,31	0,16
Объёмноцентрированная кубическая	0,68	0,25	0,17
Гранецентрированная кубическая	0,74	0,20	0,15
Типа алмаза	0,34	0,43	0,15

Доля площади, занятая кругами, равна пределу отношения площади, занятой кругами, к площади большой фигуры. Площадь, занятая кругами, равна произведению числа узлов, принадлежащих большой фигуре, на площадь одного круга. Иными словами, величина f равна отношению площади одного круга к площади, приходящейся на один узел.

Круги, построенные на рис. 6.5, имеют радиус $\frac{a}{2}$ и, следовательно, площадь $\frac{\pi a^2}{4}$. Отсюда получаем

$$f = \frac{\frac{\pi a^2}{4}}{\frac{3\sqrt{3}a^2}{4}} = \frac{\pi}{3\sqrt{3}} \approx 0,605.$$

Аналогично вычисляются коэффициенты заполнения для других решёток, причём, как видно из табл. 6.3, они меняются в широких

пределах.

Произведения $f x_y$ представлены в последнем столбце табл. 6.3. Видно, что предположение о том, что $f x_y$ не зависит от типа решётки, не выполняется точно. Однако и в группе плоских решёток, и в группе объёмных решёток, то произведение меняется мало. Отсюда следует, что с точностью порядка 10–15% справедливы формулы!

$$f x_y = 0,5 \quad (6.3)$$

для плоских решёток и

$$f x_y = 0,16 \quad (6.4)$$

для объёмных решёток.

Так как вычислить коэффициент заполнения f сравнительно просто, формулы (6.3) и (6.4) дают возможность оценить порог протекания задачи узлов для любой решётки.

Легко понять, что критическая доля объёма, заполненная белыми шарами, при которой возникает протекание, монотонно уменьшается с увеличением размерности пространства. В одномерном пространстве, т. е. в линейной цепочке узлов, протекание по белым узлам невозможно при сколь угодно малой концентрации чёрных узлов. Даже один чёрный узел запирает путь протекания, так как обойти его невозможно. В плоской (двумерной) решётке появляется возможность обхода чёрных узлов, а в трёхмерной (объёмной) решётке таких возможностей больше, так как обходные пути не ограничены плоскостью.

Идея критического объёма оказывается плодотворной не только для решёточных задач. В гл. 9 мы столкнёмся с задачей, в кото-

рой белые и чёрные шары вообще не находятся в узлах решётки, а просто беспорядочно насыпаны в банку. Нас будет интересовать вопрос о протекании по касающимся друг друга белым шарам. Оказывается, это протекание тоже возникает, когда объем, заполненный белыми шарами, составляет примерно 0,16 полного объёма. Этот результат слабо меняется, если шары отличаются друг от друга радиусом.

В гл. 10 рассмотрена задача о пространстве, которое раскрашено случайным образом белой и чёрной краской. Оказывается, что протекание по областям одного цвета возникает в плоском случае, когда доля поверхности, выкрашенной этим цветом, точно равна 0,5, а в трёхмерном случае, когда доля объёма, выкрашенного этим цветом, примерно равна 0,16.

Упражнение

1. Проверьте, правильно ли вычислены коэффициенты заполнения, приведенные в табл. 6.3.

Глава 7.

Ферромагнетик с дальним действием и задача сфер.

В задачах узлов и связей, которые рассматривались до сих пор, предполагалось, что каждый узел может быть непосредственно связан лишь со своими ближайшими соседями, а связи между далекими узлами устанавливаются по цепочке узлов, каждый из которых связан со своими ближайшими соседями. В этой главе задача узлов обобщается на случай, когда узлы, не являющиеся ближайшими соседями, непосредственно связаны друг с другом. Такая задача может оказаться практически важной, и потому полезно знать, что она достаточно хорошо разработана.

Если число узлов, с которыми связан данный узел, становится очень большим, то задача узлов превращается в совершенно новую задачу, называемую *задачей сфер*. Эта задача играет в теории протекания особенно важную роль. С её помощью пытаются понять переход к проводимости металлического типа, происходящий в полупроводниках по мере увеличения концентрации примесей. На решении этой задачи основана теория прыжковой проводимости полупроводников — важного и интересного явления, разы-

грывающегося при очень низких температурах. Поэтому задачей сфер занимались многие учёные, для неё и подобных ей задач получены интересные результаты.

Задача сфер интересна ещё и потому, что это первая *нерешёточная задача*, с которой мы сталкиваемся в этой книге. Случайные элементы, которые в ней фигурируют, заданы не на узлах периодической решётки.

7.1. Ферромагнетик с дальним действием

Вернёмся к задаче о ферромагнетике с немагнитными примесными атомами, рассматривавшейся в гл. 3. В этой главе считалось, что магнитные атомы ориентируют свои магнитные моменты параллельно, только если они являются ближайшими соседями или связаны цепочкой магнитных атомов, являющихся ближайшими соседями друг другу. Если же все ближайшие соседи магнитного атома оказывались немагнитными атомами, то магнитный момент такого атома считался ориентированным произвольно.

Такая модель основывалась на том, что взаимодействие между магнитными моментами, приводящее к параллельной ориентации, убывает с расстоянием очень быстро, — так быстро, что магнитные моменты, не являющиеся ближайшими соседями, «ничего не знают» друг о друге, т. е. не взаимодействуют между собой.

Кристаллографы обычно называют группу атомов, являющихся ближайшими соседями некоторого атома, *первой координационной группой*, а число ближайших соседей z , как уже говорилось, называют координационным числом. Совокупность равноправных атомов, являющихся вторыми по степени удалённости

соседями данного атома, называют второй координационной группой и т. д. Рассмотрим в качестве примера простую кубическую решётку (см. рис. 3.7). Первую координационную группу в этой решётке составляют 6 атомов, расположенных на рёбрах куба, выходящих из исходного атома. Вторую координационную группу образуют 12 атомов, лежащих на диагоналях граней куба, проходящих через исходный атом. Наконец, третью координационную группу составляют 8 атомов, лежащих на диагоналях куба, проходящих через исходный атом.

В гл. 3 предполагалось, что взаимодействие между магнитными моментами распространяется только на первую координационную группу. Расчёт критической доли магнитных атомов, при которой возникала (или исчезала) спонтанная намагниченность, сводился в этом предположении к задаче узлов со связями только между ближайшими соседями.

Предположение «о коротком взаимодействии» (или о короткодействии) не всегда оправдывается, и потому имеет смысл рассмотреть задачу, в которой считается, что взаимодействие между магнитными моментами распространяется на несколько координационных групп, и найти критическую долю магнитных атомов, при которой возникает спонтанная намагниченность.

Эта проблема сводится к задаче узлов, в которой связаны не только ближайшие соседи. Её формулировка не содержит по существу ничего нового. Узлы бывают белые и чёрные (магнитные и немагнитные). Два белых узла считаются связанными друг с другом, если связи распространяются на ту координационную группу, в которой находится один из них по отношению к другому. Если

узел A связан с узлом B , а узел B связан с узлом C , то A связан с C . Совокупность связанных узлов образует кластер. Порогом протекания называется доля белых узлов, при которой возникает бесконечный кластер.

Совершенно очевидно, что порог протекания x_y должен уменьшаться по мере того, как связи распространяются на более далёкие координационные группы. Чем больше связей выходит из данного белого узла, тем больше вероятность того, что хотя бы одна из них приведёт к другому белому узлу, и, соответственно, тем меньше нужно белых узлов, чтобы обеспечить протекание.

Разумеется, решить такую задачу ничуть не проще (скорее, сложнее!), чем обычную задачу узлов. Однако целый ряд задач этого рода был решён различными приближёнными методами, и в табл. 7.1 приводятся результаты одного из них. В первой графе таблицы написан тип решётки и номера координационных групп, на которые распространялись связи. Во второй графе стоит число узлов Z , с которыми связан каждый узел, т. е. суммарное число узлов, находящееся в рассматриваемых координационных группах. (В случае одной координационной группы оно совпадает с координационным числом z).

В последней графе таблицы написано произведение Z_{x_y} . Как говорилось в предыдущем разделе, в случае задачи узлов со связями в первой координационной группе это произведение сильно зависит от типа решётки. Однако, как видно из табл. 7.1, при больших значениях Z оно меняется с увеличением Z все слабее и слабее. Особенно хорошо это видно в случае объёмных решёток, где использованы большие значения Z . Произведение Z_{x_y} явно стремится

ся к числу порядка $2,6-2,7$, не зависящему от типа решётки.

В трёхмерном случае число $2,7$ и считается в настоящее время наиболее достоверным значением (с точностью $\pm 0,1$) величины B_c , определенной как предел Z_{xy} при больших Z .

$$B_c = \lim_{Z \rightarrow \infty} Z_{xy}. \quad (7.1)$$

Из следующих разделов станет ясно, почему такой предел существует и почему он не зависит от типа решётки, а зависит лишь от размерности пространства, т. е. от того, является рассматриваемая решётка плоской или объёмной.

Чтобы это понять, нужно получить представление о задаче сфер.

Упражнение

1. Укажите, где находятся 42 узла, принадлежащие трем первым координационным группам ГЦК-решётки.

7.2. Задача окружностей (сфер)

Рассмотрим теперь иную задачу или, говоря точнее, задачу, которая на первый взгляд кажется иной. Допустим, что на плоскости нарисованы окружности с одинаковым радиусом, равным R , центры которых распределены по плоскости совершенно хаотически и в среднем равномерно. Это значит, что обе координаты центров окружностей являются случайными числами, равномерно распределёнными в интервале от нуля до L , где L — очень большая

Таблица 7.1. Пороги протекания задачи узлов со связями между удалёнными соседями (Для случая одной координационной группы приведённые значения порогов не всегда точно совпадают с данными предшествующих таблиц. Это связано с тем, что в них были отобраны наиболее достоверные, с нашей точки зрения, данные. В этой же таблице сравниваются данные, полученные одним и тем же методом.)

Тип решётки	Z	x_y	Z_{x_y}
Плоские решётки			
Шестиугольная, 1	3	0,700	2,10
Квадратная, 1	4	0,590	2,36
Треугольная, 1	6	0,500	3,00
Квадратная, 1, 2	8	0,410	3,28
Треугольная, 1, 2	12	0,295	3,54
Шестиугольная, 1, 2, 3	12	0,300	3,60
Квадратная, 1, 2, 3	12	0,292	3,50
Треугольная, 1, 2, 3	18	0,225	4,05
Объёмные решётки			
Типа алмаза	4	0,425	1,70
ПК 1	6	0,307	1,84
ОЦК 1	8	0,253	1,94
ГЦК 1	12	0,195	2,34
ОЦК 1, 2	14	0,175	2,45
ЦК 1, 2	18	0,137	2,47
ГЦК 1, 2	18	0,136	2,45
ПК 1, 2, 3	26	0,097	2,52
ОЦК 1, 2, 3	26	0,095	2,47
ГЦК 1, 2, 3	42	0,061	2,56

длина (по сравнению с R), характеризующая размер рассматриваемой системы. Важная отличительная черта этой задачи состоит в том, что окружности могут сколь угодно перекрываться друг с другом. Среднее число центров окружностей, приходящееся на

единицу площади, равно N . Иными словами, N — концентрация центров окружностей.

Две окружности считаются связанными друг с другом, если центр одной из них находится внутри другой, — такие окружности иногда называют охватывающими. Если окружность A связана с окружностью B , а B связана с C , то A связана с C . Таким образом, далекие друг от друга окружности могут быть связаны по цепочке охватывающих окружностей (рис. 7.1).

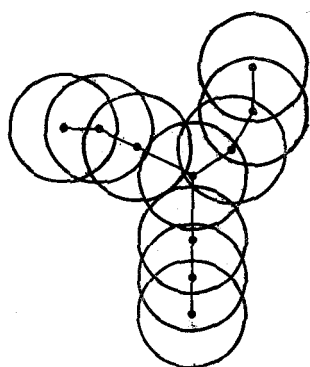


Рис. 7.1. Пути протекания по охватывающим окружностям показаны ломаными линиями. Точки изображают центры окружностей.

Задача состоит в том, чтобы найти критическое значение концентрации N , при котором возникает протекание по охватывающим окружностям, т. е. возникают пути, проходящие через всю систему и состоящие из охватывающих окружностей. (Иными словами, возникает бесконечный кластер связанных друг с другом окружностей.)

В задачу введено два параметра — концентрация N и радиус R . (Есть ещё размер системы L , но ясно, что если система достаточно

велика, критическое значение N слабо зависит от L .) Между тем, легко убедиться в том, что отсутствие или наличие протекания зависит не от двух параметров, а лишь от одного, представляющего собой безразмерное произведение NR^2 . (В случае плоской задачи размерность концентрации есть см^{-2} .) В качестве этого параметра удобно выбрать среднее число центров окружностей, находящееся внутри одной окружности. Оно равно

$$B = \pi NR^2.$$

В том, что протекание возникает при некотором значении параметра B и не зависит от того, чему при этом равны величины N и R по отдельности, легче всего убедиться следующим образом. Допустим, что дана плоскость с нарисованными на ней окружностями. Увеличим данную картину в несколько раз, например, с помощью проектора. Это и будет преобразование, меняющее N и R , но не меняющее B , поскольку среднее число центров окружностей, находящееся внутри окружности, от увеличения не меняется.

Легко также понять, что это преобразование не влияет на протекание. Если протекание отсутствовало на исходной картине, то оно отсутствует и на увеличенной, и, наоборот, если на исходной картине было протекание по охватывающим окружностям, то от увеличения оно не исчезнет.

Итак, преобразование, меняющее N и R , но не изменяющее B , не влияет на протекание. Таким образом, отсутствие или наличие протекания в системе зависит лишь от величины параметра B . При больших значениях этого параметра протекание есть, а при малых значениях его нет.

Эта новая задача, которая была только что сформулирована, называется *задачей окружностей*. Трёхмерный её аналог называется *задачей сфер*. Задача сфер формулируется следующим образом. В трёхмерном пространстве с помощью генератора случайных чисел задаются координаты центров сфер, имеющих радиус R . Две сферы называются связанными друг с другом (или охватывающими, если центр одной сферы находится внутри другой сферы). Требуется определить критическую концентрацию центров, при которой возникает протекание по охватывающим сферам.

Легко понять, что, как и в плоском случае, наличие протекания определяется только значением параметра, представляющего собой среднее число центров сфер, находящихся внутри одной сферы:

$$B = \frac{4\pi}{3}NR^3$$

где N — среднее число центров сфер в единице объема (объёмная концентрация имеет размерность см^{-3}).

Как уже говорилось, задача сфер имеет важное значение для теории электропроводности полупроводников при низких температурах. Поэтому она исследовалась многими авторами с помощью самых различных методов. По современным данным критическое значение B_c , при котором возникает протекание по сферам, равно $2,7 \pm 0,1$. Задача окружностей исследовалась менее интенсивно, и в результатах, приводимых разными авторами, имеется значительный разброс. Вероятно, значение $B_c = 4,1 \pm 0,4$.

7.3. Задача окружностей (сфер) — предельный случай задачи узлов

Вернёмся к задаче узлов, в которой связаны друг с другом не только ближайшие соседи, и объясним, почему существует предел, стоящий в правой части формулы (7.1), почему он не зависит от типа решетки и, наконец, почему он обозначен той же буквой B_c , что и пороговые значения задач окружностей и сфер.

Начнём с того, что несколько иначе сформулируем задачу узлов. Для конкретности будем говорить сначала о плоских решётках. Обобщение на объёмные решётки оказывается очень простым.

Построим вокруг каждого белого узла окружность с радиусом R , выбранным так, что он больше, чем расстояние от этого узла до узлов последней координационной группы, с которой данный узел связан, но меньше, чем расстояние до узлов следующей координационной группы. Примем, что окружности связаны друг с другом, если связаны белые узлы, вокруг которых они построены. Это означает, что окружности связаны, если центр одной из них находится внутри другой, т. е. если окружности являются охватывающими.

Появление бесконечного кластера из связанных белых узлов эквивалентно возникновению протекания по охватывающим окружностям. Иными словами при критической доле белых узлов x_y появляется протекание по окружностям радиуса R , построенным вокруг белых узлов (рис. 7.2).

Легко понять смысл величины Z_x . Величина Z представляет

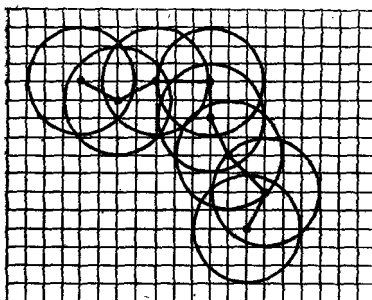


Рис. 7.2. Путь протекания по охватывающим окружностям, построенным на квадратной решётке. Взаимодействие учитывается на расстоянии, втрое большем, чем расстояние между ближайшими соседями. Путь протекания показан ломаной линией.

собой число узлов (и чёрных, и белых), находящихся внутри одной окружности. Величина Z_x есть среднее число лежащих внутри окружности центров других окружностей (или среднее число белых узлов внутри одной окружности). Величина Z_{xy} есть среднее число центров окружностей, лежащих внутри одной окружности, при котором возникает протекание, т. е. возникают бесконечные пути по охватывающим окружностям.

Отсюда видно, что произведение Z_{xy} имеет тот же смысл, что и величина B_c в задаче окружностей. Теперь, наверное, труднее понять, чем отличается задача узлов от задачи окружностей, чем понять, почему они похожи. Однако разница есть, и она очень существенна. Дело в том, что в задаче окружностей центрами окружностей могут быть любые точки на плоскости, а в задаче узлов ими являются лишь узлы рассматриваемой решётки (рис. 7.2). Если число узлов, находящихся внутри одной окружности, невели-

ко, то отличие между двумя задачами оказывается очень важным. Естественно, что при таких условиях критическое значение Z_{xy} зависит от типа решётки. Полное число узлов, находящихся внутри одной окружности, равно Z . Как видно из табл. 7.1, при не очень больших Z значения Z_{xy} для разных решёток, действительно, отличаются.

Однако, если число Z велико, то разница между двумя задачами стирается. Допустим, что мы начали с задачи окружностей, а затем переместили центр каждой окружности до ближайшего к нему узла решётки. Это уже будет задача узлов. Если узлов решётки внутри каждой окружности очень много, то с подавляющей вероятностью такой сдвиг не приведет к тому, что несвязанные окружности станут связанными, или наоборот. Отсюда видно, что при $Z \rightarrow \infty$ задача узлов и задача окружностей становятся эквивалентными. Значение B_c , определяемое формулой (7.1), не зависит от типа решётки и совпадает с величиной B_c , определенной в задаче окружностей.

Эти рассуждения полностью переносятся на трёхмерный случай. Для объёмных решёток величина B_c , определённая формулой (7.1), совпадает с величиной B_c , определённой в задаче сфер.

Итак, при большом значении Z задача узлов на любой плоской решётке сводится к задаче окружностей, а задача узлов на объёмной решётке сводится к задаче сфер. Таким образом, предел в формуле (7.1), действительно, не зависит от типа решётки, но зависит от размерности пространства, в котором решётка задана. (Величины B_c для окружностей и сфер разные!)

Глава 8.

Электропроводность примесных полупроводников и задача сфер

Теория протекания оказалась чрезвычайно полезной для понимания процессов, происходящих в полупроводниках с примесями (их называют *примесными* полупроводниками в отличие от чистых полупроводников, которые называют *собственными*). Можно, пожалуй, сказать, что в настоящее время примесные полупроводники являются одной из главных областей применения теории протекания. О примесных полупроводниках речь идет в нескольких главах книги. Настоящую главу мы начинаем с обсуждения современных представлений о них, прервав на время изложение теории протекания.

8.1. Собственные полупроводники

Поговорим сначала о чистых полупроводниках, взяв за основу полупроводниковые элементы четвертой группы таблицы Менделеева, такие как германий и кремний. Эти элементы кристаллизуются в решётку типа алмаза (рис. 6.4, а, б). На внешней оболочке ка-

ждого атома имеется четыре электрона, которые образуют связи с четырьмя ближайшими соседями. Характер движения внешних электронов таков, что их плотность не распределена равномерно вокруг атома, а сосредоточена в четырёх жгутах, направленных из центра тетраэдра, в котором находится каждый атом, к углам этого тетраэдра. Эти жгуты и стягивают атомы кристалла, не позволяя ему распадаться.

Можно представить себе, что каждый атом имеет четыре руки, которыми он берёт за руки своих четырёх соседей. При этом каждый электрон крепко связан со своими соседями, и если к полупроводнику приложить не очень сильное электрическое поле, то оно не вызовет электрический ток: электронов ровно столько, сколько требуется для образования связей, а для того чтобы вырвать электрон, осуществляющий связь, нужна довольно большая энергия. Эта энергия, точнее, минимальная работа, которая требуется, чтобы перевести электрон из связанного состояния в состояние, в котором он может свободно двигаться по кристаллу, называется *шириной энергетической щели* (или просто шириной щели) и обозначается через E_g .

Ширина энергетической щели — важнейшая характеристика полупроводника, в значительной степени определяющая все его электрические свойства. Рассмотрим, например, электропроводность. В чистом (собственном) полупроводнике электропроводность осуществляется за счёт того, что некоторое количество связей разорвано. При разрыве связи возникает электрон, способный проводить электрический ток и «покалеченный», «трёхрукий», атом. Этот «калека» также является носителем тока, только носи-

телем другого знака, чем электрон. Действительно, под действием электрического поля электрон соседнего атома может перейти на «трёхрукий», почти не изменив своей энергии. В результате «трёхруким» станет другой атом. Этот процесс может продолжаться, причём легко понять, что если электроны под действием поля движутся слева направо, то «трёхрукий» в результате этого же процесса движется справа налево. (Важно понимать, что сами атомы неподвижны. Просто за счёт движения электронов «трёхруким» становится то один, то другой атом) Поэтому «трёхрукий» атом также можно считать носителем тока, причём, в отличие от электрона, ему следует приписывать не отрицательный, а положительный заряд.

В полупроводниках, состоящих из атомов другой валентности, микроскопическая картина несколько иная, но во всех случаях при разрыве связи возникают два носителя тока, имеющих разные знаки. Один из них (отрицательный) называют электроном, а второй (положительный) — *дыркой*. Таким образом, «трёхрукий» атом является частным случаем дырки.

Энергия, необходимая для разрыва связей, черпается из энергии теплового движения атомов. В статистической физике доказывается, что при высоких температурах на один атом, совершающий малые колебания вблизи положения равновесия, приходится в среднем энергия теплового движения, равная $3kT$, где T — температура в градусах Кельвина, а k — постоянная Больцмана ($k = 1,38 \cdot 10^{-16}$ эрг/К).

Ширину щели E_g , как правило, измеряют в электрон-вольтах (эВ). Один электрон-вольт — это работа, которая совершается од-

ним электроном, когда он проходит участок пути с разностью потенциалов в 1 В: $1 \text{ эВ} = 1,6 \cdot 10^{-19} \text{ Дж} = 1,6 \cdot 10^{-12} \text{ эрг}$. В германии $E_g = 0,7 \text{ эВ}$, а в кремнии $E_g = 1,1 \text{ эВ}$. При комнатной температуре (300 К) энергия $3kT$ составляет всего лишь 0,08 эВ. Это значительно меньше, чем требуется для разрыва связи и в германии, и в кремнии.

Однако тепловое движение носит хаотический характер и в отдельные моменты времени энергия движения атома может оказаться очень большой. С каждым отдельным атомом это происходит крайне редко, но зато атомов очень много. Их число в 1 см^3 примерно составляет 10^{22} . (Напомним, что число атомов в одном моле равно постоянной Авогадро $6 \cdot 10^{23}$.)

