

ОЦЕНКА ТОЧНОСТИ МЕТОДА ОПРЕДЕЛЕНИЯ РАЗДЕЛЬНОЙ КОНЦЕНТРАЦИИ ПРИМЕСЕЙ ИЗ ИЗМЕРЕНИЙ ПОСТОЯННОЙ ХОЛЛА

Банная В. Ф., Веселова Л. И., Гершензон Е. М.,
Гусинский Э. Н., Литвак-Горская Л. Б.

На примере p -Si(B, Ga) с различной степенью компенсации проведена сравнительная оценка точности определения раздельной концентрации примесей по температурной зависимости концентрации дырок $p(T)$ в случае одной и двух легирующих примесей с энергиями ионизации, различающимися менее чем в 2 раза. Исследована функция среднеквадратичного отклонения в пространстве параметров $D(N_k, N_2)$ (N_k, N_1 и N_2 — концентрации компенсирующих примесей бора и галлия соответственно, $N_2 \gg N_1$) в предположении, что N_2 , энергии B и Ga известны. Показано, что в случае двух легирующих примесей $D(N_k, N_1)$ в окрестностях минимума имеет «овражный» рельеф и при некоторых соотношениях между N_k и N_1 разброс искомых величин превышает порядок, причем увеличение точности измерений $p(T)$ существенного улучшения в вычислении параметров не дает. При одной легирующей примеси точность вычисления параметров высокая.

Введение. Метод определения концентрации примесей в полупроводниковых материалах по температурной зависимости постоянной Холла является одним из наиболее распространенных. Он дает надежные результаты в случае, когда материал содержит одну основную примесь или несколько примесей, резко различающихся энергиями. Однако при определении раздельной концентрации примесей в полупроводниках, содержащих, например, даже два вида основных примесей с энергиями ионизации, различающимися незначительно, возникают трудности. Ошибка в определении параметров может получиться неприемлемо большой, хотя, как и в предыдущем случае, из решения уравнения нейтральности получается строгое аналитическое выражение для температурной зависимости концентрации свободных носителей (см., например, [1]).

Рассмотрим для определенности два конкретных случая, часто реализующихся на практике: кремний, легированный бором (энергия ионизации $E_1 = 4.5 \cdot 10^{-2}$ эВ), и кремний, легированный галлием ($E_2 = 7.6 \cdot 10^{-2}$ эВ) и содержащий остаточный бор (в обоих случаях существуют и компенсирующие примеси). В первом случае уравнение нейтральности запишется как

$$p + N_k = \frac{N_1}{1 + \frac{p\beta}{N_v} \exp \epsilon_1}, \quad (1)$$

где p — концентрация свободных дырок, N_1 и N_k — концентрации основной и компенсирующей примесей соответственно, β — фактор вырождения примесного уровня, N_v — плотность состояний в валентной зоне, $\epsilon_1 = E_1/kT$. Во втором случае уравнение нейтральности имеет вид

$$p + N_k = \frac{N_1}{1 + \frac{p\beta}{N_v} \exp \epsilon_1} + \frac{N_2}{1 + \frac{p\beta}{N_v} \exp \epsilon_2}, \quad (2)$$

где N_2 — концентрация галлия, $\epsilon_2 = E_2/kT$. Будем считать, что выполняется наиболее часто встречающееся на практике соотношение $N_2 \gg N_1$. Это, в частности, позволяет легко определить N_2 по области истощения примеси.

Настоящая работа посвящена анализу предельных возможностей и сравнительной оценке точности определения раздельной концентрации примесей из измерений постоянной Холла $R_H(T)$ в двух упомянутых случаях.

М а т е м а т и ч е с к о е м о д е л и р о в а н и е

С точки зрения математической обработки результатов эксперимента определение параметров полупроводниковых материалов есть задача определения параметров функции, описывающей получаемую в эксперименте зависимость, — аналитической модели. Зависимость $p(T, N_1, N_k)$, полученную из (1), будем называть «модель 1», а полученную из (2) — «модель 2». Из (1) для $p(T)$ получается квадратное уравнение, а из (2) — кубическое. Заранее не ясно, почему повышение степени уравнения должно приводить к понижению точности. Нужно сказать, что вопрос о точности определения параметров в случае нелинейных моделей теоретически еще очень мало разработан. Однако оценку точности можно произвести, исследуя вид функции среднеквадратичного отклонения в пространстве параметров. В нашем случае область определения этой функции может быть ограничена двумя переменными N_1 и N_k , так как N_2 , E_1 и E_2 считаются неизменяемыми. Тогда функция среднеквадратичного отклонения

