

МЕХАНИЗМ ГЕНЕРАЦИИ ЗАРЯДА В МДП СТРУКТУРЕ

Безлюдный С. В., Карпов В. Г., Колесников Н. В., Якименко А. Н.

Рассмотрен механизм туннельно-активационной генерации заряда в МДП структуре — активация электронов из валентной зоны на уровни поверхностных состояний и последующее туннелирование. Генерационный ток определяется редкими областями, в которых флуктуации встроенного заряда создают сильные электрические поля, уменьшающие туннельный барьер. Теория предсказывает температурную зависимость энергии активации, флуктуации генерационных токов от образца к образцу, возможность подавления генерации ионизирующим облучением. Результаты измерений времен генерации и их зависимостей от дозы облучения находятся в качественном согласии с теорией.

В работе предложена модель генерации заряда в МДП структуре. Рассмотрение процесса генерации использует представления о проколах в $p-n$ -переходах, развитые Райхом и Рузиным [1]. Некоторое отличие состоит в том, что мы учли флуктуации поверхностной плотности заряда. Модель не содержит специфических предположений о свойствах МДП структуры или дефектов в ней. Для модели существенно наличие поверхностных состояний и изгиб зон, вызванный обедняющим напряжением, как это изображено на рис. 1. Мы рассматриваем генерационные процессы, происходящие в два этапа: активация электрона из валентной зоны на уровень поверхностного состояния и последующее туннелирование (рис. 1). В результате электрон покидает структуру, оставляя избыточный положительный заряд.

Рассматриваемый механизм может реализовываться как непосредственно под электродом, так и в прилегающей поверхностной области. В первом из этих случаев пространственная координата $x=z$ (рис. 1). Во втором случае координата $x=r-R$ равна радиальному расстоянию до подэлектродной области на границе раздела диэлектрик—полупроводник.

В предлагаемом механизме важную роль играют флуктуации электрического поля. Их существование обусловлено флуктуациями в расположении заряженных примесей в диэлектрике и в полупроводнике. В области повышенного поля энергетический барьер, под которым происходит туннелирование, имеет ширину меньше среднего значения. Такие флуктуационные области, обладающие экспоненциально повышенной туннельной прозрачностью, мы, следуя [1], будем называть проколами.

Чем больше отличие прозрачности прокола от средней величины, тем реже встречается такой прокол. Существуют оптимальные проколы, для которых оптимальным образом сочетаются требования высокой прозрачности и высокой вероятности существования. Они реализуются в редких областях высокого поля и потому их относительная концентрация мала. Если площадь МДП структуры достаточно велика, то именно оптимальные проколы определяют величину генерационного тока. Для МДП структур малой площади вероятность найти оптимальные проколы мала и генерационный ток определяется наиболее прозрачными из имеющихся неоптимальных проколов. Критерий больших и малых площадей будет дан далее.

Мы рассмотрим флуктуации, создаваемые встроенным зарядом в тонком слое диэлектрика. Это эквивалентно флуктуациям плотности заряда на поверхности полупроводника. Рассмотрение произведем в два этапа. На первом

этапе зададимся энергией ε барьера, под которым происходит туннелирование, определим параметры оптимальных проколов и соответствующий генерационный ток. На втором этапе учтем необходимость термоактивации, предшествующей туннелированию. Условие равенства активационного и туннельного токов позволит определить оптимальную энергию ε_T барьера, зависящую от температуры T .

Итак, мы рассматриваем структуру, в которой на фоне средней поверхностной плотности $q\sigma_0$ заряда реализуются флуктуации $q\delta\sigma$. Соответствующие флуктуации электрической силы можно оценить как

$$\delta F = \frac{q^2}{\varkappa} \delta\sigma, \quad (1)$$

где q — заряд электрона, \varkappa — диэлектрическая постоянная. Здесь и далее мы опускаем в формулах все численные множители порядка единицы, на опре-