Поэтому концентрация электронов и дырок, появившихся за счёт разрыва связей, гораздо меньше, чем концентрация атомов, но все же настолько велика, что в определенной области температур может обеспечить заметную электропроводность. Расчёт показывает, что в германии при $T = 300 \text{ К}$ концентрация электронов порядка 10^{13} см^{-3} . В кремнии, где щель шире, при той же температуре концентрация гораздо меньше (10^{10} см^{-3}). С понижением температуры концентрация электронов резко падает и соответственно резко растёт сопротивление собственного полупроводника.

Чем больше ширина энергетической щели, тем меньше носителей тока при заданной температуре. Вся разница между полупроводником и диэлектриком состоит лишь в ширине щели. Вещества с E_g порядка 5 эВ и больше, как правило, называют диэлектриками. При комнатной температуре в них практически нет носителей

тока.

Если разница между диэлектриком и полупроводником скорее количественная, чем качественная, то разница между металлом и диэлектриком носит принципиальный характер. В металле ширина энергетической щели равна нулю, а концентрация носителей тока велика даже вблизи абсолютного нуля температур.

8.2. Примесные полупроводники

Допустим, что в германий или кремний введены в качестве примесей элементы пятой группы таблицы Менделеева, такие как фосфор, сурьма, мышьяк. Атомы этих элементов имеют пять внешних электронов. Если такой атом замещает, например, атом германия, находящийся в центре тетраэдра, то четыре его электрона идут на образование связей с четырьмя соседями, а пятый электрон остается лишним. Этот электрон связан со своим атомом, потому что если электрон уйдет, то атом останется положительно заряженным и будет притягивать свой электрон обратно. Таким образом, для того чтобы удалить электрон на очень большое расстояние от атома, нужно совершить работу против сил притяжения. Эта работа и называется *энергией* связи электрона с атомом.

Как показано чуть ниже, энергия связи лишнего электрона относительно мала, так что этот электрон находится довольно далеко от примесного атома (по сравнению с периодом кристаллической решётки). Поэтому вся структура примесного атома напоминает простейший из атомов атом водорода. Напомним, что атом водорода состоит из положительно заряженного тяжёлого ядра и

лёгкого отрицательного электрона, причём размеры ядра ничтожно малы по сравнению с расстоянием от ядра до электрона.

Аналогичную структуру имеет и примесный атом. При этом роль водородного ядра играет не только ядро примесного атома, но и его внутренние электроны, и четыре электрона, образовавшие связи. Размер всего этого образования, как будет сейчас видно, мал по сравнению с расстоянием до лишнего электрона, а заряд образования положителен и по абсолютной величине равен заряду электрона.

Как известно, в атоме водорода электрон находится от ядра на расстоянии порядка так называемого боровского радиуса a_B , который выражается формулой

$$a_B = \frac{\hbar^2}{me^2} \approx 0,53 \cdot 10^{-8} \text{ см.} \quad (8.1)$$

Здесь $\hbar = 1,05 \cdot 10^{-27}$ эрг · с — постоянная Планка (деленная на 2π), $m = 9,8 \cdot 10^{-28}$ г — масса электрона, $e = 4,8 \cdot 10^{-10}$ ед. СГСЭ — заряд электрона.

Фундаментальная константа \hbar была введена немецким физиком М. Планком в 1901 г. в связи с гипотезой о квантовом характере электромагнитных волн. Эта константа входит в квантовомеханическое уравнение, описывающее движение электрона около ядра. Кроме неё в уравнение входят только константы e и m . Боровский радиус есть единственная величина, имеющая размерность длины, которую можно составить из e , m и \hbar . Энергия связи электрона в атоме водорода равна

$$E_B = \frac{me^4}{2\hbar^2} \approx 13,6 \text{ эВ.} \quad (8.2)$$

По порядку величины эту формулу можно интерпретировать следующим образом. Положительное ядро создает на расстоянии r потенциал $\frac{e}{r}$. На расстоянии порядка a_B он равен $\frac{e^2}{a_B}$, а на бесконечно большом расстоянии — нулю. Поэтому работа, которую нужно затратить, чтобы перенести электрон, находящийся на расстоянии порядка a_B от ядра, на бесконечно большое расстояние, равна по порядку величины $\frac{e^2}{a_B}$ (разность потенциалов, помноженная на заряд электрона).

Вернемся теперь к примесному атому. Формулы, написанные для атома водорода, необходимо модифицировать, чтобы учесть, что лишний электрон движется не в пустом пространстве, а в кристалле полупроводника. При этом меняется вид закона Кулона. Сила, действующая на электрон, находящийся на расстоянии r от ядра, равна теперь не $\frac{e^2}{r^2}$, а $\frac{e^2}{\varepsilon r^2}$ (ед. СГСЭ), где ε — диэлектрическая проницаемость кристалла. Это обстоятельство можно учесть, заменив в выражениях для a_B и E_B величину e^2 на $\frac{e^2}{\varepsilon}$. Кроме того, следует иметь в виду, что масса, описывающая движение электрона по кристаллу, не совпадает с массой свободного электрона m .

Дело в том, что в кристалле существует периодический электрический потенциал, создаваемый ядрами атомов полупроводника и внутренними электронами. Один из самых интересных выводов квантовой теории твердого тела состоит в том, что если этот потенциал точно периодический, то электрон почти не замечает его. Это «почти» состоит в том, что, выбрасывая периодический потенциал из уравнения движения электрона, нужно одновременно заменить массу электрона m на массу m^* , зависящую от свойств кристалла. Величина m^* называется *эффективной массой*.

Примесный атом создает не периодический потенциал, который никак нельзя выбросить. Но описывая движение электрона в этом потенциале, можно не учитывать периодический потенциал кристалла, заменив m на m^* .

Итак, уравнение, описывающее движение лишнего электрона в кристалле вокруг заряженного примесного атома, отличается от уравнения движения электрона в атоме водорода заменами $e^2 \rightarrow \frac{e^2}{\varepsilon}$ и $m \rightarrow m^*$

Обозначим характерное расстояние, на которое удален лишний электрон от примесного атома, через a_B^* , а энергию связи этого электрона — через E_B^* . Пользуясь формулами (8.1) и (8.2), получим

$$a_B^* = 0,53 \cdot 10^{-8} \left(\frac{m}{m^*} \right) \varepsilon \text{ [см]}, \quad (8.3)$$

$$E_B^* = 13,6 \left(\frac{m^*}{m} \right) \frac{1}{\varepsilon^2} \text{ [эВ]}. \quad (8.4)$$

Как правило, в полупроводниках эффективные массы значительно меньше, чем масса свободного электрона, а диэлектрическая проницаемость равна 10–15 (например, в германии $m^* \approx 0,1m$, а $\varepsilon = 16$). Благодаря этому характерное расстояние, на котором находится лишний электрон, составляет в типичных полупроводниках от нескольких десятков до сотен ангстрем ($1 \text{ \AA} = 10^{-8} \text{ см}$). что значительно больше, чем межатомное расстояние, (в германии оно равно $2,45 \text{ \AA}$).

Итак, мы получили, что в примесном атоме лишние электроны удалены от атома на значительное расстояние и удерживаются на

этом расстоянии силами притяжения к ядру. Уже исходя из того, что расстояние a_B^* велико, должно быть ясно, что энергия связи E_B^* мала. Действительно, из приведенных данных видно, что энергия связи типичных полупроводников составляет от нескольких сотых до нескольких тысячных электрон-вольта, т. е. она значительно меньше, чем ширина энергетической щели E_g (0,7 эВ для германия).

Это естественно, так как оторвать лишний электрон от примесного атома легче, чем оторвать электрон, осуществляющий связь между атомами. Тепловая энергия kT сравнивается с 0,01 эВ при температуре порядка 100 К. При такой температуре, как правило, значительная часть лишних электронов отрывается от своих примесных атомов и принимает участие в переносе электрического тока. Таким образом, примесные атомы пятой группы таблицы Менделеева сравнительно легко отдают свои лишние электроны. Поэтому они называются *донорами* (в переводе с английского слово **donor** означает «отдающий»).

Допустим теперь, что примесными атомами являются элементы третьей группы таблицы Менделеева, такие как бор, алюминий, галлий и т. д. У этих атомов на внешней оболочке оказываются три электрона, так что для образования связей с четырьмя соседями им *не хватает* одного электрона. Этот электрон легко заимствуется у соседних атомов полупроводника, но при этом один из соседних атомов становится «трёхруким» или, иными словами, в окрестности примесного атома появляется дырка. Захватив четвёртый электрон, примесный атом становится отрицательно заряженным. Поэтому дырка притягивается к этому атому электрическими си-

лами, и для того чтобы её оторвать, нужно совершить работу. Эта работа называется энергией связи дырки с примесным атомом.

Вычисление энергии связи дырки снова приводит к задаче о водородоподобном атоме, только теперь вокруг отрицательно заряженного ядра движется положительно заряженная дырка. Роль ядра играет примесный атом вместе с захваченным дополнительным электроном. Этот дополнительный электрон превращается в жгут, образующий связь между атомами, и потому область его движения не превышает межатомного расстояния. В то же время дырка связана с примесным атомом гораздо слабее. Расстояние от атома, на котором она находится, и энергия связи определяются формулами (8.3) и (8.4). Нужно только принимать во внимание, что эффективная масса дырки, которую следует использовать в этом случае, вообще говоря, отличается от эффективной массы электрона. Как правило, она также значительно меньше, чем масса свободного электрона, так что дырочный водородоподобный атом тоже имеет размеры порядка десятков ангстрем, а энергию связи порядка нескольких сотых электрон-вольта.

При температуре около 100 К тепловое движение отрывает дырки от примесных атомов, после чего дырки ведут «самостоятельный образ жизни» и, если приложить электрическое поле, принимают участие в переносе тока.

Таким образом, примесные атомы элементов третьей группы обладают способностью легко принимать электрон и образовывать дырку. Поэтому они называются *акцепторами* (в переводе с английского слово *acceptor* означает «принимающий»).

Подведём итоги сказанному. Подвижные носители тока образу-

ются в полупроводниках только за счёт энергии теплового движения. Они могут образовываться за счёт разрыва связей решётки. Для этого нужно совершить работу, равную ширине энергетической щели E_g . При этом одновременно появляются электрон и дырка. Кроме того, они могут рождаться по одиночке за счёт отрыва электрона или дырки от примесного атома. Так как энергия связи электрона и дырки с примесным атомом значительно меньше, чем ширина энергетической щели, то вероятность того, что электрон оторвётся от примесного атома, гораздо больше, чем вероятность, что разорвётся какая-либо связь. С другой стороны, концентрации примесных атомов, как правило, намного порядков ниже, чем концентрация основных атомов. Поэтому при повышении температуры сначала отрываются все носители, находящиеся на примесях, а концентрация собственных носителей (электронов и дырок, возникающих из-за обрыва связей) ничтожно мала. Эту температурную область называют областью примесной проводимости. Однако при дальнейшем повышении температуры концентрация собственных носителей сравнивается с концентрацией примесей и начинает её превосходить. Эта область называется областью собственной проводимости.

Самый важный для дальнейшего изложения вывод состоит в том, что при очень низкой температуре, когда энергия теплового движения kT мала по сравнению с энергией связи электронов на примесях, *в полупроводнике нет никаких носителей, ни примесных, ни собственных*. Связи находятся там, где они должны быть, а примесные электроны и дырки локализованы в окрестности своих атомов. Электропроводность полупроводника в этой

области очень резко обращается в нуль при понижении температуры. Этим полупроводник отличается от металла, в котором концентрация подвижных носителей остается большой при сколь угодно низкой температуре.

8.3. Переход к металлической электропроводности при повышении концентрации примесей

Это принципиальное отличие полупроводника от металла неожиданно исчезает при увеличении концентрации примесей. Если концентрация превышает некоторое критическое значение N_c , то при сколь угодно низких температурах электропроводность остается сравнительно большой и слабо зависит от температуры. Электропроводность с таким поведением называют *металлической*. Это совсем не значит, что по порядку величины электропроводность полупроводника сравнивается с электропроводностью хороших металлов. Электропроводность полупроводника всегда остается на много порядков ниже. Название отражает лишь характер поведения электропроводности при низких температурах. *Переход к металлической электропроводности, происходящий при увеличении концентрации примесей называют переходом диэлектрик-металл или переходом Мотта* (по имени известного английского физика Н. Мотта).

Эксперименты показали, что критическая концентрация примесей N_c , при которой происходит переход Мотта, довольно сильно изменяется от одного полупроводника к другому. Хорошую оцен-

ку величины N_c можно получить с помощью соотношения

$$N_c (a_B^*)^3 \approx 0,02. \quad (8.5)$$

Например, в германии $N_c = 10^{17}$, а в кремнии $N_c = 3 \cdot 10^{18} \text{ см}^{-3}$.

Удовлетворительной математической теории перехода Мотта в полупроводниках до сих пор не существует. Это одна из сложнейших задач физики твёрдого тела. Качественные соображения состоят в следующем.

По своей структуре примесные атомы похожи на одновалентные элементы первой группы таблицы Менделеева (H, Li, Na, K). Подобно этим элементам, примесный атом имеет один электрон на внешней оболочке. Кристаллизуясь, элементы первой группы образуют хорошие металлы. Водород образует молекулярный кристалл и является диэлектриком. Однако есть все основания думать, что при достаточно большом давлении он тоже становится металлом. (В настоящее время в этом направлении ведётся столь интенсивная работа, что она могла бы стать темой самостоятельной книги.)

Совершенно естественно, что водородоподобные примесные атомы, если они расположены внутри полупроводника достаточно плотно, тоже должны образовывать металлическую систему.

На первый взгляд кажется странным, что это может иметь отношение к переходу Мотта. Действительно, концентрации атомов в металлическом натрии порядка 10^{22} см^{-3} , что примерно в 100 тысяч раз больше концентрации примесей в германии, при которой происходит переход Мотта. Ясно, что при столь низких концентрациях натрий не может быть металлом.

Однако металлизация системы определяется не просто величиной концентрации, а степенью перекрытия электронных оболочек соседних атомов. Если атомы находятся на столь большом расстоянии, что области пространства, в которых движутся их внешние электроны, далеки друг от друга, то это просто изолированные атомы. Если же ядра атомов так близки друг к другу, что области движения их внешних электронов перекрываются (накладываются), то атомы теряют свою индивидуальность. Их внешние электроны «не могут знать», какому из ядер они принадлежат. В системе из большого числа атомов происходит обобществление, коллективизация электронов. Электроны образуют самостоятельную систему, способную проводить электрический ток. Такое вещество является металлом.

Для водородоподобных атомов величина перекрытия определяется безразмерным параметром $N (a_B^*)^3$. (Величина $\frac{4}{3}\pi N (a_B^*)^3$ представляет среднее число ядер атомов, помещающихся внутри сферы с радиусом, равным эффективному боровскому радиусу.) Как уже говорилось, благодаря тому, что примесный атом находится внутри полупроводника, его радиус аномально велик. По этой причине условие (8.5) выполняется уже при $N = 10^{17} \text{ см}^{-3}$ (для германия).

Из приведённых соображений должно быть ясно, почему переход Мотта во всех проводниках происходит примерно при одном и том же значении параметра $N (a_B^*)^3$, хотя критические концентрации N_c могут сильно меняться. Ведь именно этот параметр определяет перекрытие соседних атомов!

Можно поставить вопрос: какой концентрацией атомов должен

обладать кристалл водорода, чтобы перекрытие его атомов было таким же, как перекрытие примесных атомов, находящихся в полупроводнике, при концентрации, соответствующей переходу Мотта? Перекрытие определяется формулой (8.5), но теперь в неё естественно подставить боровский радиус, полученный из формулы (8.1). Тогда получится чудовищная концентрация, равная 10^{23} см^{-3} . Таким, образом, перекрытие, соответствующее переходу Мотта, является очень большим.

Поэтому нет оснований удивляться, что примесные атомы при концентрации, большей, чем N_c , образуют металлическую систему. Этот факт лежит в основе современной теории полупроводников с большой концентрацией примесей.

8.4. Переход Мотта и задача сфер

Теория протекания предлагает упрощённое описание перехода Мотта, учитывающее, что примеси в полупроводнике расположены хаотически и могут образовывать сгущения и разряжения. Допустим, что в некоторой области примесные атомы расположены так близко друг к другу, что их электронные оболочки сильно перекрываются и внешние электроны обобществляются. Такая область представляет собой кусок металла: если к ней приложить разность потенциалов, то потечёт электрический ток.

Однако существование таких областей недостаточно для того, чтобы массивный образец вёл себя как металл. Если металлические области встречаются редко, они не сообщаются друг с другом, а представляют собой лишь изолированные металлические

вкрапления в веществе, которое при низких температурах является диэлектриком. В целом такое вещество ведет себя как диэлектрик.

С повышением концентрации примесей доля пространства, занятая металлическими областями, увеличивается и при некоторой критической концентрации N_c металлические области образуют связанную систему «озёр и каналов», пронизывающую весь полупроводниковый кристалл. Начиная с этой концентрации электропроводность массивного образца приобретает металлический характер.

Ясно, что математическая формулировка изложенных выше представлений должна быть связана с теорией протекания, однако постановка задачи не проста. Главная трудность состоит в том, что нам не известно, при какой концентрации примесей данную область можно считать металлической.

Простейшая модель перехода Мотта, сформулированная в шестидесятых годах, состоит в следующем. Представим, что каждый примесный атом является металлическим шариком с некоторым радиусом r_0 . Шарик могут проникать друг в друга (рис. 8.1, а), т. е. занимаемые ими области пространства могут перекрываться. Шарик образуют цепочки (рис. 8.1, б) и области более сложной формы, которые по определению считаются металлическими. Нужно найти концентрацию шариков N_c , начиная с которой металлические области обеспечивают электропроводность массивного образца.

Из каких же соображений следует выбирать радиус r_0 ? Водородоподобный примесный атом не имеет четких границ. Вероят-

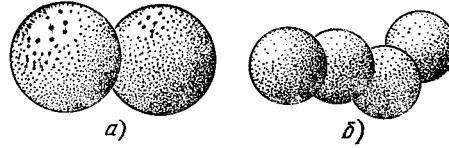


Рис. 8.1. а) Перекрывающиеся шары, изображающие атомы с обобществленными электронами; б) цепочка из перекрывающихся шаров — металлический канал, по которому может распространяться электрический ток.

ность найти электрон на расстоянии r от ядра убывает с увеличением r по закону $e^{\frac{-2r}{a_B^*}}$, где $e \approx 2,718$ — основание натурального логарифма. Вероятность найти электрон при $r = \frac{a_B^*}{2}$ в 2,7 раза меньше, чем при $r = 0$; при $r = a_B^*$ в 7,4 раза меньше, чем при $r = 0$; при $r = 1,5a_B^*$ в 20 раз меньше. Определённо можно сказать, что эффективный радиус атома должен быть пропорционален длине a_B^* :

$$r_0 = qa_B^*. \quad (8.6)$$

Это важное утверждение. Оно означает, что численный коэффициент q должен обладать определенной универсальностью: он сравнительно слабо меняется при переходе от полупроводника к полупроводнику, в то время как величина a_B^* меняется сильно.

Заранее вычислить величину q из каких-либо физических соображений очень сложно. Более простой путь состоит в том, чтобы в рамках сформулированной выше модели найти критическую концентрацию N_c соответствующую переходу Мотта. Она будет выражена через радиус шарика r_0 . Затем, пользуясь формулой (8.6), нужно выразить N_c через q и a_B^* . Согласно экспери-

ментальным данным величина N_c определяется формулой (8.5). Сравнивая теоретически выражение с формулой (8.5), можно найти величину q , т. е. как бы определить её из экспериментальных данных.

Перейдём к осуществлению этой программы. Нужно решить следующую задачу теории протекания. В трёх мерном пространстве построены шары с радиусом r_0 центры которых распределены в пространстве хаотически и в среднем равномерно. Среднее число центров шаров, приходящееся на единицу объёма, равно N . Два шара считаются связанными, если они перекрываются друг с другом (рис. 8.1, б). Нужно найти критическое значение концентрации N_c , при котором возникает протекание по перекрывающимся шарам, т. е. возникают пути, проходящие через всю систему и состоящие из перекрывающихся шаров (рис. 8.1, б).

Эта задача отличается от задачи сфер, сформулированной в предыдущей главе, тем, что в задаче сфер связанными считались охватывающие, а не перекрывающиеся сферы, т. е. сферы, центры которых находятся на расстоянии, меньшем, чем r_0 , а не $2r_0$. Эта разница, однако, малосущественна — результаты одной задача без труда переносятся на другую. Действительно, если при некоторой концентрации центров существует бесконечный путь по охватывающим сферам радиуса $2r_0$, то существует и путь по перекрывающимся сферам радиуса r_0 . Это просто тот же самый путь, т. е. путь по тем же центрам. На рис. 8.2 эта мысль иллюстрируется для случая плоской задачи. (Легко понять, как должен выглядеть соответствующий чертёж для объёмной задачи.) Если нет пути по охватывающим сферам радиуса $2r_0$, то нет и пути по перекры-

вающимся сферам радиуса r_0 . Отсюда следует, что критическая концентрация, соответствующая порогу протекания по охватывающим сферам радиуса $2r_0$, равна критической концентрации, соответствующей порогу протекания по перекрывающимся сферам радиуса r_0 .

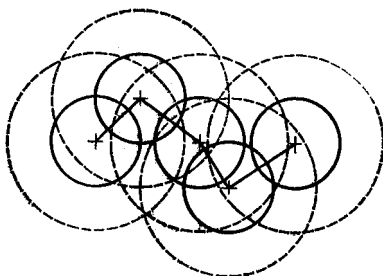


Рис. 8.2. Штриховыми линиями показана цепочка охватывающих окружностей с радиусом $2r_0$. Центры окружностей помечены крестами. Путь протекания по центрам показан ломаной линией. По этим же центрам идет протекания по перекрывающимся окружностям радиуса r_0 .

Согласно результатам предыдущей главы критическая концентрация для протекания по охватывающим сферам радиуса $2r_0$ определяется условием

$$\frac{4\pi}{3}N_c(2r_0)^3 = B_c \approx 2,7. \quad (8.7)$$

Подставляя выражение (8.6) в (8.7), получим

$$N_c(a_B^*) = \frac{0,08}{q^3}. \quad (8.8)$$

Коэффициент q можно теперь найти, сравнивая выражение (8.8) с полученным из экспериментальных данных выражением (8.5). Для того чтобы эти выражения совпадали, следует положить

$q = 1,6$. Следовательно, эффективный радиус r_0 равен $1,6a_B^*$. Вероятность найти электрон в точке, находящейся на расстоянии $r = r_0$ от ядра, в 24 раза меньше, чем в точке $r = 0$.

Самое главное в этом описании перехода диэлектрик–металл состоит, конечно, не в вычислении величины q , определяющей эффективный радиус атома, а в том, что теория протекания позволяет представить, как устроено полупроводник при концентрации примесей, близкой к N_c

Если предложенное описание правильно, то в полупроводнике имеется система металлических каналов, пронизывающая его насквозь. По этим каналам, как по проводам, течёт электрический ток. Как будет показано в третьей части книги, доля объёма полупроводника, занятая этими каналами, очень мала, если концентрация примесей близка к N_c . С этим связаны специфические свойства электропроводности и других важных характеристик полупроводника.