$$D = \sqrt{\sum_{i=1}^Q [p(T_i) - p_i]^2 / Q}$$

(Q — число точек суммирования) описывает некоторую поверхность. Здесь $\{p_i, T_i\}$ — данные эксперимента, $p(T)$ — совокупность значений, определяемых моделью. Координаты минимума этой поверхности дают набор параметров, оптимальным образом описывающий результаты эксперимента, а поведение D в окрестностях минимума оказывает решающее влияние на точность определения параметров. Если D в окрестностях минимума меняется быстро, причем приблизительно одинаково быстро по всем направлениям («котловинный» рельеф), точность определения параметров будет высокой. Возможны, однако, случаи, когда поверхность D в окрестности минимума вытянута вдоль какого-то одного направления («овражный» тип рельефа). В предельном случае истинного оврага минимум достигается не в единственной точке, а на множестве точек. Решением задачи в этом случае пришлось бы признать множество значений параметров, описывающее линию дна оврага.

Таким образом, точность определения параметров зависит от вида поверхности D , а вид поверхности связан с особенностями аналитической модели. Исследуем поведение D в окрестностях минимума для моделей 1 и 2 и сравним результаты. Будем получать совокупность значений $\{p_i, T_i\}$ из аналитической модели при заданных значениях N_1, N_k . Перебирая значения N_1 и N_k , построим поверхность D .

Представление о рельефе поверхности в окрестностях минимума можно получить, рассекая ее плоскостями, перпендикулярными координатной плоскости N_1ON_k и проходящими через точку минимума. След пересечения данной плоскости к поверхности D показывает, насколько быстро меняется D в данном направлении. В результате такого исследования было установлено, что для модели 1 рельеф близок к котловинному, причем практически не меняется при изменении самих координат минимума. На рис. 1, а для этой модели представлены характерные следы сечения поверхности D плоскостями: в направлении $N_k/N_1 = \text{const}$ (штриховая кривая) и перпендикулярном ему (сплошная). Отметим, что для совмещения минимумов, имеющих разные координаты (соответствующие разным параметрам N_1 и N_k), на всех рисунках масштабы по оси абсцисс выбраны в относительных единицах (т. е. в единицах $N_k/N_{a, \text{min}}$). Для модели 2 имеются направления, вдоль которых поверхность сильно вытянута. Как показывают расчеты, при $N_1 > N_k$ такое направление определяется условием $N_k/N_1 = \text{const}$. При $N_1 < N_k$ направление наиболее медленного изменения D определяется условием $N_k - N_1 = \text{const}$ (рис. 1, б, в). В направлениях, перпендикулярных указанным, D меняется очень быстро, так что в целом для поверхности D обнаруживается овражный тип рельефа. На рисунке видно, что овраг не является истинным, так как минимум достигается в единственной точке: его координаты соответствуют тем значениям N_1 и N_k , при которых про-

изводилась имитация эксперимента, а само минимальное значение D равно нулю.

На рисунке видно, что скорость изменения D в окрестностях минимума сильно меняется в зависимости от величины N_k/N_1 при фиксированном направлении $N_k/N_1 = \text{const}$ (при $N_1 > N_k$, рис. 1, б) и от $N_1 - N_k$ (при $N_1 < N_k$, рис. 1, а). При $N_1 > N_k$ минимум выражен достаточно хорошо при малых (0.1, кривая 2) и больших (0.8, кривая 5) значениях N_k/N_1 . При промежуточных значениях (например, 0.2, кривая 3) вытянутость весьма значительная. При $N_1 < N_k$ приемлемый рельеф получается только при близких друг к другу