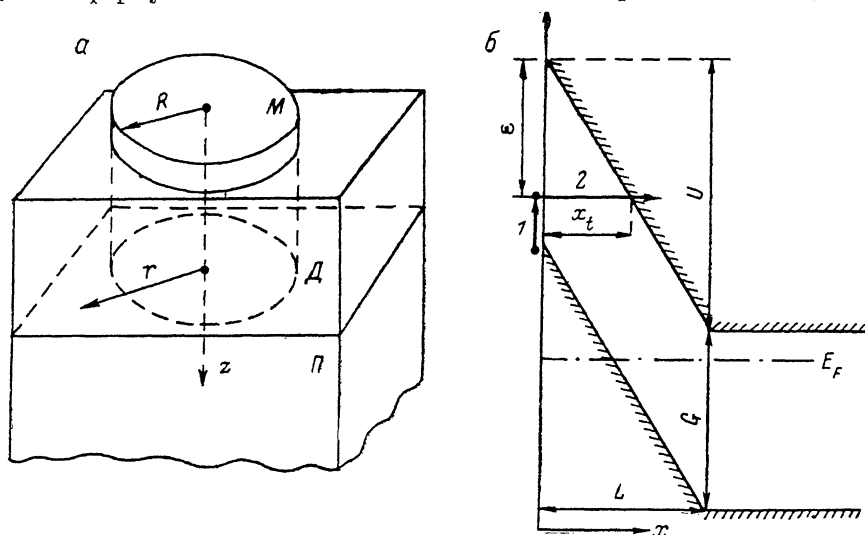


Рис. 1.

а — структура металл—диэлектрик—полупроводник; координата z направлена в глубину полупроводника, координата r лежит в плоскости раздела диэлектрик—полупроводник; R — радиус металлического электрода. *б* — зонная диаграмма полупроводника n -типа в режиме глубокого обеднения; стрелками 1, 2 показаны процессы активации туннелирования; E_F — уровень Ферми в объеме полупроводника.

деление которых излагаемая теория не претендует. Геометрический размер флуктуации порядка длины туннелирования x_t (рис. 1), так что

$$Fx_t = \varepsilon. \quad (2)$$

При заданном значении силы F вероятность туннелирования дается стандартным квантово-механическим выражением

$$P_F = \nu \exp\{-\Phi\}, \quad \Phi = \frac{(m\varepsilon^3)^{1/2}}{\hbar F}, \quad (3)$$

где m — эффективная масса туннелирующего носителя, ν — частотный фактор. С другой стороны, вероятность осуществления прокола с локальной плотностью заряда $q\sigma = q(\sigma_0 + \delta\sigma)$, отвечающей силе F_σ , дается распределением Пуассона

$$P_\sigma = \frac{(x_t^2 \sigma_0)^{x_t^2 \sigma}}{(x_t^2 \sigma)!} \exp(-x_t^2 \sigma_0). \quad (4)$$

Мы должны оптимизировать величину $P_t = P_F^\# P_\sigma$.

Предварительно введем обозначения

$$a_B^\# = \frac{\hbar^2 \varkappa}{mq^2}, \quad \varepsilon_B = \frac{mq^4}{\hbar^2 \varkappa^2}, \quad \zeta = \frac{F_0 z}{q^2 \sigma_0} \quad (5)$$

для эффективных боровских радиуса a_B и энергии ε_B . Здесь F_0 — среднее значение силы F . Оптимизацию произведем по отдельности для случаев малых ($\delta\sigma/\sigma_0 \gg \xi$) и больших ($\delta\sigma/\sigma_0 \ll \xi$) флуктуаций. Это дает

$$P_t(\varepsilon) = \nu \exp \left\{ -\Phi_0 \left[1 - \left(\frac{\varepsilon_B}{\varepsilon} \right)^{1/2} \frac{1}{\xi} \right] \right\} \quad \text{при } \xi \gg \left(\frac{\varepsilon_B}{\varepsilon} \right)^{1/2}, \quad (6)$$

$$P_t(\varepsilon) = \nu \exp \left\{ -\frac{1}{\varepsilon_0 a_B^2} \left(\frac{\varepsilon}{\varepsilon_B} \right)^2 (1 - \xi) \right\} \quad \text{при } \xi \ll \left(\frac{\varepsilon_B}{\varepsilon} \right)^{1/2}, \quad (7)$$

где Φ_0 — среднее значение величины Φ . С учетом определения ξ легко видеть, что формулы (6) или (7) справедливы соответственно, если изгиб зон U в полупроводнике велик или мал по сравнению с характерной энергией

$$\tilde{U}(\varepsilon) = (\varepsilon \varepsilon_B)^{1/2} \varepsilon_0 a_B L. \quad (8)$$

Здесь длина экранирования $L = U/F_0$.

Определим теперь энергию ε с учетом предшествующей туннелированию активации электронов. Будем считать, что образовавшаяся в момент активации дырка мгновенно уносится полем из области прокола, равно как и прошедший под барьером электрон. Тогда условие непрерывности тока через прокол принимает вид

$$\frac{1 - n(\varepsilon)}{\tau_a} = \frac{n(\varepsilon)}{\tau_t}, \quad (9)$$

где $n(\varepsilon)$ — фактор заполнения поверхностных состояний с энергией ε ,

$$\tau_a = \nu^{-1} \exp \left(\frac{G - \varepsilon}{T} \right), \quad \tau_t = \nu^{-1} \exp \{ \Phi(\varepsilon) \} \quad (10)$$