Упражнение

1. Антимонид индия (**InSb**) представляет собой полупроводник с очень малой шириной щели (0,18 эВ при комнатной температуре). У таких полупроводников малы и эффективные массы. Считая, что эффективная масса m^* электрона равна $0,015m$, а диэлектрическая проницаемость $\varepsilon = 18$, вычислите эффективный боровский радиус a_B^* и критическую концентрацию N_c , соответствующую переходу Мотта диэлектрик–металл.

Глава 9.

Различные обобщения задачи сфер

9.1. Охватывающие фигуры произвольной формы

Как уже говорилось, движение электрона в кристалле описывается так называемой эффективной массой m^* , которая может существенно отличаться от массы свободного электрона, т. е. массы электрона, находящегося в пустом пространстве. Оказывается, что эффективная масса может отличаться от массы свободного электрона не только по величине. Дело в том, что в кристалле не все направления эквивалентны. Например, характер движения электрона вдоль ребер куба может отличаться от характера его движения вдоль диагоналей куба и диагоналей граней куба. Поэтому эффективная масса может быть разной в разных направлениях. Тогда область, в которой движется электрон в окрестности донорного атома, уже не будет иметь вид шара. Она может быть эллипсоидом или ещё более сложной фигурой.

В связи с этим задача сфер была обобщена на случай фигур произвольной формы. Новая задача формулируется так: заданы узлы, распределенные в пространстве хаотически и в среднем равномерно-

но. Концентрация узлов равна N . Вокруг каждого узла построена одна и та же замкнутая поверхность произвольной формы.

Поверхности, построенные вокруг разных узлов, не только одинаковы по форме, но и одинаково ориентированы в пространстве. Если, например, поверхности имеют форму рыбы, то хвосты всех рыб должны смотреть в одну сторону.

Объём, находящийся внутри одной поверхности, равен V . Два узла считаются связанными друг с другом, если один из них находится внутри поверхности, построенной вокруг другого (охватывающие поверхности). Нужно найти критическое значение концентрации N , при котором возникает протекание по связанным узлам.

Как и в задаче сфер, наличие или отсутствие протекания определяется только значением параметра B , представляющего собой произведение VN . Это есть среднее число узлов, находящихся в объёме, ограниченном одной поверхностью. Оно не изменится, если увеличить или уменьшить масштабы вдоль всех трёх направлений, т. е. умножить или разделить координаты всех узлов системы и всех точек, лежащих на поверхностях, на одно и то же число. Ясно также, что такое изменение масштабов не уничтожит протекание, если таковое имеется, и не создаст его. Таким образом, протекание «не реагирует» на изменения N и V , при которых не меняется величина B . Поэтому, как и в задаче сфер, удобно говорить не о критической концентрации N_c , а о критическом значении параметра B :

$$B_c = N_c V. \quad (9.1)$$

Если, не меняя формы поверхности, построенной вокруг каждого узла, увеличить объём V , ограниченный этой поверхностью, например, в два раза, то критическая концентрация N_c уменьшится в два раза, а параметр B_c не изменится. Он зависит лишь от формы поверхности. Формула (9.1) обобщает формулу

$$B_c = \frac{4}{3}\pi N_c R^3, \quad (9.2)$$

использованную ранее для задачи сфер.

В общем случае критическое значение B_c , при котором возникает протекание, не должно совпадать со значением $B_c = 2,7$, полученным для задачи сфер. Вопрос о значениях B_c для разных фигур является в настоящее время предметом интенсивного изучения.

9.2. Задача эллипсоидов

Советские ученые Я. Г. Синай и Б. И. Шкловский доказали, что некоторые разные поверхности имеют одно и то же значение B_c . В частности, такими поверхностями являются эллипсоид и сфера (а в плоском случае эллипс и окружность).

Напомним сначала, что такое эллипс и эллипсоид. Эллипсом называют изображённую на плоскости замкнутую кривую, которая получается из окружности растяжением (или сжатием) вдоль одной из осей (рис. 9.1). Для того чтобы произвести растяжение или сжатие, нужно преобразовать каждую точку M , лежащую на окружности и имеющую координаты x, y , в точку M' с коорди-

натами $x' = kx, y' = y$, где k — коэффициент растяжения ($k > 1$ соответствует растяжению, а $k < 1$ — сжатию).

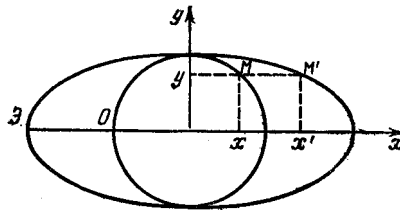


Рис. 9.1. Эллипс Э, полученный из окружности O растяжением вдоль оси x .

Поверхность, которая получается из сферы растяжением (или сжатием) вдоль одной из осей, называется эллипсоидом вращения. Для того чтобы произвести растяжение или сжатие вдоль оси z , необходимо каждую точку, лежащую на сфере и имеющую координаты x, y, z , преобразовать в точку с координатами $x' = x, y' = y, z' = kz$. Фигура, которая при этом получается, изображена на рис. 9.2. Она называется эллипсоидом вращения, потому что переходит сама в себя при повороте на любой угол вокруг оси z (называемой осью вращения). Если рассечь её плоскостью, проходящей через ось z , то в сечении получится эллипс. Сечение плоскостью, перпендикулярной оси z , представляет собой окружность.

Эллипсоид общего вида получается из эллипсоида вращения путём растяжения (или сжатия) вдоль одной из осей, перпендикулярных оси вращения. Новый коэффициент растяжения может отличаться от k . Сечение эллипсоида общего вида любой из плоскостей, параллельной плоскостям xOy, xOz, yOz , представляет

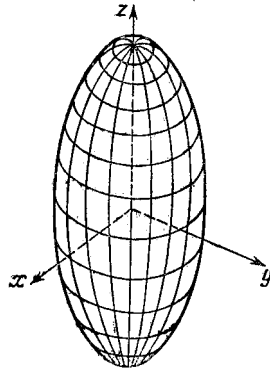


Рис. 9.2. Эллипсоид вращения.

собой эллипс.

Докажем теперь, что значения B_c для эллипсоидов и сфер одинаковы. Пусть заданы узлы, хаотически распределённые в пространстве со средней концентрацией, равной N . Вокруг каждого узла построена сфера с радиусом R . Если $B = \frac{4}{3}\pi NR^3 > B_c^{CB}$, то протекание по сферам есть, а при выполнении обратного неравенства протекание по сферам нет. Здесь $B_c^{CB} = 2,7 \pm 0,1$.

Произведём растяжение вдоль осей z и x по закону $y' = y$, $x' = k_1 x_1$, $z' = k_2 z$ так, чтобы этому преобразованию подверглись и координаты узлов, и координаты точек, лежащих на сферах. (Если бы речь шла о плоской задаче, то можно было бы представить, что узлы и окружности нанесены на резиновой полосе, которую затем растягивают в одном направлении. А можно вообразить и трёхмерное «резиновое пространство», внутри которого помечены точки, изображающие узлы и поверхности сфер. Затем это «резиновое пространство» растягивается в двух направлениях.)

Если x — координаты узлов представляют случайные числа, равномерно распределённые в интервале от 0 до L , где L — раз-

мер системы, то новые координаты x' , полученные умножением на k_1 , также представляют собой случайные числа, равномерно распределённые в интервале от 0 до k_1L . Координаты z' равномерно распределены в интервале $(0, k_2L)$, а координаты y не изменились. Концентрация узлов N' стала отличаться от концентрации N .

Все сферы преобразовались в эллипсоиды с объёмом V . (Можно показать, что $V' = k_1k_2 \cdot \frac{4}{3}\pi R^3$. Для дальнейшего вывода это не существенно.) Поэтому в результате из задачи сфер получилась задача эллипсоидов. Для неё нужно ввести величину $B' = N'V'$ и найти критическое значение B' , при котором возникает протекание по эллипсоидам. Обозначим его через $B_C^{\text{э}}$.

Дальнейшее доказательство разбивается на следующие пункты:

1. $B = B'$. Все узлы, которые до преобразования лежали внутри некоторой сферы, после преобразования оказываются внутри эллипсоида, получившегося из этой сферы. Действительно, если сфера растягивается (или сжимается), то на любом этапе растяжения внутренние по отношению к ней точки остаются внутренними, а наружные — наружными. Следовательно, среднее число узлов B , которое до преобразования находилось внутри одной сферы, равно среднему числу узлов B' , которое после преобразования находится внутри одного эллипсоида.

2.

$$\text{Если } B > B_C^{\text{сф}}, \text{ то } B > B_C^{\text{э}} \quad (9.3)$$

$$\text{Если } B < B_C^{\text{сф}}, \text{ то } B < B_C^{\text{э}} \quad (9.4)$$

Действительно, если до преобразования две сферы были связаны, т. е. центр одной из них находился внутри другой, то два эллипсоида, возникшие из этих сфер, также будут связанными, поскольку в результате преобразования внутренние точки останутся внутренними, а наружные — наружными. Если две сферы не были связаны, то получившиеся из них эллипсоиды также не будут связанными. Отсюда следует, что если $B > B_c^{\text{сф}}$, т. е. существуют бесконечные пути протекания по связанным сферам, то существуют бесконечные пути протекания по связанным эллипсоидам, что в свою очередь означает, что B' больше, чем пороговое значение $B_c^{\text{э}}$. Так как $B' = B$ отсюда получаем условие (9.3). Если $B < B_c^{\text{сф}}$, т. е. протекания по связанным сферам не существует, то не существует и протекания по связанным эллипсоидам, т. е. $B' < B_c^{\text{э}}$. Отсюда следует условие (9.4).

3. Так как условия (9.3) и (9.4) должны выполняться при любом значении B , то из них вытекает, что $B_c^{\text{сф}} = B_c^{\text{э}}$, что и требовалось доказать.

9.3. Другие поверхности

Задача эллипсоидов очень важна для физики полупроводников, однако в ряде случаев ее недостаточно. Поэтому физиков очень интересовал вопрос о том, как зависит критическое значение B_c от формы поверхности в тех случаях, когда точные соотношения отсутствуют. С этой целью производились расчеты на ЭВМ по ме-

тоту Монте-Карло, которые продемонстрировали удивительный и очень важный факт: B_c очень слабо зависит от формы поверхности.

Изучались фигуры, имеющие вид куба и тетраэдра. Оказалось, что в пределах точности расчёта ($\pm 0,1$) критические значения для этих фигур не отличаются ни друг от друга, ни от значения B_c для задачи сфер. Старались найти фигуру, максимально «непохожую» на сферу. В качестве такой фигуры выбрали «трёхмерный крест» — фигуру, образованную тремя вытянутыми параллелепипедами, пересекающимися в начале координат (рис. 9.3). Оказалось, что для такой поверхности значение B_c всего на 20% меньше, чем для сферы.

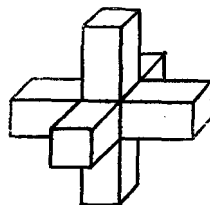


Рис. 9.3. "Трёхмерный крест"

То же самое выяснилось и для плоских фигур. Исследования показали, что значение B_c для квадратов отличается от значения B_c для окружностей всего на 2%.

Итак, в классе фигур одной размерности (т. е. либо плоских, либо объёмных) величина B_c является с хорошей точностью универсальной. Может быть, за этим кроются какие-либо соотношения, подобные доказанному в предыдущем разделе, и там, где расчёты на ЭВМ обнаруживают слабое отличие или не обнаруживают от-

личия в пределах точности, на самом деле имеют место строгие равенства. (К слабым отличиям, полученным с помощью ЭВМ, следует относиться критически, так как правильно оценить точность результатов не всегда просто.) К сожалению, ничего, кроме сказанного, в настоящее время не известно.

9.4. Ещё один эксперимент на домашней кухне и задача твердых сфер

Три студента Гарвардского университета (США) выполнили следующий простой опыт. В сосуд положили 5000 маленьких шариков, часть которых была сделана из алюминия, а другая часть — из пластика. Шарики были заранее тщательно перемешаны. Сосуд хорошенько потрясли, чтобы добиться максимально плотной упаковки шариков. На дне сосуда был электрод из фольги, сверху положили ещё один электрод и сдавили систему нагрузкой в 15 кг.

Алюминий является хорошим металлом, а пластик — диэлектриком. Цель опыта состояла в том, чтобы найти критическую долю x_c шариков из алюминия, при которой возникает ток между электродами, т. е. возникают пути до касающимся друг друга алюминиевым шарикам. Оказалось, что $x_c \approx 0,25$.

Кроме того, опыт позволял исследовать электропроводность системы в зависимости от x при $x > x_c$.

Речь идёт ещё об одной задаче теории протекания — задаче твёрдых сфер. Пожалуй, ближе всего эта новая задача стоит к задаче узлов. Вспомним построение, обсуждавшееся в гл. 6 в связи

с приближенной оценкой порога протекания задачи узлов. Вокруг каждого узла решётки строится сфера с радиусом, равным половине расстояния до ближайшего соседа. Сферы, построенные вокруг белых узлов, называются белыми, а построенные вокруг чёрных узлов, — чёрными. Протекание по белым узлам эквивалентно существованию путей по касающимся друг друга белым сферам (см рис. 6.5).

Различие между задачей узлов и новой задачей состоит в том, что в задаче узлов центрами сфер являются узлы регулярной решётки, а в новой задаче центры сфер могут находиться где угодно. Ниже показано, что это различие не очень важно.

Разница между новой задачей и задачей сфер, обсуждавшейся в предыдущих главах и предыдущих разделах этой главы, состоит в том, что сферы, фигурирующие в новой задаче, предполагаются твёрдыми: они не перекрываются друг с другом. Это различие очень существенно.

В гл. 6 показано, что в случае задачи узлов протекание по белым узлам возникает, когда доля пространства, заполненная построенными вокруг этих узлов шарами, равна примерно 0,16. Оказалось, что это число почти не зависит от типа решётки. Естественно предположить, что раз оно почти не зависит от типа решётки, то оно не должно сильно зависеть от того, есть ли вообще решётка. Если это правильно, то доля объёма, заполненного металлическими шариками, при которой по этим шарикам возникает протекание, должна примерно равняться 0,16.

Как и в гл. 6, обозначим через f коэффициент заполнения, т. е. долю объёма, занятую всеми шариками — и алюминиевыми, и пла-

стикowymi. Величина x по определению представляет отношение числа алюминиевых шариков к полному числу шариков. Поэтому доля объёма, занятая алюминиевыми шариками, равна fx . Если на пороге протекания эта доля объёма равна 0,16, то критическое значение x_c можно найти из условия $fx_c = 0,16$.

Коэффициент заполнения для системы плотно упакованных, но случайно расположенных шариков хорошо известен. Подобная система знакома человечеству с античных времён. Когда нужно было отмерить определённое количество зерна или какого-нибудь другого сыпучего товара, его насыпали в специальный сосуд (меру), трясли и уминали. В современной науке эта система служит моделью расположения атомов в аморфных металлах. Аморфные металлы или, как их ещё называют, металлические стекла — это вещества с металлической электропроводностью, но без кристаллической структуры. Оказалось, что расположение атомов в аморфном металле очень напоминает расположение плотно упакованных несжимаемых шаров. Благодаря этому, свойства плотно упакованных случайно расположенных шаров изучались с большой тщательностью (главным образом, с помощью ЭВМ). Выяснилось, что доля объёма, занимаемого шарами, равна $f = 0,637$.

Вернемся к задаче о протекании по металлическим шарам. Определив x_c по формуле $x_c = \frac{0,16}{f}$, получим $x_c = 0,25$, что совпадает с результатом трёх студентов.

Впоследствии эксперименты по определению порога протекания многократно повторялись с использованием более совершенной техники. Ставились опыты, в которых фигурировала смесь шаров разного радиуса. В этой смеси менялись в широких пре-

делах и радиусы металлических, и радиусы диэлектрических шаров. Оказалось, что в этом случае критическая доля объёма $f x_c$ приблизительно равна 0,17 и не отличается в пределах точности эксперимента от случая шаров одного радиуса.

Итак, порог протекания в задаче твёрдых сфер можно сравнительно легко оценить, считая, что критическая доля объёма, заполненная металлическими шарами, равна примерно 0,16. Заметим, что в задаче перекрывающихся сфер это число значительно больше.

Задача твёрдых сфер оказалась очень важной в прикладном отношении. Она лежит в основе теории гетерогенных материалов, представляющих собой диэлектрики с очень мелкими металлическими вкраплениями. В настоящее время такие материалы широко изучаются. Они изготавливаются и используются как в виде тонких плёнок, так и в виде объёмных образцов. Вблизи порога протекания эти вещества обладают удивительными электрическими свойствами. Например ёмкость конденсатора, заполненного таким материалом, неограниченно растёт, если доля объёма, занятого металлом, приближается к порогу протекания. Это явление связано с аномальной взаимной ёмкостью больших металлических кластеров. Описание электрических свойств гетерогенных материалов в настоящее время превращается в самостоятельный раздел теории протекания.

Глава 10.

Уровень протекания

10.1. «Всемирный потоп»

Ходит ветер шесть дней, семь ночей,
Потопом буря покрывает землю,
При наступлении дня седьмого
Буря с потопом войну прекратили,
Успокоилось море, утих ураган — потоп прекратился,
Я открыл отдушину — свет упал на лицо мне,
Я взглянул на море — тишь настала.
И всё человечество стало глиной!
Плоской, как крыша, сделалась равнина.

Так описывает всемирный потоп шумерский миф о Гильгамеше — древнейшее литературное произведение, относящееся примерно ко второму тысячелетию до нашей эры (перевод И. Дьяконова).

Когда вода стала спадать, показались над её поверхностью самые высокие на Земле горы. Вода опускалась всё ниже и ниже, пока не достигла своего обычного уровня. Вообразите грандиозную картину: огромная горная система, например Тянь-Шань, постепенно выходит из воды. Сначала появились как острова самые

высокие вершины, потом показались зоны альпийских лугов, и, наконец, вода опустилась в предгорье.

Допустим, что нас интересует вопрос: до какого уровня должна опуститься вода, чтобы исчез последний водный путь, проходящий через всю горную систему (рис. 10.1)? Ясно, что такой путь существует, пока определенная часть перевалов не вышла из воды, а затем он исчезает.

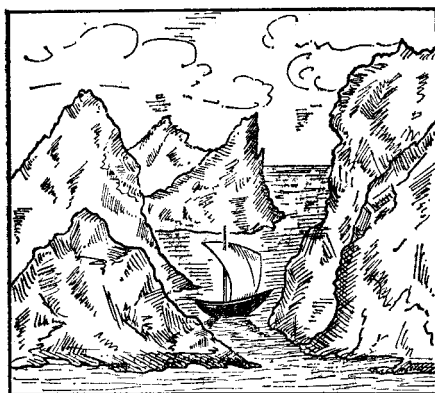


Рис. 10.1. Путешествие во время потопа.

На рис. 10.2 показана географическая карта горной системы, на которой тонкими линиями изображены горизонталы (линии одинаковой высоты). Толстыми линиями обведены горизонталы, соответствующие уровню воды. Они отделяют воду от суши. Участки, залитые водой, на рисунке заштрихованы. На рис. 10.2, а показано самое начало потопа: водные озёра ещё не сообщаются друг с другом. Затем вода поднялась выше, и на рис. 10.2, б уже возникли водные пути, пересекающие всю систему. На рис. 10.2, в из воды поднимаются лишь отдельные горные вершины.

Уровень воды, при котором появляется (или исчезает) водный путь, называется *уровнем протекания*. Задача о его определении

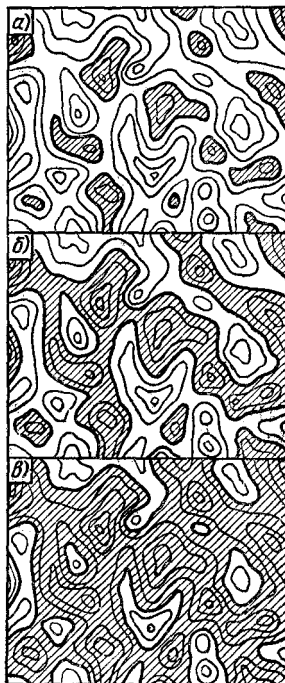


Рис. 10.2. Карта горной системы, залитой водой. Залитые области заштрихованы.

представляет собой плоскую задачу теории протекания. Она может быть сформулирована и иначе. Допустим, что плоскость хаотически раскрашена белой и чёрной краской. Пусть доля площади, раскрашенной белой краской, равна x . При малых x белые куски образуют изолированные острова, а при значениях x , близких к единице, изолированы, наоборот, чёрные куски. Нужно найти критическое значение x , при котором исчезает (или появляется) сквозной путь через всю систему, идущий только по белым областям.

Аналогичным образом можно сформулировать и объемную задачу, заполнив большой объём белым и чёрным веществом. Затем нужно менять долю объёма, заполненную одним из веществ, пока

по нему не возникнет протекание.

10.2. Построение случайной функции**

Новая задача не является решёточной. Разберем её математическую формулировку в плоском случае. Необходимо задать на всей плоскости случайную функцию $V(X, Y)$, где X и Y — координаты. В задаче о «всемирном потопе» эта функция представляет высоту над уровнем моря в точке земной поверхности с заданными координатами. Ограничимся случаем гауссовых случайных функций (названных по имени великого математика К. Гаусса).

Простейший способ построения гауссовой случайной функции состоит в том, чтобы приписать каждой точке пространства случайное число, никак не связанное с соседним случайным числом. Такая случайная функция называется «белый шум». Её значения в соседних точках резко отличаются, т. е. она является разрывной. Чтобы получить непрерывную функцию, белый шум «сглаживают». Эта процедура состоит в том, что каждой точке пространства приписывается величина, представляющая собой среднее от значений, принимаемых функцией «белый шум» в некоторой области пространства, окружающей данную точку. Эти величины и образуют непрерывную случайную функцию. При переходе к соседней точке значения функции меняются мало, поскольку мало меняется область пространства, по которой усредняется белый шум. Обозначим через r_0 размер области пространства, по которой ведется усреднение. Этот размер называют радиусом корреляции случайной функции. Главное его свойство состоит в том, что при изме-

нении аргумента функции на величину, малую по сравнению с r_0 , значение функции изменяется мало.

Можно ввести функцию распределения величин V . Обозначим её через $f(V)$. По определению вероятность того, что во взятой наугад точке пространства функция $V(X, Y)$ имеет значение, заключенное в малый интервал от V_1 до $V_1 + \Delta V$, равна $f(V_1)\Delta V$.

Функцию $V(X, Y)$ всегда можно построить так, что её среднее по всем точкам плоскости значение равно нулю. Для этого нужно при построении функции «белый шум» использовать случайные числа, симметрично распределенные относительно нуля. Можно показать, что функция $V(X, Y)$, полученная из такого «белого шума», обладает гауссовым распределением (см. гл. 2) вида

$$f(V) = \frac{1}{\delta\sqrt{2\pi}} e^{-\frac{V^2}{2\delta^2}}. \quad (10.1)$$

Заметим, что плотность вероятности (10.1) симметрична относительно положительных и отрицательных значений V . На «языке гор» это означает, что пики и впадины встречаются с одинаковой вероятностью.

Теперь легко сформулировать задачу протекания. Горизонталы, изображённые на рис. 10.2, определяются условием $V(X, Y) = \text{const}$. Для того чтобы задать уровень воды, нужно ввести число t , изменяющееся в интервале от $-\infty$ до ∞ . Участки плоскости, в которых $V(X, Y) < t$, назовем белыми (залитыми водой), а участки, где $V(X, Y) > t$, — чёрными (выступающими из воды). Толстые линии на рис. 10.2 определяются условием $V(X, Y) = t$.