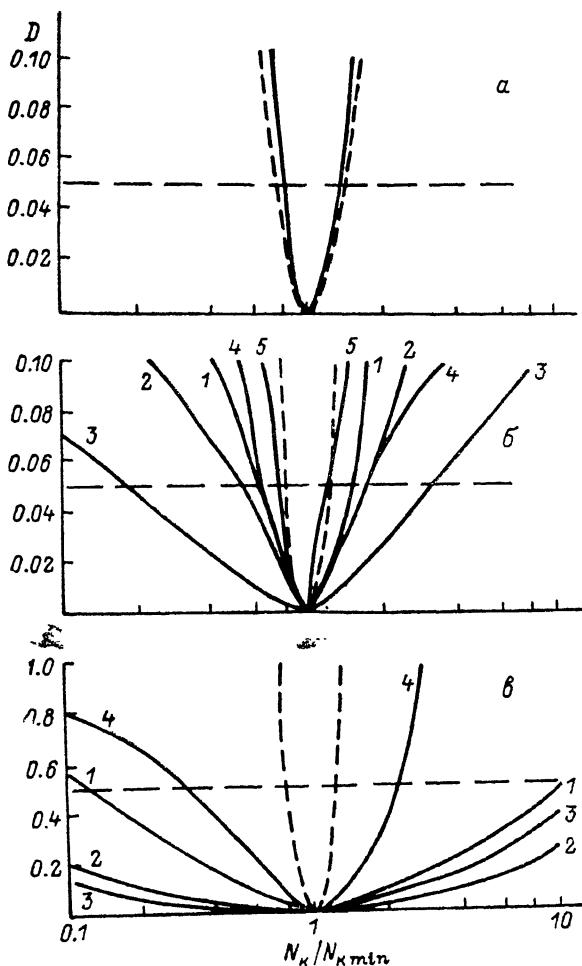


Рис. 1. Следы сечений поверхности $D(N_1, N_k)$ плоскостями в направлениях наибольшего (штриховые кривые) и медленного (сплошные) изменений.

а) Si(B), модель 1 б) Si(B, Ga), модель 2, $N_1 > N_k$. Номера кривых соответствуют разным значениям $K=N_k/N_1$: 1 — 0.05, 2 — 0.1, 3 — 0.2, 4 — 0.5, 5 — 0.8. в) Si(B, Ga), модель 2, $N_1 < N_k$. Номера кривых соответствуют разным значениям $1/K$: 1 — 0.1, 2 — 0.2, 3 — 0.5, 4 — 1.0.

значениях N_1 и N_k . Таким образом, скорость изменения D в окрестности минимума для модели 2 зависит от соотношения N_1 , N_2 и N_k .

Для сравнения оценки точности определения параметров по моделям 1, 2 предположим, что условием окончания поиска координат минимума является $D \leq 0.05$. Тогда разброс искомых значений можно охарактеризовать величиной области пересечений плоскости $D=0.05$ (она параллельна координатной плоскости) с поверхностью $D(N_1, N_k)$, след ее — прямые линии на рис. 1, а, б. Как видно из рисунка, для модели 1 это дает весьма незначительный разброс относительных значений ($\sim 0.9-1.05$). Для модели 2 ситуация сложнее. При $N_1 > N_k$ точность определения параметров сильно зависит от степени компенсации $K=N_k/N_1$. В случае слабой ($K \leq 0.1$) и сильной ($K \geq 0.5$) компенсаций

разброс значений $N_1/N_{K \min}$ составляет ≤ 3 . Наиболее «неблагоприятным» является случай средних компенсаций. Так, при $K=0.2$ $0.15 \leq N_K/N_{K \min} \leq 3$, т. е. разброс составляет более 1 порядка величины. Уменьшение дисперсии приводит к сокращению интервала значений $N_K/N_{K \min}$ для всех компенсаций. Однако для $K=0.2$ и $D=0.01$ разброс $N_K/N_{K \min} \approx 2$ (рис. 1, б).

При $N_1 < N_K$ и $D=0.05$ практически для всех значений K (кроме $K=1$) разброс параметров составляет несколько порядков, при $K=0.2$ вообще можно указать только верхний предел значений N_1 и N_K , причем уменьшение D улучшает точность определения параметров только при $K \approx 1$ (рис. 1, б).