Уровнем протекания называется критическое значение t_c , при котором белые области образуют пути, пронизывающие всю си-

стему. Можно также говорить о критической доле пространства x_c , заполненной белыми областями в момент возникновения протекания. Эта доля пространства равна вероятности того, что непрерывная случайная величина V принимает какое-либо значение в интервале $-\infty < V < t_c$. По смыслу функции распределения

$$x_c = \int_{-\infty}^{t_c} f(V) dV. \quad (10.2)$$

Формула (10.2) связывает критическую долю пространства x_c с уровнем протекания t_c .

10.3. Аналогия с задачей узлов**

Представьте себе, что на географическую карту, показанную на рис. 10.2, нанесена плоская решётка (всё равно какая) с малым периодом. Пусть период решётки будет значительно меньше, чем характерные размеры областей суши и воды. Назовём узлы решётки, попавшие в области, залитые водой, белыми, а узлы, попавшие на сушу, — черными. Так же, как в обычной задаче узлов, будем считать белые узлы связанными, если они являются ближайшими соседями. Критическая доля площади, заполненная водой, равна критической доле белых узлов, при которой по ним возникнет протекание.

Однако новая задача не тождественна задаче узлов. Вспомним гл. 4, где подробно описано построение системы белых и черных узлов. (В этой главе назывались блокированными и неблокированными.) Важным моментом этого построения было то, что каждый

узел становился белым или чёрным по воле генератора случайных чисел и без всякой связи с тем, какой цвет имели узлы, являющиеся его ближайшими соседями. Вследствие этого белые и чёрные узлы были хорошо перемешаны друг с другом. В новой задаче это совсем не так. Поскольку период решётки очень мал, чёрные и белые узлы располагаются большими блоками. Соседом белого узла с подавляющей вероятностью является белый узел, а соседом чёрного узла — чёрный узел. (Размеры блоков определяются радиусом корреляции r_0).

Следует понимать, что введение решётки в задачу об уровне протекания является чисто формальным приёмом. По существу задача не является решёточной и её ответы не должны зависеть ни от периода решётки (если он достаточно мал), ни от типа решётки. Однако такой приём позволяет воспользоваться хорошо разработанным аппаратом решёточных задач.

В гл. 4 приведён алгоритм решения задачи узлов методом Монте-Карло на ЭВМ. Этот алгоритм полностью переносится на задачу об уровне протекания. Тем же способом, который описан в гл. 4, с помощью случайной функции $V(X, Y)$ составляется массив $K(X, Y)$, состоящий из нулей и единиц. Белым узлам соответствуют единицы, а чёрным — нули. Тем же способом осуществляется поиск путей протекания и определение критической доли пространства, при которой возникает протекание.

10.4. Уровни протекания в плоской в трёхмерной задачах**

Плоская задача имеет точное решение, если свойства случайной функции $V(X, Y)$ в среднем симметричны относительно значения $V = 0$. Такими свойствами обладают, в частности, описанные выше гауссовы функции.

Для того чтобы получить точное решение, следует воспользоваться симметричной формулировкой задачи протекания, которая была дана в гл. 5. Как и раньше, будем называть области, в которых $V(X, Y) < t$, белыми, а области, где $V(X, Y) > t$, — чёрными. Кроме уровня протекания по белым областям t_c , можно ввести уровень протекания по чёрным областям t_c . Дальнейшее доказательство разбивается на следующие пункты:

1. Из симметрии функции $V(X, Y)$ следует, что $t_c = -t'_c$. Действительно, заменим в каждой точке плоскости V на $-V$. При этом получится функция $V' = -V$, в среднем обладающая теми же свойствами, так что уровни протекания, вычисленные с её помощью, должны быть такими же, как уровни, вычисленные с помощью исходной функции. Из неравенства $V < t_c$ следует неравенство $-V > -t_c$, т. е. $V' > -t_c$. При $t = t_c$ белые с точки зрения функции V области ($V < t_c$) образуют бесконечный кластер. Эти же самые области пространства являются чёрными, если рассматривать функцию V' и $t = -t_c$. Ведь в них $V' > -t_c$. Следовательно, значение $t = -t_c$ является уровнем протекания по чёрным областям для функции V' . Но, как уже говорилось, уровни протекания у функций V и V'

должны быть одинаковыми. Следовательно, и для функции V уровнем протекания по чёрным областям является $t'_c = -t_c$.

2. Если $t_c < 0$, то по мере увеличения t сначала появляется протекание по белым областям (при $t = t_c < 0$), а затем (при $t = -t_c > 0$) исчезает протекание по чёрным областям. В области $t_c < t < -t_c$ существует протекание и по белым, и по чёрным областям. Если $t_c > 0$, то сначала исчезает протекание по чёрным областям, а затем появляется протекание по белым областям. При этом в области $-t_c < t < t_c$ нет никакого протекания.
3. В случае плоской задачи протекание по чёрным областям исключает протекание по белым областям, и наоборот. Действительно, если по горной системе можно проплыть с запада на восток на корабле, значит, по ней нельзя пройти с севера на юг посуху. Следовательно, случай $t_c < 0$ исключается. С другой стороны, отсутствие протекания по белым областям обязательно означает наличие протекания по чёрным областям с севера на юг. (В этом можно убедиться, изучая картинки типа рис. 10.2.) Таким образом, исключается и случай $t_c > 0$. Следовательно, остается одна возможность $t_c = 0$. В этом и состоит результат: уровень протекания равен нулю.

С помощью формулы (10.2) можно вычислить критическую долю площади x_c . Из условия нормировки функции распределения (см. формулу (2.1)) следует, что при $t_c = \infty$ правая часть формулы (10.2) равна единице. Из симметрии функции $f(y)$ вытекает, что при $t_c = 0$ доля площади $x_c = 0,5$.

В трёхмерном случае протекание по белым областям с запада на восток не исключает протекания по чёрным областям с севера на юг, так как каналы протекания легко могут быть развязаны. (Вспомните развязки на шоссе, которые делаются на разных уровнях высоты.) Поэтому в трёхмерном случае $t_c < 0$ и, согласно формуле (10.2), $x_c < 0,5$. Расчёт, произведенный методом Монте-Карло по указанной выше схеме для гауссовых случайных функций, показал, что в трёхмерном случае $x_c = 0,16 \pm 0,01$.

Для приближённой оценки величины x_c можно воспользоваться методом, с помощью которого в гл. 6 оценивались пороги протекания задачи узлов. Метод состоит в том, что вокруг каждого узла строятся шары (или круги в плоском случае) с радиусом, равным расстоянию до ближайшего соседа. Шар, построенный вокруг белого узла, считается белым, а построенный вокруг чёрного узла — чёрным. Оказалось, что протекание по касающимся друг друга белым шарам возникает, когда доля пространства, заполненного этими шарами, примерно одинакова для всех решёток. Естественно предположить, что эта доля должна быть близка к значению x_c , которое фигурирует в задаче об уровне протекания. Согласно формулам (6.3) и (6.4), в плоском случае доля поверхности, занятой белыми кругами, примерно равна 0,5, а в трёхмерном случае доля объёма, занятого белыми шарами, примерно равна 0,16. Таким образом, обе оценки удивительно точно совпадают с решением задачи об уровне протекания. Можно надеяться, что $x_c = 0,16$ даёт неплохую оценку и для негауссовых случайных функций.

10.5. Компенсация примесей в полупроводниках

Уровень протекания играет очень важную роль в теории полупроводников с примесями. Допустим, что в полупроводник введены примерно в одинаковом количестве и донорные, и акцепторные примеси. Донорные примеси имеют лишний электрон на внешней оболочке, а акцепторным примесям, наоборот, не хватает одного электрона. Поэтому доноры охотно отдают свои электроны акцепторам и становятся при этом положительно заряженными. Акцепторы принимают электроны и заряжаются отрицательно. (Это явление называют компенсацией примесей.) Так как примеси в полупроводниках расположены неупорядоченно, то возникает хаотическая система положительно и отрицательно расположенных зарядов. Каждый заряд создает вокруг себя электрический потенциал, равный $\pm \frac{e}{\epsilon r}$, где e — абсолютная величина заряда электрона, ϵ — диэлектрическая проницаемость, r — расстояние от заряда. Знак потенциала определяется знаком заряда.

Потенциал любой точки пространства представляет собой сумму потенциалов, создаваемых всеми примесями. Так как примеси расположены в пространстве случайно, то потенциал также представляет собой случайную функцию.

Если доноров немного больше, чем акцепторов, то часть электронов остается на донорах. Если энергия связи электронов с донорами относительно мала, то тепловое движение атомов легко отрывает эти электроны. В принципе они могут участвовать в переносе электрического тока, но этот перенос затруднен тем, что в пространстве существует электрический потенциал, создаваемый

заряженными примесями. Электрический потенциал, помноженный на заряд электрона, представляет собой потенциальную энергию взаимодействия электрона с электрическим полем примесей. Потенциальная энергия существенным образом меняет характер движения электрона.

10.6. Движение частицы при наличии потенциальной энергии

Полная энергия частицы E складывается из кинетической энергии, равной $\frac{mv^2}{2}$, где m — масса частицы, а v — скорость, и потенциальной энергии $V(\mathcal{R})$ зависящей от координат точки, в которой находится частица:

$$E = \frac{mv^2}{2} + V(\mathcal{R}), \quad (10.3)$$

Здесь \mathcal{R} — вектор, проведенный из начала координат в точку, где находится частица.

Движение частицы описывается функцией $\mathcal{W}(t)$, где t — время. Основным принцип механики состоит в том, что во время движения полная энергия частиц E не меняется. Это значит, что при изменении скорость частицы меняется таким образом, чтобы скомпенсировать изменение функции $V(\mathcal{R})$. Закон сохранения энергии, а также тот факт, что кинетическая энергия $\frac{mv^2}{2}$ должна быть величиной положительной, накладывают важные ограничения на движение частицы.

Допустим для простоты, что V зависит лишь от одной координаты (от X) и скорость тоже направлена вдоль оси X . Пусть (X)

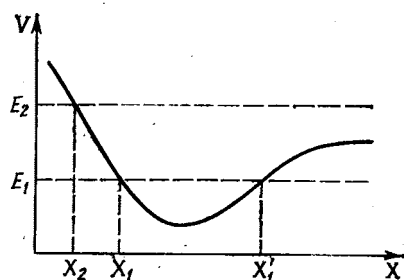


Рис. 10.3. Потенциальная энергия электрона V как функция координаты X (сплошная линия).

имеет вид, изображённый на рис. 10.3. Характер движения частицы определяется величиной её полной энергии. Если энергия равна E_1 , то кинетическая энергия положительна лишь в области $X_1 < X < X_1'$. Точки X_1 и X_1' определяют границы движения. Скорость в этих точках обращается в нуль, и потому эти точки называются *точками остановки*. Частица, имеющая энергию E_1 , захвачена *потенциальной ямой*, находящейся между точками остановки, и не может выйти из области этой ямы. Область её движения ограничена с двух сторон. Частица, имеющая энергию E_2 , не может проникнуть левее точки X_2 , зато область её движения не ограничена с правой стороны.

Таким образом, в зависимости от соотношения между полной и потенциальной энергиями движение частицы может быть ограниченным или неограниченным.

Заметим, что квантовая механика допускает проникновение частицы в область с отрицательной кинетической энергией. Однако, если функция $V(\mathcal{R})$ достаточно плавная (а именно так и обстоит дело в случае полупроводников с большой концентрацией компен-

сированных примесей) это проникновение не играет особой роли.

10.7. Движение электрона в поле примесей

Потенциальная энергия электрона, связанная с потенциалами хаотически расположенных примесей, является случайной функцией координат. Если полная энергия E электрона невелика, то он может совершать движение лишь в ограниченной области пространства, окруженной со всех сторон точками остановки. При достаточно большой энергии электрон получает возможность двигаться по всему пространству. Только в этом случае он может участвовать в переносе электрического тока.

Следует понимать различие между одномерной задачей, обсуждавшейся в предыдущем разделе, и двумерной или трёхмерной задачей. Если частица может двигаться только вдоль некоторой прямой, то для того чтобы движение было неограниченным в обоих направлениях, энергия частицы должна быть больше всех значений, которые принимает на этой прямой потенциальная энергия. Таким образом, должно выполняться жесткое условие $E > V_{\max}$, где V_{\max} — максимальное значение потенциальной энергии.

В двумерном и трёхмерном случаях это условие не требуется. Частица может *огibtать* области, в которых её движение запрещено. Нужно только, чтобы разрешенные для движения области образовывали систему «озер и каналов», по которым частица могла бы уйти в бесконечной системе на бесконечное расстояние.

Отсюда видно, что возникла задача об определении уровня протекания. Пусть $V(\mathcal{R})$ — случайная функция, представляю-

щая собой потенциальную энергию электрона. Зафиксировав полную энергию E , назовем белыми областями пространства области, в которых $E > V(\mathcal{R})$ (кинетическая энергия положительна), а чёрными — области, где $E < V(\mathcal{R})$. Нужно найти уровень протекания, т. е. критическое значение E_c , при котором возникает протекание по белым областям.

Только те электроны, которые имеют энергию, большую, чем E_c , являются свободными и принимают участие в переносе электрического тока (эту энергию иногда называют *порогом подвижности*). При низких температурах энергия E_c значительно больше, чем энергия теплового движения атомов kT . Поэтому вероятность того, что электрон приобретает энергию E_c , мала. Соответственно мала и концентрация электронов, способных переносить ток. С повышением температуры эта концентрация резко возрастает, вследствие чего возрастает и электропроводность системы. Таким образом, порог протекания E_c определяет температурную зависимость электропроводное полупроводника с компенсированными примесями.

При очень низких температурах электроны скапливаются в самых глубоких потенциальных ямах и не принимают участия в переносе тока. Поэтому при низких температурах полупроводник с компенсированными примесями превращается в диэлектрик.

Как говорилось в гл. 8, при достаточно большой концентрации примесей одного сорта (например, доноров) полупроводник приобретает металлическую проводимость, слабо зависящую от температуры вплоть до абсолютного нуля температур (переход Мотта). Компенсация примесей (например, добавление акцепторов)

приводит к росту хаотической потенциальной энергии и к срыву металлической проводимости. Существует детальная теория этого явления, основанная на представлениях об уровне протекания.

Часть III

**Критическое поведение различных
величин вблизи порога протекания и
геометрия бесконечного кластера**

В этой части речь идёт о наиболее интересных с точки зрения физики аспектах теории протекания — о поведении различных величин в непосредственной близости к порогу протекания. В предшествующих частях говорилось, что такие физические величины, как спонтанная намагниченность ферромагнетика с примесями или электропроводность сетки с заблокированными узлами обращаются в нуль в пороговой точке. В этой части обсуждаются законы, которые описывают их поведение вблизи порога. Для вывода этих законов необходимо понимать геометрические свойства бесконечного кластера.

Глава 11.

Решётка Бете**

В гл. 5 было показано, что пороги протекания некоторых плоских задач могут быть определены точно. Однако нигде не говорилось, что можно найти также функцию $P(x)$ — вероятность того, что некоторый узел принадлежит бесконечному кластеру. В настоящее время точные выражения для этой функции (так же, как и для электропроводности сетки) неизвестны ни для плоских, ни для трёхмерных задач. Исключение представляет так называемая решётка Бете, которую, как будет показано ниже, следует классифицировать как решётку в пространстве с бесконечным числом измерений. Ниже излагаются постановка и решение задачи узлов на решётке Бете.

11.1. Слухи

«Что ж за вздор, в самом деле, разнесли по городу? Что же за направление такое, что не успеешь повернуться, а тут уже и выпустят историю, а хотя бы какой-нибудь смысл был...» — возмущался Николай Васильевич Гоголь, описывая, как нелепая

сплетня, распущенная двумя дамами, погубила многообещающую авантюру Чичикова. — «Это предприятие удалось произвести им с небольшим в полчаса. Город был решительно взбунтован: всё пришло в брожение, и хотя бы кто-нибудь мог что-либо понять».

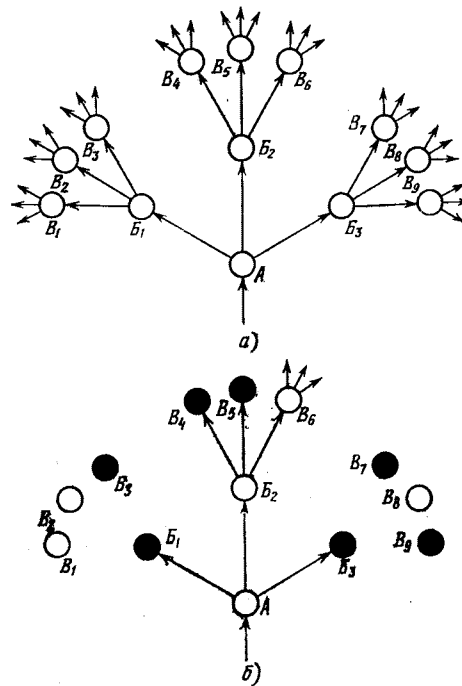


Рис. 11.1. Решётка Бете с $q = 3$. Светлые кружки — люди категории S , тёмные — категории T .

Слухи, действительно, распространяются со сказочной быстротой. И это перестаёт быть удивительным, если рассмотреть предлагаемую математическую модель.

Модель изображена на рис. 11.1, а. Допустим, что некая «во всех отношениях приятная дама», обозначенная кружком A , общилась сведения трём своим знакомым: B_1 , B_2 и B_3 . Каждый из них передал их ещё трём своим знакомым, так что «из вторых

рук» информацию получили 9 человек, обозначенные кружками B . Каждый из этих 9 человек тоже передал информацию трём своим знакомым, в результате чего с ней ознакомились ещё 27 человек. Легко подсчитать, что «из десятых рук» сведения получают $3^{10} = 59049$ человек! Если предположить, что на передачу информации трём своим знакомым каждый человек затратит 20 мин, то получится, что всё дело займёт $200 \text{ мин} = 3 \text{ ч } 20 \text{ мин}$.

Разумеется, эта модель сильно упрощает реально происходящий процесс. Считается, что число знакомых у всех людей одно и то же. Кроме того, предполагается, что каждый человек получает сведения только от одного человека. Это значит, что в каждый кружок (рис. 11.1, а) входит только одна линия. Благодаря этому свойству модель напоминает дерево, бесконечно разветвляющееся во все стороны. Каждый кружок можно считать основанием своего дерева, причём деревья, выросшие, например, из кружков B_1 , B_2 и B_3 , не имеют между собой ни одного общего кружка. То же самое можно сказать и о деревьях, основаниями которых являются кружки B , и т. д.

В научной литературе эта модель так и называется — *дерево*. Её называют также *решёткой Бете* по имени знаменитого физика Ганса Бете. Кружки, изображённые на рис. 11.1, а являются узлами этой решётки. Число линий, выходящих из каждого узла решётки Бете, может быть произвольным (но одинаковым для всех узлов). Обозначим это число через q . На рис. 11.1, а изображена решётка с $q = 3$.

Вспомним теперь, что большинство людей имеет собственную точку зрения и не участвует в распространении слухов. Разде-

лим всех людей на две категории: к категории C , изображаемой на рисунке светлыми кружками, отнесём людей, передающих полученную информацию своим знакомым (рис. 11.1, б). Из этих кружков может выходить q стрелок. К категории T , изображаемой чёрными кружками, отнесём людей, которые не участвуют в распространении слухов. Из чёрных кружков (рис. 11.1, б) не выходит ни одной стрелки.

Наличие чёрных кружков сильно влияет на продвижение слуха. Рассмотрим конфигурацию, изображённую на рис. 11.1, б. Из трёх кружков, B_1, B_2, B_3 , лишь один оказался светлым и передал слух дальше. Кружки B_1, B_2 и B_3 с удовольствием бы посплетничали, но B_1 и B_3 ничего им не передали. Из B_4, B_5, B_6 к C принадлежит лишь B_6 . Таким образом, из вторых рук вместо 9 человек слух дошёл лишь до трёх, причём в следующие руки его передаст лишь B_6 .

Допустим, что рассматриваемая система ничем не ограничена и имеет бесконечное количество кружков. Тогда можно поставить следующий вопрос. Умрёт после конечного числа передач вышедший из точки A слух или он уйдёт на бесконечное расстояние от A и в бесконечной системе станет достоянием бесконечного числа лиц? Как видно из рис. 11.1, б, это зависит от относительного количества светлых и тёмных кружков и от конфигураций, возникающих в окрестности узла.

По существу речь идёт о задаче узлов теории протекания, только сформулированной на решётке Бете. Пусть доля людей категории C равна x . Это значит, что выбранный наугад человек с вероятностью x окажется принадлежащим к категории C и с

вероятностью $1 - x$ — к категории T . Вопрос, который нужно решить, состоит в следующем. Какова вероятность $P(x)$ того, что слух, сообщённый выбранному наугад человеку, станет достоянием бесконечного числа лиц? Ясно, что при малых значениях x эта вероятность равна нулю, но, однако, она становится отличной от нуля, начиная с некоторого критического значения $x = x_c$.

11.2. Решение задачи узлов на решётке Бете

Вместо функции $P(x)$ удобно ввести вероятность того, что слух, сообщённый выбранному наугад человеку, не сделается достоянием бесконечного числа лиц. Обозначим эту вероятность через $Q(x)$. Очевидно, что

$$Q(x) = 1 - P(x), \quad (11.1)$$

так как рассматриваемые события образуют полную систему событий.

Для $Q(x)$ можно составить алгебраическое уравнение. Рассуждать нужно следующим образом. Передача слуха может прерываться по двум несовместимым причинам. Первая состоит в том, что выбранный наугад человек окажется из категории T . Вторая причина состоит в том, что хотя человек окажется из категории C и передаст слух q людям, все каналы, ведущие от этих людей, на разных этапах прервутся. Таким образом, вероятность Q представляет сумму вероятностей двух несовместимых событий. Вероятность того, что выбранный наугад человек окажется из категории T , равна $1 - x$. Вероятность того, что он окажется из категории C ,

но передача слуха прервётся на более далёких этапах, обозначим пока через W' . Тогда

$$Q = 1 - x + W'.$$

Перейдем теперь к вероятности W' . Исход, который она описывает, является следствием одновременного осуществления двух событий:

1. выбранный наугад человек оказывается C (вероятность этого события равна x),
2. все q каналов, ведущие от знакомых выбранного наугад человека, на каком-нибудь этапе прерываются.

Очевидно, что эти два события независимы. Поэтому вероятность W' равна произведению вероятностей: $W' = xW(x)$, так что

$$Q = 1 - x + xW(x), \quad (11.2)$$

где $W(x)$ — вероятность того, что все q каналов на каком-то этапе прерываются (разумеется, этот этап может быть разным у разных каналов).

Рассмотрим один из q каналов, начинающихся с одного из знакомых выбранного наугад человека. Событие, состоящее в том, что этот канал где-то прерывается, эквивалентно тому, что слух, сообщенный данному знакомому, не станет достоянием бесконечного числа лиц. По определению вероятность этого события равна $Q(x)$.

Для дальнейшего оказывается очень важным, что деревья, основаниями которых являются q знакомых выбранного наугад чело-

века, не имеют общих кружков. Отсюда следует, что если у одного дерева имеется определённая конфигурация светлых и тёмных кружков, то это не влияет на вероятность любой конфигурации кружков у других деревьев. (Очевидно, что при наличии общих кружков последнее утверждение было бы неверным.) Поэтому события, состоящие в том, что слух прерывается в одном и в другом каналах, независимы.

Таким образом, вероятность того, что все q каналов прерываются, равна произведению вероятностей того, что прерывается по отдельности каждый из q каналов:

$$W(x) = [Q(x)]^q. \quad (11.3)$$

Подставляя формулу (11.3) в (11.2), получаем уравнение для $Q(x)$:

$$Q(x) = 1 - x + x[Q(x)]^q. \quad (11.4)$$

Заметим, что решающим обстоятельством, позволившим свести задачу к алгебраическому уравнению (11.4), явилась независимость различных каналов. Это свойство присуще исключительно решётке Бете, и потому использованный выше метод в применении к обычным решёткам не приводит к успеху, хотя часто употребляется для получения приближённо правильного решения.

Перейдём к анализу уравнения (11.4). Оно имеет смысл при всех x в интервале $0 \leq x \leq 1$. Перепишем его через $P(x) = 1 - Q(x)$. Получим

$$[1 - P(x)]^q x + P(x) - x = 0. \quad (11.5)$$

Одним из решений (11.5) является $P(x) = 0$ при всех x , однако уравнение (11.5) при $q > 1$ нелинейное и имеет другие решения. В частности, при $x = 1$ решением является также $P(1) = 1$, причем физический смысл имеет именно второе решение, так как если все кружки светлые, то вероятность P должна равняться единице, а не нулю.

Решения уравнения (11.5) легко найти при $q = 2$. Решений в этом случае два: $P(x) = 0$ и $P(x) = 2 - \frac{1}{x}$. При $x > \frac{1}{2}$ физический смысл имеет второе решение. При $x < \frac{1}{2}$ оно становится отрицательным и не имеет смысла. Итак, при $q = 2$ смысл имеет следующее решение:

$$P(x) = \begin{cases} 0 & \text{при } 0 < x < \frac{1}{2} \\ 2 - \frac{1}{x} & \text{при } \frac{1}{2} < x < 1 \end{cases} \quad (11.6)$$

Порог протекания x_c равен в этом случае $\frac{1}{2}$.

Аналогичное решение существует при всех $q > 1$, однако порог протекания x_c зависит от q . В общем случае можно найти x_c и вид $P(x)$ при x , близком к x_c , заранее предположив $P(x) \ll 1$, что всегда справедливо в окрестности порога протекания. Член $(1 - P)^q$ в уравнении (11.5) можно разложить по формуле бинома

$$(1 - P)^q = 1 - qP + \frac{q(q - 1)}{2}P^2 - \dots \quad (11.7)$$

Так как $P \neq 1$, каждый последующий член значительно меньше предыдущего. Поэтому подставим формулу (11.7) в уравнение (11.5), считая, что в правой части (11.7) есть только три написанных члена. Получим

$$\frac{q(q - 1)}{2}xP^2 = qPx - P.$$

Предполагая $P \neq 0$, поделим обе части этого равенства на P и найдём

$$P = \frac{(x - \frac{1}{q})2}{x(q - 1)} \quad (11.8)$$

При $q = 2$ формула (11.8) совпадает с (11.6). Это решение обращается в нуль при $x = \frac{1}{q}$. Отсюда следует, что $x_c = \frac{1}{q}$. Решение (11.8) имеет смысл при $q > 1$, $x > \frac{1}{q}$ и только при значениях x , очень близких к $\frac{1}{q}$. Поэтому в знаменателе выражения (11.8) можно положить $x = \frac{1}{q}$. Окончательно получаем

$$P(x) = \frac{(x - \frac{1}{q})2q}{q - 1}. \quad (11.9)$$

Выражение (11.9) описывает функцию $P(x)$ вблизи порога протекания.

11.3. Обсуждение результатов

При $q = 1$ функция $P(x) = 0$ при всех x в интервале $0 \leq x < 1$, и никаких других решений уравнение (11.5) в этом интервале не имеет. При $q = 1$ и $x = 1$ уравнение (11.5) удовлетворяется любым значением P .

При $q = 1$ решётка Бете превращается в линейную цепочку узлов. В такой цепочке сколь угодно малое количество чёрных узлов прерывает протекание по белым узлами. Поэтому естественно, что при всех x в интервале $0 \leq x < 1$ распространение слухов обязательно прерывается, т. е. $P(x) = 0$. Значение x_c для линейной цепочки равно единице.

Как было показано в предыдущем разделе, в общем случае $x_c = \frac{1}{q}$. Этот результат можно было предвидеть заранее. В рассматриваемой задаче каждый человек передаёт слух q своим знакомым. Среднее число людей категории C среди этих знакомых равно qx . Значит, после каждой передачи вместо одного источника информации в среднем возникает qx источников. Таким образом, величина qx является коэффициентом размножения. Для того чтобы процесс не прекращался, необходимо, чтобы коэффициент размножения был больше единицы. Отсюда следует, что критическая концентрация x_c получается из условия $qx_c = 1$, т. е. $x_c = \frac{1}{q}$.

Напомним, что аналогичным образом определяется условие, необходимое для поддержания цепной реакции деления урана. Процесс распространения слухов, по существу, тоже является цепной реакцией и описывается по той же схеме, что и ядерный взрыв.

Интересно сравнить полученное значение x_c с результатами решёточных задач в пространствах с большим числом измерений. Приближённое вычисление порога протекания задачи узлов производилось для так называемых гиперрешёток. Это решётки такого же типа, как квадратная и простая кубическая, но только в пространстве с большим числом измерений. Координационное число z (число ближайших соседей) для таких решёток определяется формулой $z = 2d$, где d — размерность пространства (при $d = 2$ получаем $z = 4$, а при $d = 3$ получаем $z = 6$). Вычисления порога протекания x_y производились при $d = 4, 5, 6$. Оказалось, что результаты хорошо описываются формулой

$$x_y = \left(1 + \frac{6,3}{d^2}\right) \frac{1}{z - 1} \quad (11.10)$$

При достаточно больших d вторым членом в скобке можно пренебречь, и $x_y = \frac{1}{z-1}$. Но для решётки Бете $z - 1 = q$. Ведь в каждый узел этой решётки одна связь входит и q связей из него выходит.

Отсюда видно, что порог протекания на решётке Бете ($x_c = \frac{1}{q}$) такой же, как на гиперрешётке с очень большим числом измерений. Таким образом, решётка Бете как бы соответствует бесконечномерному пространству.

Решётка Бете является единственной системой, для которой удалось точно найти вид $P(x)$ вблизи порога протекания. Оказалось (см. (11.9)), что в этом случае $P(x)$ обращается в нуль при $x \rightarrow x_c$ по линейному закону $P \sim (x - x_c)$. Как будет видно из дальнейшего, это специфическое свойство решётки Бете, а также всех решёток в пространстве с большим числом измерений.

Упражнение

1. Разберите задачу связей на решётке Бете. Все узлы считайте одинаковыми, а связи целыми и разорванными. Пусть доля целых связей равна x . Найдите функцию $P(x)$, определённую так же, как и раньше.

Глава 12.

Структура бесконечного кластера

12.1. Модель Шкловского — де Жена

Рассмотрим для определённости задачу узлов и допустим, что концентрация неблокированных узлов немного выше пороговой, так что существует бесконечный кластер. Он представляет собой бесконечные цепочки из связанных друг с другом узлов. Если соединить все связанные узлы бесконечного кластера отрезками прямых, то получится набор пересекающихся друг с другом ломаных линий (см. рис. 4.2, где показана одна такая линия).

Структурой бесконечного кластера называют его геометрию в масштабах гораздо больших, чем период решётки. В таких масштабах изломы, происходящие в отдельных узлах решётки, не воспринимаются глазом и цепочка представляется плавно изогнутой линией.

На рис. 12.1 изображён небольшой фрагмент бесконечного кластера. На концах A и B кластер не кончается — он уходит налево и направо на бесконечное расстояние. Введём теперь следующую классификацию точек и линий бесконечного кластера. Участки

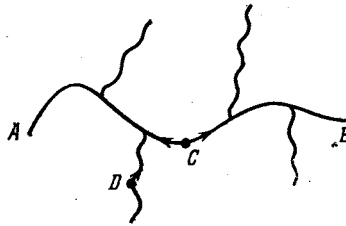


Рис. 12.1. Фрагмент бесконечного кластера с мёртвыми концами.

бесконечного кластера делятся на *скелет* и *мёртвые концы*.

Считается, что точка принадлежит скелету бесконечного кластера, если по крайней мере два пути, выходящие из неё в разные стороны, позволяют уйти на бесконечное расстояние. Такой точкой является, например, точка C на рис. 12.1. Из неё можно уйти на бесконечное расстояние, двинувшись и в правую, и в левую стороны. Если только один путь, выходящий из точки, ведёт на бесконечное расстояние, то эта точка принадлежит мёртвому концу. Например, из точки D на рис. 12.1 можно уйти на бесконечное расстояние, двигаясь только вверх. Движение вниз приводит в тупик. Поэтому считается, что точка D лежит на мёртвом конце.

Отбросим мысленно все мёртвые концы и постараемся представить, как встроен скелет бесконечного кластера.

Простейшая модель скелета была предложена независимо друг от друга советским физиком Б. И. Шкловским и французским физиком П. де Женом. Для плоской задачи эта модель представляет собой нечто вроде очень большой рыболовной сети, старой и изрядно потрёпанной. Она уже потеряла строгую периодичность, её верёвки не натянуты, некоторые узлы в ней порваны, другие съехали со своего места, но тем не менее «в среднем» это сеть

(рис. 12.2).

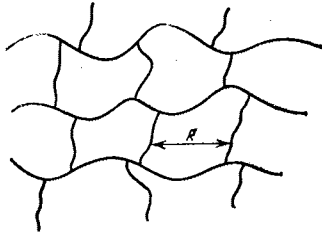


Рис. 12.2. Скелет бесконечного кластера.

Характерный линейный размер ячейки этой сети R называется *радиусом корреляции* бесконечного кластера. Он резко возрастает с приближением к порогу протекания:

$$R = \frac{l}{|x - x_c|^\nu}. \quad (12.1)$$

Здесь l — длина, равная по порядку величины периоду решётки, а ν — положительное число, которое называется *индексом радиуса корреляции*. Таким образом, по мере приближения к порогу протекания сетка становится всё более и более редкой.

Существование обращающегося в бесконечность радиуса корреляции является общим свойством всех критических явлений. То, что он обращается в бесконечность именно по степенному закону (12.1), не является строго доказанным, но лежит в основе современных представлений о критических явлениях и, по видимому, хорошо подтверждается экспериментальными данными.

Радиус корреляции имеет смысл и при $x < x_c$, т. е. ниже порога. В этой области он описывает *максимальный размер конечных кластеров*. Если $x \rightarrow x_c$ со стороны меньших значений ($x < x_c$),

то радиус корреляции тоже обращается в бесконечность по закону (12.1). Это означает, что при подходе к порогу протекания снизу конечные кластеры неограниченно увеличивают свои размеры и при $x = x_c$ сливаются в бесконечный кластер. Таким образом, зависимость $R(x)$ имеет вид, схематически показанный на рис. 12.3.

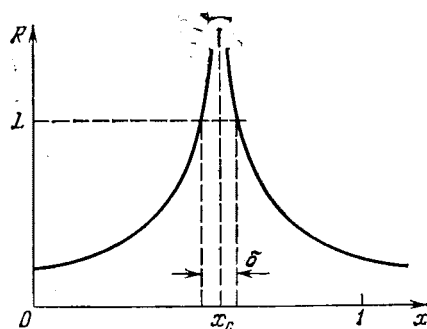


Рис. 12.3. Зависимость радиуса корреляции от x . Показана ширина критической области δ для квадрата $L \times L$ (см. следующий раздел).

В случае объёмных задач модель Шкловского — де Жена имеет аналогичный вид. Она похожа на сильно испорченный проволочный каркас трёхмерной решётки, причём средняя длина одной ячейки выражается формулой (12.1). Следует только иметь в виду, что численные значения индексов радиуса корреляции для плоских и объёмных задач разные.

Рассмотрим теперь, к каким следствиям приводит представление о сеточной структуре бесконечного кластера.

12.2. Роль размеров системы

В гл. 1, 2 и 3 подчёркивалось, что понятие порога протекания имеет смысл лишь в бесконечной системе. В конечной системе порог протекания меняется от образца к образцу, т. е. является величиной случайной. Однако, значения, которые принимает эта случайная величина, с подавляющей вероятностью попадают в некоторую область с шириной $\delta(\mathcal{N})$, которая называется *критической областью*. При увеличении числа узлов в системе \mathcal{N} ширина этой области уменьшается по степенному закону (см. формулу (1.8)), так что при $\mathcal{N} \rightarrow \infty$ порог протекания приобретает чёткий смысл, превращаясь из случайной величины в величину достоверную.

В начальных главах книги эти сведения сообщались практически без всякого вывода. Представление о радиусе корреляции позволяет их понять и получить формулу (1.8).

Для определённости обсудим эксперимент с экранной сеткой, имеющей размеры $L \times L$, схема которого изображена на рис. 1.1. Допустим, что сделано много опытов, использующих разные случайные последовательности блокируемых узлов, результатом которых явился набор порогов протекания. Напомним, что конфигурации блокированных узлов, полученные в разных опытах, совершенно не похожи друг на друга.

Удобный способ рассуждения состоит в следующем. Вообразим *бесконечную* экранную сетку с заданной долей x неблокированных узлов. Представим, что на разные участки этой сетки накладывается квадрат, имеющий размеры $L \times L$, и изучается

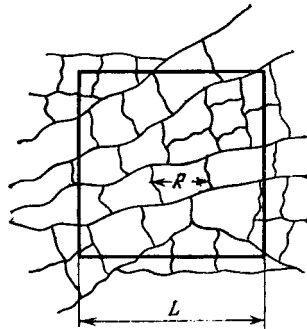


Рис. 12.4. Квадрат, наложенный на бесконечную сетку при $R \ll L$ и $x > x_c$

протекание левой стороны этого квадрата на правую по неблокированным узлам, оказавшимся внутри этого квадрата (рис. 12.4). Накладывая квадрат на разные участки бесконечной сетки, можно перебрать результаты разных опытов с конечной сеткой.

В бесконечной экранной сетке протекание возникает точно при $x = x_c$, но, как мы сейчас увидим, это совершенно не означает, что при $x > x_c$ обязательно есть протекание в квадрате $L \times L$.

При $x > x_c$ в бесконечной системе существует бесконечный кластер. Изобразим его скелет в виде рыболовной сети, показанной на рис. 12.2. Для дальнейшего крайне важно соотношение между радиусом корреляции R и длиной квадрата L . Примем сначала, что L значительно превосходит R . Тогда (рис. 12.4) внутри квадрата находится много ячеек сети бесконечного кластера, который обеспечивает протекание между сторонами квадрата. Эти ячейки могут иметь разные размеры, в сети бесконечного кластера могут быть большие дыры, но если в квадрате в среднем должно быть много ячеек, то вероятность того, что в кластере имеется дыра размером в целый квадрат, ничтожно мала. Поэтому делается следующий вывод:

<p>Если $x > x_c$, то порог протекания квадрата не может находиться в области значений x, удовлетворяющей сильному неравенству $L \gg R(x)$. Эта область должна быть выше порога.</p>	(I)
---	-----

Согласно формуле (12.1) при стремлении x к x_c радиус корреляции неограниченно возрастает и при каком-то значении x неизбежно сравнивается с L . Теперь о протекании внутри квадрата ничего определённого сказать нельзя. Всё зависит от конкретной конфигурации заблокированных узлов внутри него.

Пусть теперь $x < x_c$ и радиус корреляции значительно меньше, чем L . При $x < x_c$ радиус корреляции представляет собой максимальную длину конечных кластеров. Если $R \ll L$, то не существует такого кластера, который мог бы связать стороны квадрата. Поэтому делается ещё один определённый вывод:

<p>Если $x < x_c$, то порог протекания квадрата тоже не может находиться в области значений x, удовлетворяющей сильному неравенству $L \gg R(x)$. Эта область должна быть ниже порога.</p>	(II)
--	------

Если $x < x_c$, но величина x очень близка к x_c , то радиус корреляции становится больше, чем L . В этом случае о протекании в квадрате нельзя сказать ничего определённого. В бесконечной системе существуют конечные кластеры размера, большего, чем L , но внутри них есть дыры такого же размера, и всё зависит от конкретной конфигурации заблокированных узлов внутри квадрата.

Теперь можно оценить размер критической области, в которой могут находиться значения порога протекания квадрата $L \times L$. Согласно выводам I и II эта область должна определяться условием $L \lesssim R$. Как видно из рис. 12.3, чем больше L , тем уже эта область и тем теснее она прижата к порогу протекания для бесконечной системы. Ширина области δ определяется условием $R(\delta) = L$. С помощью формулы (12.1) получаем $l/\delta^\nu = L$ или

$$\delta = \left(\frac{l}{L} \right)^{\frac{1}{\nu}} \quad (12.2)$$

Внутри критической области, т. е. при $x - x_c \ll \delta$ пороги протекания квадратов с длиной L распределены однородно (см. рис. 2.4, где показана функция распределения порогов протекания). Точка $x = x_c$ внутри этой области ничем не выделена. Действительно, это точка, в которой наступает протекание в бесконечной системе. Но установить, есть такое протекание или нет, работая с квадратом конечного размера, невозможно. Если $L < R$, то, накладывая квадрат на разные участки бесконечной сетки, нельзя сказать, существуют в этой сетке только конечные кластеры или они уже слились и образуют бесконечный кластер. *Изучение протекания в квадрате конечного размера позволяет лишь определить ширину критической области.*

В этом разделе обсуждались лишь плоские задачи. Однако всё сказанное полностью переносится на задачи объёмные. Ширина критической области для объёмных задач также определяется формулой (12.1). Небольшая разница возникает, если выразить ширину δ не через размер системы L , а через полное число узлов

\mathcal{N} . Дело в том, что $\mathcal{N} = \left(\frac{l}{a}\right)^d$, где a — период решётки, а d — размерность пространства. Поэтому согласно (12.2)

$$\delta(\mathcal{N}) = \frac{C}{\mathcal{N}^{\frac{1}{\nu d}}}, \quad (12.3)$$

где C — численный коэффициент, который не может быть определён из столь простых соображений. В плоском случае ($d = 2$) формула (12.3) совпадает, с формулой (1.8). Именно с помощью этой формулы в результате исследования найденной на ЭВМ зависимости $\delta(\mathcal{N})$ был впервые определен индекс радиуса корреляции плоской задачи. Оказалось, что $\nu_2 = 1,33$. (Здесь и ниже число 2 показывает, что речь идет об индексе двумерной системы.) Для трёхмерных задач индекс ν иной: $\nu_3 = 0,8 \div 0,9$. (Число 3 означает, что индекс относится к трёхмерным задачам.)

12.3. Электропроводность вблизи порога протекания

Рассмотрим для определённости двумерные или трёхмерные сетки с заблокированными узлами. Как говорилось в начальных главах книги, электропроводность таких сеток отлична от нуля при $x > x_c$ и обращается в нуль на пороге протекания x_c . Экспериментальные данные, а также данные, полученные с помощью расчёта на ЭВМ, показывают, что удельная электропроводность сеток обращается в нуль по закону

$$\sigma(x) = \sigma_0(x - x_c)^t, \quad (12.4)$$

где множитель σ_0 по порядку величины равен удельной электропроводности сетки без заблокированных узлов. Величина t называ-

ется критическим индексом электропроводности и является предметом очень тщательного изучения, преимущественно с помощью расчетов на ЭВМ. (В одном из последних расчетов, например, использовалась квадратная сетка, имеющая 800×800 узлов.) Установлено, что для двумерных сеток $t_2 = 1,3$, а для трёхмерных $t_3 = 1,6 \div 1,7$.

Сеточная модель бесконечного кластера позволяет вывести формулу (12.4) и связать индекс t с индексом радиуса корреляции. Электрический ток течёт только по бесконечному кластеру, причём именно по его скелету. В мёртвых концах, прикреплённых к скелету лишь с одной стороны, тока нет. Если сделать электрический ток достаточно сильным, так чтобы проволока, по которой он течёт, светилась, то в темноте скелет бесконечного кластера можно наблюдать визуально: освещенные каналы на тёмном фоне. Вдали от порога вся сетка светится более или менее равномерно, вблизи порога расстояние между освещенными каналами увеличивается и, наконец, на самом пороге свечение совсем прекращается — ток через систему прервался.

Вычислим удельную электропроводность скелета бесконечного кластера. Следует иметь в виду, что это вычисление не может претендовать на правильный учёт численных множителей. Оно позволяет лишь получить зависимость σ от $x - x_c$. Эта зависимость не изменится, если мысленно заменить неправильную и нерегулярную сеть идеальной сеткой с периодом, равным R .

Рассмотрим сначала плоский случай (см. рис. 12.4). Удельное сопротивление равно сопротивлению квадрата с единичной длиной. Число проволочек, пересекающих этот квадрат, равно $\frac{1}{R}$, где

R — расстояние между проволочками, которое выражается формулой (12.1). Обозначим сопротивление одной проволочки, имеющей единичную длину, через ρ_0 . Все проволочки включены параллельно. Следовательно, удельное сопротивление

$$\rho = \frac{\rho_0}{\frac{1}{R}} = \rho_0 R,$$

а удельная электропроводность

$$\sigma = \rho^{-1} = \rho_0^{-1} R^{-1} \quad (12.5)$$

Подставляя (12.2), получим

$$\sigma = \sigma_2 (x - x_c)^\nu \quad (12.6)$$

где $\sigma_2 = \rho_0^{-1} l^{-1}$.

В трёхмерном случае нужно вычислить удельное сопротивление проволочного каркаса, изображающего, например, простую кубическую решётку с периодом R (от типа решётки зависит лишь численный коэффициент). Удельное сопротивление равно сопротивлению кубика с единичной длиной ребра. Число параллельно соединённых проволочек; проходящих через грань такого кубика, равно $\frac{1}{R^2}$. Поэтому удельное сопротивление

$$\rho = \frac{\rho_0}{\left(\frac{1}{R}\right)^2} = \rho_0 R^2, \quad (12.7)$$

а удельная электропроводность равна

$$\sigma = \rho_0^{-1} R^{-2} = \sigma_3 (x - x_c)^{2\nu}, \quad (12.8)$$

где $\sigma_3 = \rho_0^{-1} l^{-2}$.

Во избежание недоразумений следует обратить внимание на то, что удельная электропроводность σ в двумерном и трёхмерном случаях имеет разную размерность. В двумерном случае она измеряется в Ом^{-1} , а в трёхмерном — в $\text{Ом}^{-1} \cdot \text{см}^{-1}$.

Множители σ_2 и σ_3 порядку величины представляют собой удельные электропроводности двумерной и трёхмерной сеток без заблокированных узлов. Действительно, как видно из формул (12.5) и (12.8), удельная электропроводность $\sigma(x)$ превращается в σ_2 или в σ_3 при $R = l$, т. е. когда сетка бесконечного кластера совпадает с исходной сеткой, на которой ставится задача. Таким образом, величина σ_0 в формуле (12.4) в двумерном случае равна σ_2 , а в трёхмерном — σ_3 .

Сравнивая формулы (12.5) и (12.6) с формулой (12.4), получаем, что в двумерном случае $t = \nu$, а в трёхмерном $t = 2\nu$. Используя $\nu_2 = 1,3$ и $\nu_3 = 0,8 \div 0,9$, получим $t_2 = 1,3$, $t_3 = 1,6 \div 1,8$, что очень близко к приведенным выше данным. Это совпадение свидетельствует в пользу модели Шкловского — де Жена.