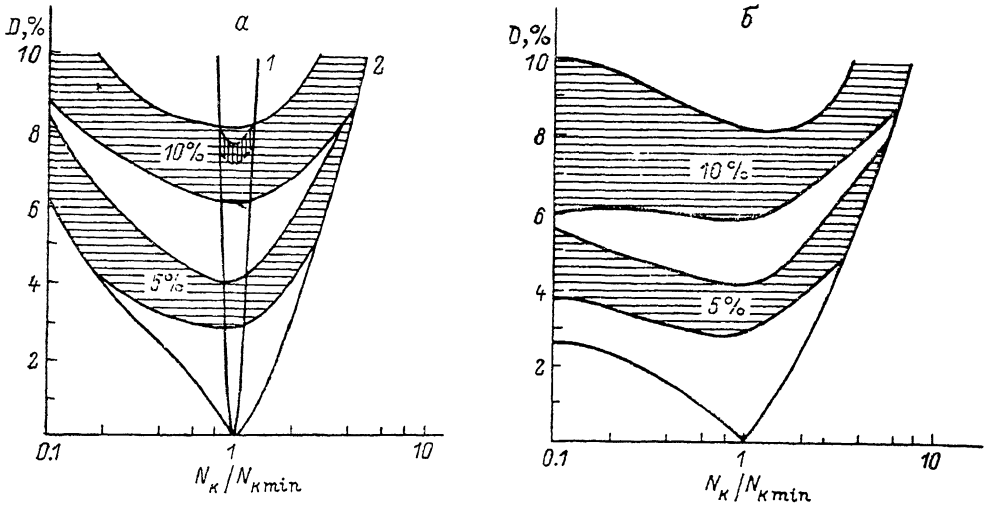


Рис. 2. Следы сечений поверхности $D(N_1, N_K)$ плоскостями в направлениях наиболее медленного (сплошные кривые) изменения.

Заштрихованные области — совокупность следов аналогичных сечений для разных случайных разбросов (p_i, T_i). а) 1 — Si(B, $N_K=5 \cdot 10^{18}$, $N_1=10^{18}$ см $^{-3}$); 2 — Si(B, Ga), $N_1 > N_K$, $N_2=10^{18}$ см $^{-3}$ ($N_K=10^{18}$, $N_1=4 \cdot 10^{18}$ см $^{-3}$); б) Si(Ga, B), $N_K > N_1$, $N_2=10^{18}$ см $^{-3}$ ($N_K=10^{18}$, $N_1=2 \cdot 10^{18}$ см $^{-3}$).

Оценим теперь влияние точности проведения эксперимента на точность определения параметров. Повторим описанную выше процедуру расчета, предварительно многократно подвергнув значения p_i , получаемые по модели 1, случайному разбросу. Таким образом, имитируются многократные измерения с той или иной степенью точности одного и того же полупроводникового образца. Заштрихованные области на рис. 2 включают в себя совокупность следов, полученных для разных случайных наборов $\{p_i, T_i\}$. Как видно, координаты минимума для разных наборов при 5%-м разбросе значений p_i , соответствующих модели 1, меняются очень незначительно (рис. 2, а). Для модели 2 при $N_1 > N_K$ и таком же разбросе значений p_i координаты минимума различаются сильнее (рис. 2, б), а при 10%-м разбросе могут различаться более чем на порядок. При $N_1 > N_K$ уже при 5%-м разбросе определить параметры практически невозможно (рис. 2, б).

Обсуждение результатов. Сопоставление с экспериментом

Таким образом, точность определения параметров для модели 1 существенно выше, чем для модели 2. Рассмотрим, какими особенностями моделей объясняется это различие.

Для модели 1 при низких температурах $p \ll N_K$ квадратное уравнение относительно $p(T)$ переходит в линейное с единственным параметром N_1/N_K . Дополнительную информацию для определения раздельной концентрации примесей дают измерения при высоких температурах.

Для модели 2 при низких температурах и $N_1 > N_K$ можно считать, что глубокий уровень полностью нейтрализован. В этих условиях, как и для модели 1, из измерения $p(x)$ определяется отношение N_1/N_K . При $N_1 < N_K$ основной мел-

кий уровень полностью оголен, и приближенное решение содержит параметр $(N_k - N_1)/2$. В области высоких температур, учитывая сделанные предположения, можно определить N_2 . Если проводить измерения только в области низких и высоких температур, легко видеть, что раздельная концентрация примесей определена быть не может: при $N_1 > N_k$ будут определены N_2 и N_1/N_k , а при $N_1 < N_k$ — N_2 и $N_k - N_1$. Это означает, что вдоль соответствующих направлений существуют истинные овраги. Отсюда следует, что план проведения эксперимента должен предусматривать необходимость иметь большое число экспериментальных точек в промежуточной области температур, где сравнимы вклады всех членов в уравнение нейтральности. Отметим, что ширина и положение этой области существенно зависят от параметров материала.