Упражнение

1. Модель скелета бесконечного кластера может быть обобщена следующим образом. Представим, что проволочки, образующие скелет, очень извилисты (рис. 12.5). Расстояние между точками их пересечения по-прежнему равно $R(x)$ и выражается формулой (12.1). Однако, если выпрямить отрезок проволочки, заключенный между двумя точками пересечений, то окажется, что его длина значительно превосходит R . Пред-

ставим эту длину \mathcal{L} в виде

$$\mathcal{L} = \frac{l}{(x - x_c)^\zeta} \quad (x > x_c),$$

где $\zeta > \nu$. При $x \rightarrow x_c$ отношение $\frac{\mathcal{L}}{R} = (x - x_c)^{-(\zeta - \nu)}$ стремится к бесконечности.

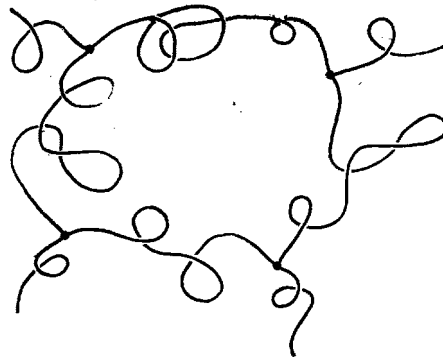


Рис. 12.5. Скелет бесконечного кластера с учётом «извилистости».

Длина \mathcal{L} имеет смысл только при условии $\mathcal{L} \gg R$, т. е. $\zeta > \nu$. Однако существует доказательство того, что $\zeta = 1$. Поэтому предлагаемое обобщение модели Шкловского – де Жена (на самом деле это обобщение было сделано самими авторами модели) разумно, лишь если $\nu < 1$. В двумерном случае $\nu = 1,3$, и потому никакой «извилистости» быть не должно. Но в трёхмерном случае $\nu < 1$, и есть основания думать, что «извилистость» скелет бесконечного кластера действительно существует.

Выразите критический индекс электропроводности t через индексы ζ , и ν в трёхмерном случае, используя обобщённую модель.

12.4. Функция $P(x)$ вблизи порога протекания. Роль мёртвых концов

Как и электропроводность, функция $P(x)$, представляющая долю узлов, принадлежащих бесконечному кластеру обращается в нуль при $x = x_c$. Исследования показали, что вблизи порога эта функция имеет вид

$$P(x) = D(x - x_c)^\beta, \quad (12.9)$$

где D — численный коэффициент порядка единицы, а β — еще один критический индекс. Установлено, что для двумерных задач $\beta_2 = 0,14$, а для трёхмерных $\beta_3 = 0,4$. Эти результаты получены главным образом с помощью ЭВМ.

В функцию $P(x)$ дают вклад все узлы бесконечного кластера — и принадлежащие скелету, и принадлежащие мёртвым концам. С помощью модели бесконечного кластера можно определить, каких узлов больше. Допустим сначала, что мёртвых концов совсем нет, и вычислим вклад в $P(x)$ от скелета бесконечного кластера.

В двумерном случае на каждую ячейку бесконечного кластера приходится порядка $\frac{R}{a}$ узлов, принадлежащих скелету, где a — период решётки (как и в предыдущем разделе, здесь делается оценка, не претендующая на установление численных коэффициентов). Площадь ячейки порядка R^2 и, следовательно, полное число всех узлов в ячейке порядка $\frac{R^2}{a^2}$. Отсюда следует, что доля узлов, принадлежащих скелету бесконечного кластера,

$$P_{\text{СК}}(x) \sim \frac{a}{R} \sim (x - x_c)^\nu \quad (12.10)$$

Здесь означает равенство по порядку величины (без учёта численных коэффициентов порядка единицы).

В трёхмерном случае на каждую ячейку бесконечного кластера тоже приходится $\frac{R}{a}$ узлов, принадлежащих скелету, но полное число узлов в ячейке порядка $(\frac{R}{a})$. Поэтому в трёхмерном случае

$$P_{\text{СК}}(x) \sim \left(\frac{a}{R}\right) \sim (x - x_c)^{2\nu}. \quad (12.11)$$

Сравнивая формулы (12.10) и (12.11) с формулами (12.6) и (12.8), получаем, что доля узлов, принадлежащих скелету бесконечного кластера, по порядку величины совпадает с функцией $\frac{\sigma(x)}{\sigma_0} = (x - x_c)^t$

Сравнивая (12.10) и (12.11) с формулой (12.9), видим, что

$$\frac{P_{\text{СК}}(x)}{P(x)} \sim (x - x_c)^{\nu_2 - \beta_2}$$

в двумерном случае, а в трёхмерном случае

$$\frac{P_{\text{СК}}(x)}{P(x)} \sim (x - x_c)^{2\nu_3 - \beta_3}$$

Вспомним, что $\nu_2 = 1,3$, а $\nu_3 \approx 0,9$. Следовательно, $\nu_2 - \beta_2 = 1,2$, а $2\nu_3 - \beta_3 \approx 1,4$. Таким образом, и в двумерном, и в трёхмерном случаях отношение $\frac{P_{\text{СК}}(x)}{P(x)}$ быстро стремится к нулю при $x \rightarrow x_c$. Это значит, что узлы, образующие скелет бесконечного кластера, составляют ничтожную долю от полного числа узлов, принадлежащих бесконечному кластеру. Основная «масса» бесконечного кластера сосредоточена в мёртвых концах и совершенно бесполезна с точки зрения электропроводности. Поэтому вблизи порога протекания $\frac{\sigma(x)}{\sigma_0} \ll P(x)$ (см. рис. 3.5). Однако именно

мёртвые концы определяют спонтанную намагниченность ферромагнетика с примесями вблизи порога протекания (см. гл. 3).

12.5. Универсальность критических индексов

Мы познакомились с тремя критическими индексами: ν , t и P , которые описывают поведение различных величин в окрестности порога протекания. Это поведение также называется критическим, так как функции $R(x)$, $\sigma(x)$ и $\beta(x)$ имеют при $x = x_c$ особые точки. Например, функция $R(x)$ обращается в бесконечность, у функции $P(x)$ первая производная в точке x_c имеет разрыв. С левой стороны она равна нулю, а с правой обращается в бесконечность. У функции $\sigma(x)$ имеет разрыв вторая производная. В теории протекания известно много других величин, имеющих критическое поведение, и, соответственно, много других критических индексов.

Для каждого из обсуждавшихся выше критических индексов приводились два значения: для двумерной и для трёхмерной задач. Но ведь существует великое множество и двумерных, и трёхмерных задач. Например, существуют трёхмерные задачи узлов, связей, сфер, задача об определении уровня протекания в случайном потенциале и многие другие. О каких же индексах шла речь? Теперь мы подходим, пожалуй, к самому интересному моменту в теории протекания. Согласно современным представлениям *критические индексы для всех задач в пространстве с одной и той же размерностью одинаковы*. (Среди задач, упомянутых в этой книге, исключение представляет только задача об ориентирован-

ном протекании.) Утверждение об универсальности критических индексов является скорее принятым, чем строго доказанным. Однако многочисленные проверки этого утверждения, выполненные с помощью ЭВМ, не смогли его опровергнуть.

В чем физические причины универсальности индексов? Видимо, в том, что индексы определяются структурой кластеров в окрестности порога протекания. Основную роль при этом играют геометрические свойства кластеров, проявляющиеся на больших расстояниях (порядка радиуса корреляции). Вблизи пороговой точки эти расстояния намного превосходят период решётки (в случае решёточных задач) или радиус сферы (в случае задачи сфер). Поэтому геометрия кластеров не зависит от того, на какой решётке задана задача. Задача может быть вообще не решёточной, а заданной на узлах, случайно расположенных в пространстве, и это тоже не повлияет на структуру больших кластеров. Но, разумеется, размерность пространства очень сильно сказывается на геометрии кластеров, так как обеспечить, например, «развязку» линий в трёхмерном пространстве гораздо легче, чем в двумерном. По этим причинам критические индексы не зависят от типа задачи, но зависят от размерности пространства.

Интересно заметить, что изменение критических индексов с увеличением размерности пространства происходит до размерности 6. Начиная с шестимерного пространства, индексы не меняются с увеличением размерности, причём индекс $\beta = 1$, как у решётки Бете. При $d \geq 6$ задача о критических индексах значительно упрощается и допускает точное решение.

Итак, в отличие от порогов протекания, которые существенно

зависят от типа задачи, критические индексы обладают определенной универсальностью. Отсюда следует простой вывод. Если результаты физического эксперимента трактуются с помощью теории протекания, а микроскопическая структура исследуемой системы не вполне ясна, то прежде всего следует сравнивать с теорией критические индексы, так как они почти ни от чего не зависят. Именно так и поступают при анализе экспериментальных данных по электропроводности в гетерогенных материалах (см. гл. 9).

Идея универсальности критических индексов заимствована теорией протекания у теории фазовых переходов второго рода (к фазовым переходам второго рода относятся, например, происходящие при повышении температуры переход металла из сверхпроводящего состояния в нормальное и переход ферромагнетика в неферромагнитное состояние). Вблизи точки фазового перехода второго рода так же, как вблизи порога протекания, образуются области большого размера, отличающиеся друг от друга своими свойствами. Разница состоит в том, что границы этих областей не «заморожены», как в теории протекания, а меняются со временем благодаря тепловому движению. Размер областей также называется радиусом корреляции и выражается формулой (12.1).

Из теории фазовых переходов пришла и другая важная идея — гипотеза подобия. В применении к теории протекания она состоит в следующем. Допустим, что сделаны фотографии различных участков сетки с заблокированными узлами при двух значениях a , равных x_1 и x_2 . Оба значения находятся по одну сторону от x_c и близки к x_c . Пусть, например, $x_1 - x_c > x_2 - x_c > 0$. При $x = x_1$ радиус корреляции $R_1 = l(x - x_c)^{-\nu}$, а при $x = x_2$ он равен

$R_2 = l(x - x_c)^{-\nu}$. По условию $R_1 < R_2$ Подобие состоит в том, что если фотографии, сделанные при $x = x_1$ увеличить в отношении $\frac{R_2}{R_1}$ то «в среднем» они ничем не будут отличаться от фотографий, сделанных при $x = x_2$. При этом фотографии предполагаются столь «мелкими», что на них не видны отдельные узлы и связи, а видны лишь большие блоки. Итак, гипотеза подобия заключается в том, что при приближении к порогу протекания крупномасштабная геометрия системы преобразуется подобным образом, причём все линейные размеры увеличиваются пропорционально радиусу корреляции.

Заметим, что модель Шкловского — де Жена удовлетворяет гипотезе подобия, однако гипотеза подобия гораздо шире. Она касается не только скелета бесконечного кластера и вообще не предполагает разбиения на скелет и мёртвые концы.

Математическая формулировка гипотезы подобия позволяет установить связь между критическими индексами ν , β и третьим индексом, который здесь не вводился. Расчёты показывают, что эта связь хорошо выполняется.

Представления о подобии, впервые введённые советскими физиками А. З. Паташинским и В. Л. Покровским, а также американским физиком Л. Кадановым, лежат в основе современной теории фазовых переходов и теории протекания.

С представлениями о подобии связан ряд новых математических методов вычисления критических индексы. Эти методы бурно развивались в последнее десятилетие и достигли определённого совершенства. В настоящее время почти все основные индексы вычислены с их помощью. Однако вычисления столь сложны, что их

трудно изложить в этой книге.

Глава 13.

Прыжковая электропроводность

В предыдущих главах было подробно описано, как применяется теория протекания для вычисления электропроводности систем, состоящих из случайно составленных элементов. При этом каждый элемент был либо проводящим, либо непроводящим, но сопротивление всех проводящих элементов предполагалось одинаковым. Примерами таких систем являются сетки с разорванными связями или блокированными узлами, смесь металлических и диэлектрических шариков и т. д.

Теперь мы обратимся к другому классу систем, которые тоже случайно составлены из различных элементов, но сопротивления этих элементов могут принимать любые значения, заключенные в очень широком интервале от самых малых до чудовищно больших. Оказывается, задача о вычислении сопротивления систем, состоящих из очень большого числа этих элементов, тоже решается с помощью теории протекания. В 1971 г. была построена теория прыжковой электропроводности полупроводников, в основу которой легли идеи протекания. В этой главе мы познакомимся с явлением прыжковой электропроводности и его математическим

описанием.

13.1. Механизм прыжковой электропроводности

Рассмотрим примесный полупроводник, содержащий, например, донорные примеси. Концентрацию примесей будем считать малой по сравнению с критической концентрацией N_c , при которой происходит моттовский переход металл — диэлектрик (см. гл. 8). При этих условиях перекрытие электронных оболочек соседних примесных атомов мало. Поэтому каждый донор можно рассматривать как изолированный водородоподобный атом, внешний электрон которого находится на расстоянии порядка a_B^* от ядра и имеет энергию связи порядка E_B^* (гл. 8). Допустим, что температура столь низка, что тепловой энергии колеблющихся атомов недостаточно, чтобы отрывать электроны от доноров. Как же осуществляется при этом электропроводность?

Представим, что часть доноров не имеет внешнего электрона. Обычно так получается вследствие компенсации примесей (гл. 10). Если полупроводник содержит не только донорные, но и акцепторные примеси, то каждый акцептор забирает электрон у донора. Если акцепторов меньше, чем доноров, то часть доноров сохранит внешний электрон, а другая часть останется без электрона и будет иметь положительный заряд.

Механизм прыжковой электропроводности состоит в том, что электрон «перепрыгивает» с одного донора на другой, который не имел до этого электрона.

Сейчас мы рассмотрим случай, когда потенциальная энергия

внешнего донорного электрона слабо зависит от положения донора в пространстве, т. е. от того, какова конфигурация заряженных примесей, окружающих данный донор. Этот случай соответствует малой концентрации примесей. В типичных для прыжковой проводимости условиях разброс энергий внешних электронов, находящихся на разных донорах, составляет примерно $\frac{1}{10}$ от энергии связи E_B^* .

В такой ситуации основное препятствие для электронного «прыжка» между донорами представляет притяжение электрона к тому донору, которому он принадлежит в начальном состоянии. Если рассуждать с позиций классической механики, то, для того чтобы перевести электрон с внешней орбиты одного донора на внешнюю орбиту другого донора, нужно совершить работу против сил притяжения, так как необходимо удалить электрон на половину расстояния между донорами. После этого электрон начнет притягиваться ко второму донору. Если доноры стоят редко, то эта работа мало отличается от работы по удалению электрона, принадлежащего изолированному донору, на бесконечность, т. е. от энергии связи E_B^* .

В чем же преимущество прыжковой электропроводности перед электропроводностью свободных электронов? Дело в том, что *прыжковая электропроводность — явление квантовое*. Квантовая механика разрешает электрону перейти с одного донора на другой, минуя свободное состояние и не занимая энергию у теплового движения атомов. Такой переход называется туннельным. Закон сохранения энергии остается в силе и для туннельных переходов. Он требует, чтобы энергия электронов в начальном и конечном

состояниях была одинаковой. Поэтому, если благодаря потенциалам окружающих примесей, энергии электрона на первом и втором доноре отличаются, недостающую энергию все равно придется занимать, а лишнюю отдавать. Но эта энергия в 10 раз меньше, чем E_B^* . Поэтому в низкотемпературных условиях прыжковая электропроводность выигрывает у электропроводности свободных электронов.

13.2. Сетка сопротивлений

Следует, однако, иметь в виду, что в рассматриваемых условиях туннельный переход — событие очень мало вероятное. Как говорилось в гл. 8, вероятность найти внешний электрон донора на расстоянии r от донорного ядра убывает с r по закону $e^{-\frac{2r}{a_B^*}}$, где $e \approx 2,72$. Если доноры 1 и 2 находятся друг от друга на расстоянии r_{12} , то вероятность найти внешний электрон донора 1 вблизи ядра донора 2 равна $e^{-\frac{2r_{12}}{a_B^*}}$. Эту величину и содержит вероятность туннельного перехода. Как следует из гл. 8, если концентрация доноров много меньше, чем концентрация N_c , соответствующая переходу диэлектрик — металл, то среднее расстояние между донорами гораздо больше, чем a_B^* , и, как правило, величина $e^{-\frac{2r_{12}}{a_B^*}} = \frac{1}{e^{\frac{2r_{12}}{a_B^*}}}$ очень мала.

Тем не менее между соседними донорами время от времени происходят туннельные переходы. Если к полупроводнику приложить электрическое поле, то окажется, что переходы в направлении

против поля (по направлению действующей силы) происходят чаще, чем переходы по полю. В результате возникает электрический ток, пропорциональный электрическому полю. В этом и состоит интересующее нас явление.

Для вычисления удельного сопротивления полупроводника используют модель, которая называется «сетка сопротивлений». Эта модель формулируется не на языке атомов и туннельных переходов, а на языке обыкновенных электротехнических сопротивлений. Представим, что между каждой парой доноров включено сопротивление. Сами доноры можно вообразить в виде крохотных металлических шариков, к которым припаяны проволочки от множества сопротивлений. Второй конец каждого из этих сопротивлений припаян к другому шарика. В результате получится нечто вроде беспорядочной трёхмерной сетки сопротивлений. Разумеется, если вы захотите на самом деле её собрать, придётся сильно увеличить масштабы системы. Ведь среднее расстояние между донорами порядка 10^{-5} см.

Сопротивления, включённые между донорами, должны быть вычислены, исходя из туннельного тока, который течёт между этими донорами при заданном электрическом поле. Это сводится к решению квантовомеханической задачи, которое мы не будем здесь приводить. Заметим только, что согласно сказанному выше, чем дальше доноры друг от друга, тем меньший ток переносится туннельными переходами между этими донорами при одном и том же поле.

В соответствии с этим, сопротивление \mathcal{R} , связывающее доноры, находящиеся на расстоянии r друг от друга, можно представить в

виде

$$\mathcal{R}(r) = \mathcal{R}_0 e^{\frac{2r_{12}}{a_B^*}}, \quad (13.1)$$

где величина \mathcal{R}_0 одинакова для всех сопротивлений (она может быть порядка 1 Ом). Задача состоит в том, чтобы найти удельное сопротивление системы, состоящей из очень большого (10^{16} – 10^{19}) числа доноров, случайно разбросанных в пространстве.

13.3. Свойства сетки сопротивлений

Главная черта обсуждаемой модели состоит в том, что сопротивления, определяемые формулой (13.1), чудовищно отличаются друг от друга. Среднее расстояние между донорами r_D связано с концентрацией доноров N_D соотношением $\frac{4}{3}\pi r_D^3 N_D = 1$, означающим, что в сфере радиуса r_D в среднем находится один донор. В условиях, в которых обычно наблюдается прыжковая электропроводность, величина r_D превышает a_B^* в 6–12 раз.

Допустим, например, что $r_D = 10a_B^*$. Тогда сопротивление, включенное между донорами, находящимися на расстоянии $1,5r_D$, относится к сопротивлению, включенному между донорами с расстоянием r_D , как $e^{\frac{r_D}{a_B^*}} = e^{10} = 2,2 \cdot 10^4$.

Доноры с взаимным расстоянием r_D встречаются почти так же часто, как и с расстоянием $1,5r_D$. Таким образом, незначительное изменение расстояния между донорами ведет к огромному изменению включенного между ними сопротивления.

В принципе модель сетки предполагает, что сопротивление включено между каждой парой доноров. Однако сопротивления,

включенные между удалёнными донорами, столь велики, что ими несомненно можно пренебрегать. Как правило, эти же доноры связаны цепочкой сопротивлений, идущей по донорам, являющимся ближайшими соседями друг друга. Хотя длина этой цепочки больше, чем кратчайшее расстояние между удалёнными донорами, сопротивление цепочки значительно меньше, чем одно сопротивление, включённое между этими донорами. Таковы уж свойства экспоненты: если $x_1 \gg 1$ и $x_2 \gg 1$, то $e^{x_1+x_2} = e^{x_1} \cdot e^{x_2} \gg e^{x_1} + e^{x_2}$. Таким образом, фактически в рассматриваемой сетке сопротивлений достаточно оставить сопротивления, включённые между четырьмя-пятью ближайшими соседями каждого донора.

13.4. Снова задача сфер

Перейдём теперь к вычислению сопротивления. Предлагается следующий способ рассуждений. Выключим (отпаяем) все сопротивления, которые должны быть между шариками, изображающими доноры, а затем будем их включать в определённой последовательности.

Включим сначала сопротивления, связывающие доноры, расстояние между которыми меньше, чем некоторая длина r' . Это значит, что включаются самые низкоомные (маленькие) сопротивления, величина которых меньше, чем $\mathcal{R}' = \mathcal{R}_0 e^{\frac{2r'}{a_B^*}}$. Если длина r' значительно меньше, чем среднее расстояние между донорами r_D , то будут включены лишь сопротивления между редкими, аномально близкими друг к другу донорами. Эти сопротивления, как правило, не связаны друг с другом и не могут обеспечить проте-

кание тока через систему.

Будем постепенно увеличивать длину r' , включая всё новые порции сопротивлений. Начиная с некоторого значения $r' = r_c$ включённые сопротивления образуют бесконечный кластер. При $r' > r_c$ система проводит ток и обладает конечным удельным сопротивлением.

Задача об определении r_c есть не что иное, как задача сфер (гл. 7). Действительно, построим вокруг каждого донора сферу радиуса r' . Тогда сопротивления будут включены только между такими донорами, один из которых находится внутри сферы, построенной около второго. Согласно результатам гл. 7 такие доноры впервые образуют бесконечный кластер при условии

$$B_c = \frac{4\pi}{3} N_D r_c^3 = 2,7 \pm 0,1,$$

откуда следует, что

$$r_c = (0,865 \pm 0,015) N_D^{-\frac{1}{3}} \approx 1,39 r_D. \quad (13.2)$$

13.5. Вычисление удельного сопротивления

Согласно сказанному в гл. 12, скелет бесконечного кластера можно вообразить в виде трёхмерной проволочной сетки (см. рис. 12.2) с расстоянием между узлами порядка радиуса корреляции. Для данной задачи каждую проволоку следует считать состоящей из множества шариков, связанных сопротивлениями, а

выражение (12.1) нужно переписать в виде

$$R(r') = r_D \left(\frac{r_D}{|r' - r_c|} \right)^\nu. \quad (13.3)$$

Выражение (13.3) обращается в бесконечность на пороге протекания, причём, согласно идее об универсальности, критический индекс ν такой же, как и для других трёхмерных задач теории протекания ($\nu = 0,8-0,9$). Вдали от порога, при $r' - r_c$ порядка r_D радиус корреляции становится порядка r_D .

При $r' = r_c$ расстояние между узлами сетки бесконечного кластера обращается в бесконечность. Бесконечный кластер в этой точке ещё имеет нулевую плотность. Однако при $r' > r_c$ он образует каналы, пронизывающие всю систему и обеспечивающие конечное удельное сопротивление.

Продолжим процедуру включения сопротивлений. Если включить сопротивления между донорами с расстояниями от r_c до $r_c + ga_B^*$, где g — некоторое число, меньшее, чем единица (например, $g = 0,2$), то максимальное из включенных сопротивлений практически не изменится, так как $e^{\frac{2r_c}{a_B^*+2g}} \approx e^{\frac{2r_c}{a_B^*}}$. С другой стороны, бесконечный кластер будет представлять сетку с расстоянием между узлами порядка

$$R_c = r_D \left(\frac{r_D}{ga_B^*} \right)^\nu. \quad (13.4)$$

Такая сетка называется в теории прыжковой проводимости *критической*. Именно она и ответственна за протекание тока.

Действительно, дальнейшее увеличение величины r' , т. е. включение остальных сопротивлений, не приведет к существенно-

му изменению удельного сопротивления. Как видно из формулы (13.1), сопротивления с $r - r_c \gg a_B^*$ очень велики по сравнению с сопротивлениями с $r = r_c$. Поэтому фактически через эти сопротивления ток всё равно не течёт, предпочитая критическую сетку, максимальные сопротивления которой равны

$$\mathcal{R}_{max} = \mathcal{R}_0 e^{-\frac{2r_c}{a_B^*}}.$$

Итак, нам осталось вычислять удельное сопротивление критической сетки. Подобно тому, как это делалось в гл. 12, представим эту сетку в виде проволочного каркаса, изображающего простую кубическую решетку (рис. 3.7) с периодом R_c .

Каждая проволочка, соединяющая два соседних узла решётки, состоит из большого числа последовательно соединённых исходных сопротивлений, определяемых формулой (13.1). Так как отдельные сопротивления чудовищно отличаются друг от друга, сопротивление одной проволочки следует считать равным максимальному из входящих в нее сопротивлений, т. е. \mathcal{R}_{max} . Элементарный кубик простой кубической решётки (рис. 6.1) состоит из четырёх параллельно соединённых проволочек, причём каждая проволочка принадлежит также четырём другим кубикам, так что на долю каждого кубика приходится одна проволочка с сопротивлением \mathcal{R}_{max} .

Для того чтобы получить удельное сопротивление, необходимо умножить полное сопротивление кубика \mathcal{R}_{max} на площадь его грани (R_c^2) и разделить на длину ребра (R_c). Таким образом,

$$\rho = \mathcal{R}_{max} R_c = \mathcal{R}_0 \cdot R_c e^{-\frac{2r_c}{a_B^*}}. \quad (13.5)$$

При этом мы обращаемся с кубиком так, как будто он заполнен однородным веществом с удельным сопротивлением ρ . Именно такой смысл и вкладывают в понятие удельного сопротивления системы, которая на самом деле сильно неоднородна.

13.6. Обсуждение результата

Необходимо иметь в виду, что различные части формулы (13.5) получены с разной степенью точности. Множитель, который называют предэкспоненциальным множителем, получен с точностью до численного коэффициента. Действительно, фактически мы, например, не знаем число g в выражении (13.4) для R_c . Кроме того, замена сетки бесконечного кластера периодической решёткой вполне может привести к ошибке в численном коэффициенте. Следует заметить, что в настоящее время не существует теории, позволяющей достоверно определить численный коэффициент в предэкспоненциальном множителе.

Что же касается величины, стоящей в показателе экспоненты формулы (13.5), то она известна с хорошей точностью. По смыслу вывода формулы (13.5) в показатель экспоненты вместо r_c можно было бы подставить $r_c + a_B^* g$.

Тогда вместо $\frac{2r_c}{a_B^*}$ возникла бы сумма $\frac{2r_c}{a_B^*} + 2g$. Второй член суммы фактически характеризует неопределённость, с которой нам известен показатель экспоненты. Как уже говорилось, при малых концентрациях доноров $\frac{2r_c}{a_B^*} \gg 1$, поэтому относительная погрешность мала. Фактически она отражает неопределённость численного коэффициента в предэкспоненциальном множителе. Именно

показателем экспоненты определяется зависимость удельного сопротивления ρ от концентрации доноров. Согласно (13.2), (13.5)

$$\ln \rho = \ln \mathcal{R}_0 \cdot R_c + \frac{2 \cdot 0,86}{N_D^{\frac{1}{3}} a_B^*}. \quad (13.6)$$

Первый член в правой части (13.6) зависит от N_D гораздо слабее, чем второй, и его можно считать константой. Сравнение формулы (13.6) с экспериментальными данными, выполненное для очень многих полупроводников, показало, что эта формула очень точно описывает зависимость $\ln \rho$ от N_D , что явилось важным достижением изложенной выше теории.

Метод, который описан в этой главе, с успехом применялся также для нахождения температурной зависимости прыжковой электропроводности, зависимости от внешнего магнитного поля и от ряда других параметров. Более того, этот метод применим к любой неоднородной системе, сопротивление которой меняется от точки к точке в широких пределах.

Глава 14.

Заключительная

В этой, последней, главе мы кратко обсудим некоторые приложения теории протекания, которые не попали в предыдущие главы, но представляются нам интересными, а также подведём итог всему сказанному, попытавшись определить, что общего между всеми рассмотренными задачами.

14.1. Некоторые приложения

Протекание жидкости по лабиринту. Проблема тесно связана с задачей о проникновении газа в угольную маску, положившей начало теории протекания. Представьте пористое тело, в которое под давлением нагнетается жидкость. Жидкость не смачивает тело, так что капиллярные силы препятствуют её проникновению в поры.

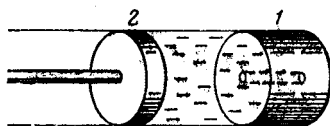


Рис. 14.1.

Лабиринт из капилляров пронизывает всё тело, но диаметры капилляров сильно отличаются друг от друга. В одних местах капилляры широкие, а в других — узкие.

Представим сначала, что в теле имеется лишь один цилиндрический капилляр. Слева от тела 1 находится жидкость при некотором давлении, создаваемом поршнем 2 а справа — воздух при атмосферном давлении (рис. 14.1). Тело закреплено и не может двигаться. Из-за поверхностного натяжения несмачивающая жидкость проникает в капилляр, только если её давление превышает атмосферное давление на величину $\frac{2\sigma}{R}$, где R — радиус капилляра, а σ — коэффициент поверхностного натяжения. При меньших давлениях жидкость образует выпуклый мениск, но не проходит через капилляр.

Теперь представим, что в теле существуют капилляры разных радиусов. При заданном значении давления жидкости некоторые из них (самые широкие) пропускают жидкость, а некоторые (более узкие) не пропускают. С увеличением давления число пропускающих жидкость капилляров увеличивается. При малых давлениях, когда пропускают лишь самые толстые капилляры, жидкость не может проникнуть в тело дальше приповерхностного слоя. Однако при некотором критическом давлении пропускающие капилляры образуют систему, пронизывающую всё тело. Начиная с этого давления, жидкость просачивается через тело.

Вычисление этого давления, а также ряда других характеристик процесса представляет очень важное с точки зрения практических нужд приложение теории протекания.

Образование полимерных гелей. Полимер — это макромолекула, состоящая из огромного числа элементарных единиц (мономеров). Находящиеся в растворе мономеры могут связываться друг с другом, образуя сложную трёхмерную сетку, пронизывающую всю систему. В результате образуется гель — твёрдообразная среда, напоминающая желе или студень.

Существует модель, удовлетворительно описывающая образование геля. Она сводится к задаче теории протекания, где белыми узлами являются молекулы мономера, а чёрными — молекулы растворителя. Связи между белыми узлами осуществляются с вероятностью, зависящей от температуры. Образование бесконечного кластера из связанных белых узлов соответствует появлению геля.

Такая задача теории протекания называется смешенной, поскольку случайными элементами являются и узлы, и связи. Пусть вероятность того, что некоторый узел является белым, равна x_1 (она равна концентрации молекул мономера), а вероятность того, что данная связь не разорвана, равна x_2 . Нужно найти область значений x_1 и x_2 , в которой существует бесконечный кластер из связанных друг с другом белых узлов.

По определению x_1 и x_2 изменяются в интервале от нуля до единицы.

Если $x_2 = 1$, т. е. все связи целые, то бесконечный кластер существует при всех x_1 в области $x_y \leq x_2 \leq 1$, где x_y — порог протекания задачи узлов. Если $x_1 = 1$, т. е. все узлы белые, то для существования бесконечного кластера необходимо условие $x_{св} \leq x_2 \leq 1$, где $x_{св}$ — порог протекания задачи связей.

Квадрат на рис. 14.2 включает всю область изменения переменных x_1 и x_2 . Сплошной кривой показан график функции $x_{min}(x_2)$, описывающей границу области существования бесконечного кластера. При каждом значении x_2 в интервале $x_{cв} < x_2 < 1$ функция $x_{min}(x_2)$ даёт минимальное значение x_1 , при котором существует бесконечный кластер. Легко понять, что $x_{min}(1) = x_y$, а $x_{min}(x_{cв}) = 1$. Сама область существования бесконечного кластера на рис. 14.2 заштрихована.

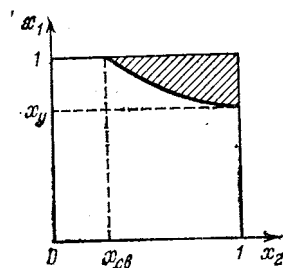


Рис. 14.2.

Если вероятность x_2 задана при всех значениях температуры, то с помощью функции $x_{min}(x_2)$ можно найти область температур и концентраций мономера, в которой образуется гель.

14.2. Что же такое теория протекания?

В этой книге нигде не фигурировало определение того, что называется теорией протекания. Дать такое определение не очень просто. Попробуем же теперь, на последней странице книги, сформулировать, что общего между всеми задачами, о которых шла речь, и что, собственно, является предметом теории протекания.

Теория протекания занимается связностью очень большого (макроскопического) числа элементов при условии, что связь каждого элемента со своими соседями носит случайный характер, но задается вполне определённым способом (например, с помощью генератора случайных чисел, обладающего конкретными свойствами).

Различные задачи теории протекания объединяются тем, что геометрия связанных элементов вблизи порога протекания у них одинакова. Для того чтобы это заметить, нужно отвлечься от мелкомасштабной структуры, определяемой характером связей и свойствами элементов, и следить только за связностью больших блоков. Универсальная крупномасштабная геометрия диктует универсальные свойства физических величин, зависящих от структуры больших кластеров. Это и объединяет столь не похожие друг на друга задачи теории протекания.

Ответы и решения

Глава 1

1. По общему правилу для нахождения среднего необходимо каждое возможное значение случайной величины умножить на вероятность этого значения и все произведения сложить. Вероятность того, что любая из граней кубика оказалась сверху, равна $\frac{1}{6}$. Следовательно,

$$\bar{a} = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = \frac{21}{6}.$$

2. Результаты отдельных опытов изменяются, но среднее значение $x_c(\mathcal{N})$, вычисленное с помощью многих опытов, остаётся тем же, так как движение слева направо в среднем осуществляется столь же вероятно, как и сверху вниз. Соответственно, не изменится и $x_c = \lim_{\mathcal{N} \rightarrow \infty} x_c(\mathcal{N})$.

3. Обозначим пороговые значения, полученные в опыте под номером i , через x'_i и x_i , причем x'_i соответствует новому определению порога, а x_i — старому (слева направо). Если по мере уменьшения x сначала исчезло протекание слева направо, а потом сверху вниз, то $x'_i = x_i$. Если же это случилось в обратной последовательности, то $x'_i > x_i$. Среднее значение порога при новом определении равно

$$x'_c(\mathcal{N}) = \frac{x'_1 + x'_2 + \dots + x'_Q}{Q},$$

а при старом определении

$$x_c(\mathcal{N}) = \frac{x_1 + x_2 + \dots + x_Q}{Q}.$$

В обеих формулах полное число опытов Q считается очень большим. Но при большом числе опытов ситуации, когда $x'_i > x_i$ обязательно возникнут. Поэтому $x'_c(\mathcal{N}) > x_c(\mathcal{N})$. Однако,

$$\lim_{\mathcal{N} \rightarrow \infty} x'_c(\mathcal{N}) = \lim_{\mathcal{N} \rightarrow \infty} x_c(\mathcal{N}) = x_c.$$

Дело в том, что порог протекания в бесконечной системе является величиной не случайной, а достоверной, которая не меняется от опыта к опыту. В то же время отличие порогов протекания в разных направлениях есть явление случайное. В среднем с точки зрения протекания все направления эквивалентны. Поэтому x_c не зависит от направления.

4. Решение аналогично предыдущему.

Обозначим пороговые значения, полученные в опыте под номером i , через x''_i и x_i , причем x''_i соответствует новому определению порога, а x_i — старому. Легко доказать, что $x''_i \leq x_i$. Рассуждая, как в предыдущей задаче, получим

$$x''_c(\mathcal{N}) < x_c(\mathcal{N}), \text{ но } \lim_{\mathcal{N} \rightarrow \infty} x''_c(\mathcal{N}) = \lim_{\mathcal{N} \rightarrow \infty} x_c(\mathcal{N}) = x_c.$$

5. Расчёт по формуле (1.8) дает $\delta = 0,01$. Это значит, что «типичные» отклонения от среднего значения равны $\pm 0,01$. Отсюда следует, что последний знак в числе 0,59, полученном Ватсоном и Лисом, с довольно большой вероятностью мог оказаться ошибочным. С гораздо меньшей вероятностью мог оказаться ошибочным первый знак. Поскольку был сделан только один опыт, сами авторы могли оценить лишь погрешность, с которой они определяли порог протекания при использованной ими последовательности блокируемых узлов. (Она оказалась $\pm 0,005$.) Однако, они ничего не могли сказать о том, как изменится результат при повторении опыта с другой случайной последовательностью. Более поздние исследования, в программу которых входило много опытов при одном значении \mathcal{N} , а также опыты с большими \mathcal{N} , привели к установлению формулы (1.8). Эти опыты также доказали, что, по-видимому, число 0,59 правильно даже в последнем знаке, что в значительной мере следует считать результатом везения.

Глава 2

1. Доля заблокированных узлов равна $1 - x = \frac{N'}{N}$, а доля неблокированных узлов равна $x = \frac{N - N'}{N}$. Если выбрать случайно Q узлов, то среди них окажется Qx неблокированных и $Q(1 - x)$ заблокированных (чем больше число Q , тем с большей точностью будет выполняться это соотношение). Поэтому вероятность того, что случайно выбранный узел окажется заблокированным, равна $\frac{Q(1-x)}{Q} = 1 - x$, а вероятность того, что он окажется неблокированным, равна $\frac{Qx}{Q} = x$. Так как узел может быть либо заблокированным, либо неблокированным, сумма вероятностей равна единице: $1 - x + x = 1$.
2. Вероятность любой последовательности из трёх фиксированных чисел равна $\frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}$. Число разных последовательностей, удовлетворяющих поставленным условиям для чисел 1, 2, 3, равно 6 (123; 213; 321; 231; 132; 312), а для чисел 1, 2, 2 равно 3 (122; 212; 221). Вероятность того, что осуществляется одна из возможных последовательностей, равна сумме вероятностей. Следовательно, в первом случае искомая вероятность равна $6 \cdot \frac{1}{216} = \frac{1}{36}$, а во втором случае $3 \cdot \frac{1}{216} = \frac{1}{72}$.
3. $(0,8)^3 \cdot (0,9)^4 = 0,336$.
4. Решение не приводится.

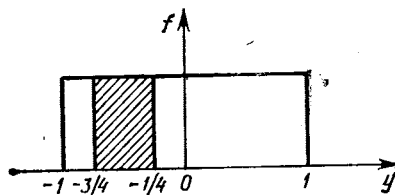


Рис. 0.1.

5. Функция распределения случайной величины a является постоянной величиной в интервале $(-1, 1)$ и равна нулю вне этого интервала (рис. 0.1). Так как полная площадь прямоугольника, ограниченная кривой $f(y)$ (в данном случае горизонтальной прямой), осью абсцисс и перпендикулярами, восстановленными в точках -1 и 1 , должна равняться единице, то $f(y) = \frac{1}{2}$ при $-1 < y < 1$.

По общему правилу искомая вероятность равна площади прямоугольника, ограниченного прямой $f(y)$, осью абсцисс и перпендикулярами, восстановленными в точках $y = -\frac{3}{4}$ и $y = -\frac{1}{4}$. (Этот прямоугольник на рис. О.1 заштрихован.) Вероятность равна $(-\frac{1}{4} - (-\frac{3}{4})) \cdot \frac{1}{2} = \frac{1}{4}$.

6. Переменная y пробегает любые значения от $-\infty$ до ∞ .

Поэтому в формулах (2.3) и (2.4) следует положить $A = -\infty$, $B = \infty$. Согласно формуле (2.3)

$$\bar{a} = \int_{-\infty}^{\infty} y f_{\mathcal{N}}(y) dy.$$

Подставляя формулу (2.6), получим

$$\bar{a} = \frac{1}{\sqrt{2\pi}} \delta_{\mathcal{N}}^{-1} \int_{-\infty}^{\infty} y \exp\left(-\frac{y^2}{2\delta_{\mathcal{N}}^2}\right) dy.$$

Под знаком интеграла стоит нечетная функция. Сделав замену переменной $y = -t$ и сравнив результат с исходной формулой, замечаем, что $\bar{a} = -\bar{a}$. Отсюда следует $\bar{a} = 0$.

Согласно формуле (2.4) дисперсия

$$\delta^2 = \int_{-\infty}^{\infty} y^2 f_{\mathcal{N}}(y) dy = \frac{1}{\sqrt{2\pi}} \delta_{\mathcal{N}}^{-1} \int_{-\infty}^{\infty} y^2 \exp\left(-\frac{y^2}{2\delta_{\mathcal{N}}^2}\right) dy.$$

Заменим переменную: $y = \sqrt{2}\delta_{\mathcal{N}}t$. Тогда

$$\delta^2 = \frac{2}{\sqrt{\pi}} \delta_{\mathcal{N}}^2 \int_{-\infty}^{\infty} t^2 e^{-t^2} dt.$$

Интеграл по t равен $\frac{\sqrt{\pi}}{2}$. Поэтому получаем $\delta^2 = \delta_{\mathcal{N}}^2$.

Глава 3

1. По определению $P(1) = 1$. При x , близких к единице, узлы могут не принадлежать к бесконечному кластеру по двум причинам:

- а) В них могут быть немагнитные атомы. Доля таких узлов равна $1 - x$.
- б) Магнитные атомы могут быть изолированы от бесконечного кластера, как, например, атом B на рис. 3.4. Но при x , близком к единице, когда немагнитных атомов мало, эта причина менее существенна, так как для такой изоляции требуется, чтобы несколько немагнитных атомов (4 в случае плоской решётки, показанной на рис. 3.4) собрались вокруг одного атома. При малом числе немагнитных атомов вероятность такого события мала. Поэтому вторую причину можно не принимать во внимание и считать, что доля атомов, принадлежащих бесконечному кластеру, равна просто доле магнитных атомов. Итак, при условии, что $1 - x \ll 1$, имеем $P(x) = x$.

2. В простой кубической решётке каждый атом имеет 6 ближайших соседей, расположенных от него в направлениях рёбер куба (рис. 3.7). Вероятность W_0 того, что все ближайшие соседи некоторого атома являются немагнитными атомами, равна произведению шести вероятностей: $W_0 = (1 - x)^6$. Вероятность W того, что хотя бы один из них является магнитным атомом, равна

$$W(x) = 1 - W_0 = 1 - (1 - x)^6.$$

Согласно формуле (3.2)

$$P_2(x) = xW(x) = x(1 - (1 - x)^6).$$

При $x \ll 1$

$$P_2(x) \approx 6x^2.$$

Легко сообразить, что для любой решётки, в которой каждый атом имеет z ближайших соседей,

$$P_2(x) = xW(x) = x(1 - (1 - x)^z),$$

а при $x \ll 1$

$$P_2(x) = zx^2.$$

3. На рис. 0.2 изображены 12 атомов, находящихся в окрестности атома 0. Все они могут принять участие в образовании кластера из трёх атомов. Такой кластер может быть образован, например, атомами 1, 0, 2, если все эти три атома окажутся магнитными. Вероятность этого события равна произведению трёх вероятностей $x \cdot x \cdot x = x^3$. Вероятность, что кластер образовали атомы 0, 4, 12 или любая другая тройка атомов, также равна x^3 .

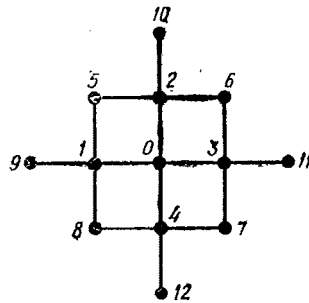


Рис. 0.2.

Прежде всего нужно ответить на вопрос, сколько таких троек существует. Сосчитаем сначала, сколько троек содержат атомы 0 и 1. Таких троек 6. Это 015; 018; 019; 103; 102; 104.

Теперь перейдем к тройкам, включающим атомы 0 и 3, но не включающим атом 1. Их пять: 036; 03 11; 037; 203; 304.

Аналогично имеется четыре тройки, включающие атомы 0 и 2, но не включающие атомы 1 и 3: 025; 026; 02 10; 204 и три тройки, включающие атомы 0 и 4, но не включающие 1, 2, 3. Это 047; 048; 04 12.

Итак, существует $6 + 5 + 4 + 3 = 18$ троек, причем вероятность каждой из них равна x^3 . Нужно найти вероятность того, что возникла хотя бы одна из них. При $x \ll 1$ с событиями, состоящими в том, что возникла одна из троек, можно обращаться как с событиями несовместимыми. Действительно, вероятность того, что образовались одновременно тройка 102 и

015, равна вероятности того, что четыре атома 0125 — магнитные, т. е. образовался кластер из четырёх атомов. Вероятность этого события равна $x^4 = x^3 \times x \ll x^3$. Итак, при $x \ll 1$ тройки можно считать несовместимыми. Тогда вероятность того, что образовалась хотя бы одна из них, равна сумме вероятностей, и

$$P_3 = 18x^3.$$

4. Вероятность того, что выбранный наугад атом принадлежит кластеру не менее, чем из двух атомов, можно представить в виде суммы вероятностей несовместимых событий:

$$P_2(x) = P_3(x) + \bar{P}_2(x), \quad (O.1)$$

где $\bar{P}_2(x)$ вероятность того, что атом принадлежит кластеру из двух атомов. Отсюда следует:

$$P_3(x) = P_2(x) - \bar{P}_2(x), \quad (O.2)$$

Функция $P_2(x)$ определяется формулой (3.3), и, следовательно, нужно только вычислить $\bar{P}_2(x)$.

Атом 0 (рис. O.2) может образовать кластер из двух атомов с атомами 1, 2, 3 или 4. Вероятность того, что кластер образован атомами 0 и 1, равна вероятности того, что оба эти атома магнитные, умноженной на вероятность того, что атомы 2, 3, 4, 8, 9 и 5 — немагнитные, т. е. равна $x^2(1-x)^6$. Точно такие же вероятности имеют события, состоящие в том, что кластер из двух атомов образован атомами 0 и 2, или 0 и 3, или 0 и 4. Все эти события являются несовместимыми, и поэтому вероятность $P_2(x)$ равна сумме четырёх вероятностей:

$$P_2(x) = 4x^2(1-x)^6. \quad (O.3)$$

Подставляя формулу (O.3) в формулу (O.2), получаем

$$P_3(x) = x(1 - (1-x)^4) - 4x^2(1-x)^6, \quad (O.4)$$

что и решает поставленную задачу.