Измерения $R_H(T)$ проводились на большом числе образцов Si⟨B⟩ с $N_1 \approx 10^{14} \div 10^{17} \text{ см}^{-3}$ и Si⟨Ga⟩, содержащего остаточный бор, с $N_2 \approx 5 \cdot 10^{15} \div 10^{17} \text{ см}^{-3}$ по стандартной холловской методике в широком диапазоне темпера-

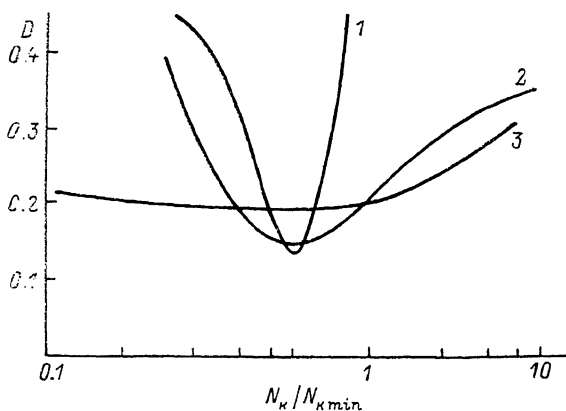


Рис. 3. Следы сечений поверхности $D(N_1, N_k)$ плоскостями в направлениях наиболее медленного ее изменения для трех измеренных образцов.

1 — Si⟨B⟩, $N_1 = 10^{16}$, $N_k = 5 \cdot 10^{12} \text{ см}^{-3}$; 2 — Si⟨Ga, B⟩, $N_1 = 4 \cdot 10^{12}$, $N_k = 2.4 \cdot 10^{12}$, $N_2 = 8 \cdot 10^{16} \text{ см}^{-3}$; 3 — Si⟨Ga, B⟩, $N_1 = 10^{14}$, $N_k = 1.2 \cdot 10^{14}$, $N_2 = 7.8 \cdot 10^{16} \text{ см}^{-3}$. Значения раздельной концентрации примесей в приведенных образцах получены из результатов холловских измерений.

тур ($T = 20 \div 300 \text{ K}$). Обработка результатов выполнялась на ЭВМ по формулам (1), (2), при этом E_1 , E_2 , N_2 и N_0/β считались известными.

При обработке результатов измерения реальных образцов получены данные, хорошо согласующиеся со сказанным выше. На рис. 3 представлены следы сечений поверхности $D(N_1, N_k)$ плоскостями в направлениях наиболее медленного ее изменения для трех образцов. В случае Si⟨B⟩ (кривая 1) параметры определяются уверенно, зависимость $D(N_1, N_k)$ имеет выраженный котловинный рельеф. Для Si⟨Ga⟩, содержащего бор, в случае $N_1 > N_k$ удается получить приемлемые результаты (кривая 2); в случае $N_1 < N_k$ (кривая 3) получаются значения $D(N_1, N_k)$, не содержащие выраженного минимума. Любые параметры, полученные для подобных образцов, будут случайными.

Таким образом, проведенное рассмотрение показывает, что определение раздельной концентрации примесей в полупроводниковых материалах, для которых справедлива модель 2, только по результатам измерения температурной зависимости постоянной Холла может наталкиваться на неустраняемые трудности.

О степени надежности получаемых при обработке результатов можно судить, исследуя форму поверхности среднеквадратичного отклонения в окрестностях минимума. При получении выраженного овражного рельефа следует, вероятно, использовать независимые методы для оценки одного из параметров N_1 или N_k . Один из таких методов предложен, например, в [2], другим можно считать способ оценки концентрации ионизированных центров в легированных и слабо компенсированных материалах по величине подвижности [3]. Так, для образца 2 этот способ дает $N_k = 5 \cdot 10^{12}$, а для образца 3 $N_k = 4 \cdot 10^{14} \text{ см}^{-3}$. Испол-

зую эти данные в совокупности с результатами холловских измерений, можно рассчитать отдельную концентрацию примесей в таких материалах.

Заметим, что для случая сложного легирования нами рассмотрена наиболее простая из возможных задач, когда число легирующих примесей ограничено двумя и энергии ионизации их известны. В противном случае задача усложняется.

Представляется, что проведенная оценка ограниченности стандартных холловских измерений для определения параметров полупроводникового материала даже в очень простой ситуации должна привлечь внимание: учет полученных результатов особенно важен при создании автоматизированных методов измерений, в программу которых обязательно должны быть заложены измерения зависимостей, дополняющих $R_H(T)$.

Список литературы

- [1] Блекмор Дж. Статистика электронов в полупроводниках. М., 1964. 392 с.
- [2] Веденеев А. С., Воронкова Г. И., Ждан А. Г., Коган Ш. М., Лифшиц Т. М., Рыльков В. В. // ФТП. 1988. Т. 22. В. 4. С. 586—592.
- [3] Банная В. Ф., Веселова Л. И., Гершензон Е. М. // ФТП. 1989. Т. 23. В. 2. С. 338—345.

Московский государственный
пединститут им. В. И. Ленина

Получена 4.05.1990
Прпнята к печати 17.08.1990