Воспользовавшись формулой бинома, легко доказать, что выражение (0.4) не содержит членов, у которых степень была бы меньше, чем третья. Если $x \ll 1$, то $P_3(x) \approx 18x^3$, что совпадает с результатом предыдущего упражнения,

Глава 4

1. 0,0085; 0,0072; 0,0051; 0,0026; 0,0006; 00000; 0,0000 ... 0,0067; 0,0044; 0,0019; 0,0003; 0,0000; 0,0000 ... 0 0032; 0,001; 0,0001; 0,0000; 0,0000 ...
2. Число b является фактически n -значным. Поэтому $b < 10^n$. Для того чтобы получить следующее число b' , нужно найти b^2 , поделить на 10^n и взять целую часть. Таким образом, $b' < \frac{b^2}{10^n}$, но $\frac{b^2}{10^n} = b \cdot \frac{b}{10^n}$. Так как $\frac{b}{10^n} < 1$, то $\frac{b^2}{10^n} < b$. Отсюда следует $b' < b$, что и нужно было доказать.
3. 5, 15, 5, 15, 5, 15 ...
4. 5, 16, 9, 8, 5, 16, 9, 8, 5 ...
5. 5, 17, 13, 1, 5, 17, 13, 1, 5 ...
6. Всюду не выполнено условие v . В упражнении 3 $c = 0$, в упражнении 5 не выполнено также условие a и т. д.
7. Пусть $X_0 = 0$. Тогда получаем 0, 3, 1, 4, 2, 0, 3, 1, 4, 2 ... При любом значении X_0 число X_1 совпадает с одним из чисел этой последовательности.
8. Функция распределения случайных чисел, которые даёт генератор, изображена на рис. 2.3. Из этих чисел составлен массив V . Доля неблокированных узлов в массиве V равна доле случайных чисел в массиве V , удовлетворяющих неравенству $V < t$. Следовательно, средняя доля неблокированных узлов равна вероятности того, что случайное число окажется меньше, чем t . По определению функции распределения эта вероятность

равна площади, ограниченной кривой $f(y)$, осью абсцисс и перпендикулярами, восстановленными в точках 0 и t . В данном случае это площадь прямоугольника, равная t . Следовательно, средняя доля неблокированных узлов x равна t .

Глава 5

1. При значениях x , близких к единице, почти все узлы принадлежат бесконечному кластеру. Не принадлежат к нему лишь такие узлы, у которых оборваны все связи, соединяющие их с остальной системой. Вероятность того, что оборвана одна определённая связь, равна $1 - x$. В случае квадратной решётки для того, чтобы один узел оказался изолированным от системы, необходимо, чтобы были порваны четыре выходящие из него связи. (Рис. О.3, а) Вероятность этого события равна произведению вероятностей, т. е. равна $(1 - x)^4$. Для того чтобы были изолированы от системы два узла, нужно, чтобы было порвано шесть связей (рис. О.3, б). Вероятность этого события равна $(1 - x)^6$. При условии, что $1 - x \ll 1$, она значительно меньше вероятности того, что изолирован один узел. Таким образом, в интересующем нас предельном случае можно считать, что все изолированные узлы располагаются по одиночке и вероятность того, что выбранный наугад узел изолирован, равна $(1 - x)^4$. Вероятность того, что выбранный наугад узел не изолирован, равна $1 - (1 - x)^4$, т. е. для квадратной решётки

$$P^{\text{CB}}(x) = 1 - (1 - x)^4.$$

Рассуждая таким же образом, получим, что для треугольной решётки

$$P^{\text{CB}}(x) = 1 - (1 - x)^6,$$

а для шестиугольной решётки

$$P^{\text{CB}}(x) = 1 - (1 - x)^3.$$

Эти результаты справедливы, если $1 - x \ll 1$.

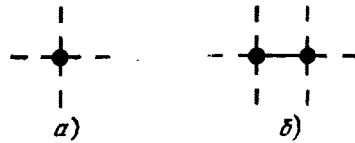


Рис. О.3. а) Один изолированный узел; б) два изолированных узла. Целая связь показана сплошной линией, а порванные связи — штриховыми.

2. Задача решается так же, как предыдущая. При условии $1 - x \ll 1$ почти все узлы принадлежат бесконечному кластеру. Выбранный наугад узел изолирован от бесконечного кластера, если все узлы, являющиеся его ближайшими соседями, содержат немагнитные атомы (для определённости мы пользуемся терминологией задачи о ферромагнетике). Как и в предыдущей задаче, вероятность того, что одним из соседей рассматриваемого узла является магнитный атом, но изолированный от бесконечного кластера, мала. Поэтому достаточно лишь вычислить вероятность того, что все соседи данного узла являются немагнитными атомами. Вероятность того, что в некотором узле находится немагнитный атом, равна $1 - x$. Число ближайших соседей равно числу связей, выходящих из данного узла. Поэтому результаты получаются такие же, как в предыдущем упражнении:

- для квадратной решетки $P^Y(x) = 1 - (1 - x)^4$;
- для треугольной $P^Y(x) = 1 - (1 - x)^6$;
- для шестиугольной $P^Y(x) = 1 - (1 - x)^3$.

Таким образом, при значениях x , близких к единице, $P^Y(x) = P^{CB}$, что не противоречит формуле (5.2).

3. Рассмотрим шестиугольную решётку с долей белых связей, равной x , и треугольную решётку с долей белых связей, равной y .

Напомним, что термин «белая связь» является синонимом термина «целая связь» а «чёрная связь» — это «разорванная связь», Термин узел «связан» с другим узлом, если это не оговорено специально, нужно понимать так, что он связан посредством неразорванных, т. е. белых связей.) Наложим

решётки друг на друга, как показано на рис. О.4. При этом узлы типа A , B , C являются общими для обеих решёток, а узлы типа D принадлежат только шестиугольной решётке. Идея дальнейших рассуждений состоит в том, что задачу о протекании на шестиугольной решётке можно свести к соответствующей задаче на треугольной решётке.

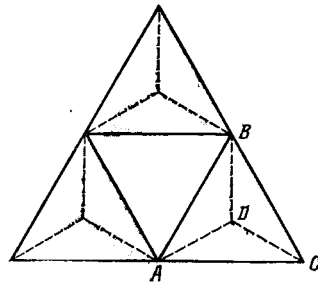


Рис. О.4. Преобразование звезды в треугольник. Штриховыми линиями показана шестиугольная решётка, а сплошными — треугольная.

Вероятности того, что узлы A , B и C связаны друг с другом, нужно выразить через величину x — долю белых связей шестиугольной решётки. При этом необходимо использовать геометрию и статистические свойства связей, выходящих из узлов типа D . После этого можно рассматривать только треугольную решётку, показанную на рис. О.4 сплошными линиями, забыв о том, что внутри каждого треугольника есть штриховые линии и узлы типа D .

Этот приём широко используется при расчёте электрических цепей и называется «преобразованием звезды в треугольник».

Фактически потребуются следующие величины:

- 1) W_{--} — вероятность того, что узел A не связан ни с узлом B , ни с узлом C . Эта вероятность равна сумме вероятностей двух несовместимых событий. Первое событие состоит в том, что связь AD — чёрная, а связи BD и DC — какие угодно. Его вероятность равна $1 - x$. Второе событие состоит в том, что связь AD — белая, а обе связи BD и DC — чёрные. Вероятность этого события равна произведению трех

вероятностей $x(1-x)(1-x)$. В результате

$$W_{--}(x) = 1 - x + (1 - x)^2. \quad (0.5)$$

- 2) W_{+-} — вероятность того, что узел A связан с B , но не связан с C . Она равна вероятности того, что связи AD и DB — белые, а связь DC — чёрная, и вычисляется как произведение вероятностей всех трех событий:

$$W_{+-} = x^2(1 - x). \quad (0.6)$$

- 3) W_{-+} — вероятность того, что узел A связан с C , но не связан с B . Легко убедиться, что она равна W_{+-} .
- 4) W_{++} — вероятность того, что узел A связан и с B , и с C , Она равна вероятности того, что все три связи AD , DB и DC — белые, и вычисляется как произведение вероятностей:

$$W_{++} = x^3. \quad (0.7)$$

Зная эти четыре вероятности, можно больше не возвращаться к шестиугольной решётке, а решать задачу протекания на треугольной решётке. Если бы её удалось решить, то было бы найдено критическое значение $x_{\text{св}}(\text{Ш})$ для шестиугольной решётки.

Решать задачу в таком виде, разумеется, несколько не легче. Однако, можно выразить те же самые вероятности W через величину y , представляющую долю белых связей на треугольной решётке. На пороге протекания эти вероятности имеют вполне определённое значение, которое пока неизвестно, но, приравняв вероятности, выраженные через x , к вероятностям, выраженным через y , можно получить соотношение, связывающее пороги шестиугольной ($x_{\text{св}}(\text{Ш})$) и треугольной ($x_{\text{св}}(\text{Т})$) решёток.

Итак, следующая задача состоит в том, чтобы выразить все четыре вероятности через y .

- 1) W_{--} . Узел A не связан ни с B , ни с C , если связи AB и AC — чёрные, а связь BC — какая угодно. Вероятность этого события равна

произведению вероятностей

$$W_{--}(y) = (1 - y)^2. \quad (O.8)$$

- 2) W_{+-} . Узел A связан с B , но не связан с C только, если связь AB — белая, а обе связи BC и AC — чёрные

$$W_{+-} = y(1 - y)^2. \quad (O.9)$$

(Если бы, например, связь BC была белой, то узел A был бы связан с C по пути ABC .)

- 3) $W_{-+} = W_{+-}$, как и в предыдущем случае.
- 4) W_{++} . Вероятность того, что узел A связан и с B , и с C , равна сумме вероятностей четырёх несовместимых событий. Первое событие состоит в том, что все три связи AB , BC и AC — белые. Его вероятность равна y^3 . Три других события состоят в том, что чёрной является лишь одна из трёх связей. Например, если чёрной связью является AB , то узел A связан с C через белую связь AC и связан с B по пути ACB . Вероятность каждого из трёх событий равна $y^2(1 - y)$. В результате получаем, что

$$W_{++}(y) = y^3 + 3y^2(1 - y). \quad (O.10)$$

На пороге протекания все $W(x)$ должны равняться $W(y)$. Поэтому возникает система уравнений:

$$W_{--}(x) = W_{--}(y); \quad 1 - x + x(1 - x)^2 = (1 - y)^2, \quad (O.11)$$

$$W_{+-}(x) = W_{+-}(y); \quad x^2(1 - x) = y(1 - y)^2, \quad (O.12)$$

$$W_{++}(x) = W_{++}(y); \quad x^3 = y^3 + 3y^2(1 - y). \quad (O.13)$$

Эти уравнения должны удовлетворяться при подстановке $x = x_{\text{св}}(\text{Ш})$ и $y = x_{\text{св}}(\text{Т})$. Кроме того, пороги протекания удовлетворяют соотношению (5.23). Из него следует, что $x_{\text{св}}(\text{Ш}) = 1 - x_{\text{св}}(\text{Т})$. Подставим в уравнения (O.11), (O.12), (O.13)

$$y = x_{\text{св}}(\text{Т}), \quad x = 1 - x_{\text{св}}(\text{Т}).$$

Уравнение (О.12) превращается при этом в тождество

$$(1 - x_{\text{СВ}}(T))^2 x_{\text{СВ}}(T) = x_{\text{СВ}}(T)(1 - x_{\text{СВ}}(T))^2,$$

а уравнения (О.11) и (О.13) приводятся к одному и тому же кубическому уравнению:

$$x_{\text{СВ}}^3(T) - 3x_{\text{СВ}}(T) + 1 = 0.$$

В интервале $0 \leq x_{\text{СВ}}(T) \leq 1$ это уравнение имеет единственный корень

$$x_{\text{СВ}}(T) = 2 \sin \frac{\pi}{18} \approx 0,347296.$$

Соответственно

$$x_{\text{СВ}}(\text{Ш}) = 1 - x_{\text{СВ}}(T) = 1 - 2 \sin \frac{\pi}{18} \approx 0,652704.$$

4. Прежде всего нужно найти площадь, приходящуюся на один узел, в каждой из трёх решёток, изображённых на рис. 5.1, при условии, что расстояние между ближайшими друг к другу узлами равно a .

К в а д р а т н а я р е ш е т к а. Каждому квадрату принадлежит 4 узла, но каждый из этих узлов принадлежит четырём разным квадратам. Следовательно, на долю каждого квадрата остаётся один узел, или, иными словами, каждому узлу принадлежит площадь одного квадрата, т. е. $S(\text{К}) = a^2$.

Т р е у г о л ь н а я р е ш е т к а. Каждому треугольнику принадлежит 3 узла, но каждый из этих узлов является собственностью шести разных треугольников. Таким образом, на один треугольник приходится половина узла, а на один узел приходится площадь, равная двум площадям треугольника. Площадь равностороннего треугольника со стороной a равна $\frac{\sqrt{3}a^2}{4}$. Следовательно, одному узлу принадлежит площадь $S(\text{Т}) = \frac{\sqrt{3}a^2}{2}$.

Ш е с т и у г о л ь н а я р е ш е т к а. Каждому шестиугольнику принадлежит 6 узлов, но каждый из этих узлов принадлежит трём разным шестиугольникам. Следовательно, на один шестиугольник приходится два узла. Площадь шестиугольника равна площади шести равносторонних треугольников со стороной a , т. е. равна $\frac{3\sqrt{3}a^2}{2}$. Следовательно, на каждый узел приходится площадь $S(\text{Ш}) = \frac{3\sqrt{3}a^2}{4}$.

Для каждой из решёток величина a определяется с помощью заданной функции $a(x)$, причём в качестве x нужно использовать порог протекания $x_{\text{св}}$ для соответствующей решётки. В результате получим

$$S(\text{T}) = \frac{\sqrt{3}}{2}(0,35a)^2; S(\text{K}) = (0,5a)^2;$$

$$S(\text{Ш}) = \frac{3\sqrt{3}}{4}(0,65a)^2.$$

Ясно, что функция $a(x)$ убывает с ростом x . (Если пары легче заражают друг друга, значит расстояние между ними уменьшается.) Легко увидеть, однако, что этого утверждения недостаточно, чтобы получить хотя бы одно неравенство между площадями, написанными выше. Действительно, у шестиугольной решётки длина a минимальна, но зато численный коэффициент в выражении для площади максимален, а у треугольной решётки — наоборот. Поэтому, для того чтобы найти решётку с минимальной площадью, надо знать функцию $a(x)$ более детально.

Глава 6

1. Вычисление коэффициентов заполнения для плоских решёток легко делается с помощью результатов упр. 4 к гл. 5. Это было продемонстрировано в тексте на примере шестиугольной решётки. Поэтому мы ограничимся здесь лишь объёмными решётками, причём всего лишь двумя, предоставив читателю сделать всё остальное самостоятельно.

Простая кубическая решётка. Каждому элементарному кубику (рис. 3.7) принадлежит 8 узлов, но каждый из этих узлов является собственностью 8 различных элементарных кубиков. Следовательно, на каждый кубик приходится один узел, а объём, приходящийся на каждый узел, равен объёму элементарного кубика, т. е. a^3 . Сферы, описанные вокруг узлов, имеют радиус $\frac{a}{2}$ и объём $\frac{4\pi a^3}{24}$. Доля объёма, занимаемая сферами, равна отношению объёма одной сферы к объёму, приходящемуся на

один узел. Следовательно, для ПК-решётки

$$f(\text{ПК}) = \frac{4\pi}{24} \approx 0,52.$$

Объемноцентрированная кубическая решетка. Каждому элементарному кубику (рис. 6.2, б) принадлежит 9 узлов, из которых 8 находятся по углам и один в центре. Узел, находящийся в центре, принадлежит только одному кубику, а каждый из угловых узлов принадлежит 8 разным кубикам. Следовательно, на один элементарный кубик приходится два узла, а объем, приходящийся на один узел, равен половине объема кубика, т. е. $\frac{a^3}{2}$. Ближайший сосед каждого узла находится от него на расстоянии, равном половине диагонали кубика, т. е. на расстоянии $\frac{\sqrt{3}a}{2}$. Радиус сферы, описанной вокруг каждого узла, равен $\frac{\sqrt{3}a}{4}$, а объем сферы равен $\frac{4\pi}{3} \left(\frac{\sqrt{3}a}{4}\right)^3 = \frac{\pi\sqrt{3}}{16} a^3$. Коэффициент заполнения равен отношению объема сферы к объёму, приходящемуся на один узел. Таким образом,

$$f(\text{ОЦК}) = \frac{\pi\sqrt{3}}{8} \approx 0,68.$$

Глава 7

1. К первой координационной группе узла 0 относятся 12 узлов типа 1 (рис. 0.5), ко второй — 6 узлов типа 2, к третьей — 24 узла типа 3 (на рисунке показаны не все узлы); $12 + 6 + 24 = 42$.

Глава 8

1. Пользуясь формулами (8.3) и (8.5), получим $a_{\text{Б}}^* = 636 \cdot 10^{-8} \text{ см} = 636 \text{ \AA}$, $N_c = 7,8 \cdot 10^{13} \text{ см}^{-3}$. Это уникально низкая критическая концентрация. Переход к металлической проводимости происходит, когда один примесный атом приходится на 10^8 атомов полупроводника! Для того чтобы получить полупроводник с такой концентрацией примесей, требуется очень сложная техника очистки кристалла.

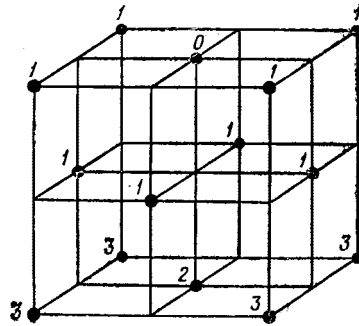


Рис. 0.5. Соседи узла 0 в ГЦК-решётке.

Глава 11

1. Введем функцию $Q(x)$ — вероятность того, что информация, попавшая в выбранный наугад узел, достигнет лишь конечное число других узлов. Как и раньше, $P(x) = 1 - Q$. Из каждого узла выходит q независимых каналов, по которым распространяется информация. Найдём вероятность того, что один такой канал прерывается на некотором этапе. Это может быть результатом одного из двух несовместимых событий: a — первая связь этого канала оказалась разорванной, b — первая связь оказалась целой, но узел, в который она ведёт, способен передать информацию лишь конечному числу других узлов. Вероятность события a равна $1 - x$, а вероятность события b равна $xQ(x)$. Вероятности несовместимых событий можно складывать. Поэтому вероятность того, что прервётся один канал, равна $1 - x + xQ(x)$. Так как все каналы независимы, то вероятность того, что все они прервутся, равна $(1 - x + xQ(x))^q$.

Отсюда получается уравнение для $Q(x)$:

$$Q(x) = (1 - x + xQ(x))^q.$$

Подстановкой $Q' = 1 - x + xQ$ оно сводится к уравнению (11.4). Окончательно получаем

$$P = \frac{\frac{x-1}{q} 2q^2}{q-1}.$$

Заметим, что пороги протекания для задачи узлов и задачи связей на решётке Бете одинаковы ($x_c = \frac{1}{q}$). Это можно было бы сказать заранее. Действительно, предположим, что узел, в который веди разорванная связь, заблокирован. После этого можно допустить, что все связи целые, и мы вернёмся от задачи связей к задаче узлов. Отсюда следует, что пороги протекания у этих задач должны быть одинаковыми.

Глава 12

1. Нужно вычислить сопротивление куба с единичной длиной ребра. Число параллельно соединённых проволочек по-прежнему равно $\frac{1}{r^2}$, но длина каждой проволочки уже не равна единице. Она больше единицы в число раз, равное отношению $\frac{\zeta}{R} = (x - x_c)^{-(\zeta - \nu)}$. Соответственно, сопротивление одной проволочки равно не ρ_0 , а $\rho_0(x - x_c)^{-(\zeta - \nu)}$. Поэтому, для того чтобы получить требуемый результат, достаточно заменить в формуле (12.6) величину ρ_0 , фигурирующую в выражении для σ_3 , на $\rho_0(x - x_c)^{-(\zeta - \nu)}$. В результате получим

$$\sigma = \sigma_3(x - x_c)^{\zeta + \nu},$$

где $\sigma_3 = \rho_0^{-1}t^{-2}$. Таким образом, $t = \zeta + \nu$.

Оглавление

Предисловие	1
I Задача узлов	7
1. Порог протекания	9
1.1. Два учёных мужа кромсают экранную сетку	9
1.2. Что такое случайная величина?	12
1.3. Среднее значение и дисперсия	13
1.4. Зачем нужна большая сетка?	18
2. Основные правила расчёта вероятностей	23
2.1. События и их вероятности	23
2.2. Сложение вероятностей	25
2.3. Умножение вероятностей	28
2.4. Порог протекания в сетке 2×2	32
2.5. Непрерывная случайная величина	35
2.6. Порог протекания как непрерывная случайная величина	38
3. Бесконечный кластер	43
3.1. Постоянный магнит	43
3.2. Ферромагнетик с примесями	47
3.3. Появление бесконечного кластера	50
3.4. Снова задача узлов	54
3.5. Кластеры при низкой концентрации	57

4. Решение методом Монте-Карло	63
4.1. Почему Монте-Карло?	63
4.2. Что такое метод Монте-Карло?	65
4.3. Как придумать случайное число?	69
4.4. Метод середины квадрата	71
4.5. Линейный конгруэнтный метод	73
4.6. Определение порога протекания	76
4.7. Поиск путей протекания	79
4.8. Определение порога	81
II Различные задачи теории протекания и их применения	85
5. Задачи на плоских решётках	87
5.1. Мы сажаем фруктовый сад (задача связей)	87
5.2. Неравенство, связывающее $x_{св}$ и x_u	92
5.3. Покрывающие и включающие решётки	95
5.4. Белое протекание и чёрное протекание	102
5.5. Дуальные решётки	107
5.6. Результаты для плоских решёток	112
5.7. Ориентированное протекание	114
6. Приближённые оценки порогов	119
6.1. Объёмные решётки	120
6.2. Пороги протекания для объёмных решёток	124
6.3. От чего зависит порог протекания задачи связей?	125
6.4. Как оценить порог протекания задачи узлов?	127
7. Ферромагнетик с дальнодействием и задача сфер.	133
7.1. Ферромагнетик с дальнодействием	134
7.2. Задача окружностей (сфер)	137
7.3. Предельный случай задачи узлов	142

8. Электропроводность примесных п.п.	145
8.1. Собственные полупроводники	145
8.2. Примесные полупроводники	149
8.3. Переход к металлической электропроводности	156
8.4. Переход Мотта и задача сфер	159
9. Различные обобщения задачи сфер	165
9.1. Охватывающие фигуры произвольной формы	165
9.2. Задача эллипсоидов	167
9.3. Другие поверхности	171
9.4. Ещё один эксперимент на домашней кухне	173
10. Уровень протекания	177
10.1. «Всемирный потоп»	177
10.2. Построение случайной функции**	180
10.3. Аналогия с задачей узлов**	182
10.4. Уровни протекания в плоской в трёхмерной задачах**	184
10.5. Компенсация примесей в полупроводниках	187
10.6. Движение частицы при наличии потенциальной энергии	188
10.7. Движение электрона в поле примесей	190
III Критическое поведение различных величин вблизи порога протекания и геометрия бесконечного кластера	193
11. Решётка Бете**	197
11.1. Слухи	197
11.2. Решение задачи узлов на решётке Бете	201
11.3. Обсуждение результатов	205
12. Структура бесконечного кластера	209
12.1. Модель Шкловского — де Жена	209
12.2. Роль размеров системы	213
12.3. Электропроводность вблизи порога протекания	217

12.4. Роль мёртвых концов	222
12.5. Универсальность критических индексов	224
13. Прыжковая электропроводность	229
13.1. Механизм прыжковой электропроводности	230
13.2. Сетка сопротивлений	232
13.3. Свойства сетки сопротивлений	234
13.4. Снова задача сфер	235
13.5. Вычисление удельного сопротивления	236
13.6. Обсуждение результата	239
14. Заключительная	241
14.1. Некоторые приложения	241
14.2. Что же такое теория протекания?	244
Ответы и решения	247
Глава 1	247
Глава 2	249
Глава 3	251
Глава 4	254
Глава 5	255
Глава 6	261
Глава 7	262
Глава 8	262
Глава 11	263
Глава 12	264