— времена активационного и туннельного процессов. Из (9) следует, что генерационный ток пропорционален величине

$$(\tau_a + \tau_t)^{-1}. \quad (11)$$

Время τ_a экспоненциально убывает, а время τ_t экспоненциально растет с ростом ε . Поэтому условие максимума генерационного тока, определяющее искомую энергию $\varepsilon = \varepsilon_T$, равносильно условию минимума $\tau_a + \tau_t$. Его можно приближенно представить в виде равенства показателей экспонент

$$\frac{G - \varepsilon}{T} = \Phi(\varepsilon). \quad (12)$$

Решение этого уравнения дает

$$\varepsilon_T = \frac{\varepsilon_B^2}{2T} \frac{\varepsilon_0 a_B^2}{1 - \xi} \left[\left(1 + \frac{4GT}{\varepsilon_B^2} \frac{1 - \xi}{\varepsilon_0 a_B^2} \right)^{1/2} - 1 \right] \quad \text{при } U \gg \tilde{U}(\varepsilon_T), \quad (13a)$$

$$\varepsilon_T \approx \frac{\varepsilon_F G}{T^*} \left[\frac{\varepsilon_F}{T^*} + \left(\frac{G}{T^*} \right)^{1/2} \right]^{-1} \quad \text{при } U \ll \tilde{U}(\varepsilon_T), \quad (13b)$$

где

$$\varepsilon_F = \left(\frac{\hbar^2 F_0}{m} \right)^{1/2}, \quad \frac{1}{T^*} = \frac{1}{T} - \frac{1}{\varepsilon_B \xi^2 \varepsilon_0 a_B^2}. \quad (14)$$

Здесь формула (13b) является аппроксимирующей между зависимостями, найденными в предельных случаях низких и высоких значений эффективной температуры T^* .

Пусть размер структуры достаточно велик, чтобы в площади генерации реализовались $N \gg 1$ оптимальных проколов. В этом случае генерационный ток

$$I \approx Sg(\varepsilon_T) Tq P_t(\varepsilon_T), \quad (15)$$

где $g(\varepsilon)$ — плотность поверхностных состояний, S — площадь генерации.

Критерий больших площадей $N \gg 1$ можно представить в виде $g(\varepsilon_T) \chi \times STP_s \gg 1$. Раскрывая выражение P_s для случаев больших и малых U и учитывая равенство (12), можно записать критерий больших площадей как

$$g(\varepsilon_T) ST \exp \left\{ -2 \frac{G - \varepsilon_T}{T} \left(1 + \frac{U}{\bar{U}(\varepsilon_T)} \right)^{-1} \right\} \gg 1. \quad (16)$$

Выражение для показателя экспоненты (16) является аппроксимирующим между предельными случаями $U \ll \bar{U}(\varepsilon_T)$ и $U \gg \bar{U}(\varepsilon_T)$.

Если критерий больших площадей не выполнен и реализуется неравенство, противоположное (16), то генерационный ток сильно флуктуирует от образца

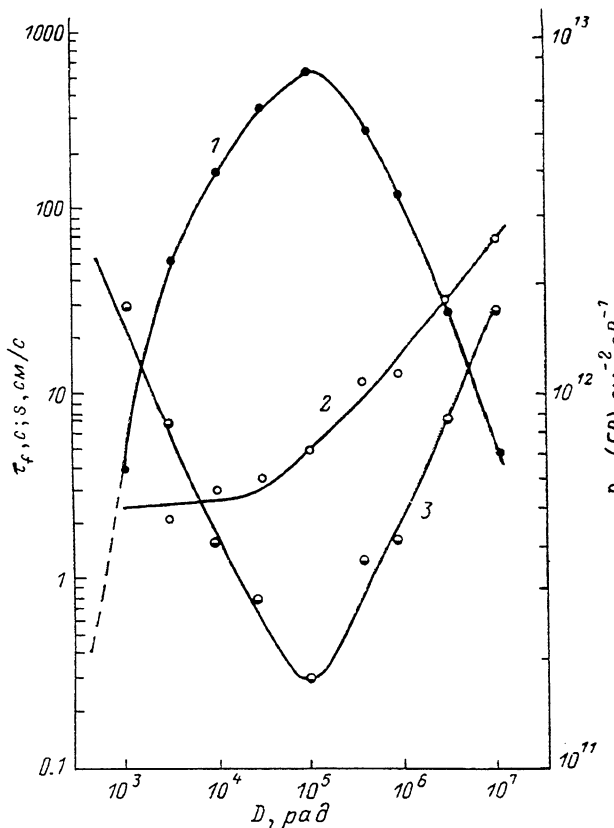


Рис. 2. Дозовые зависимости релаксационных характеристик: времени релаксации τ_f (1), скорости поверхностной генерации s (2), плотности поверхностных состояний D_{ss} (FB) (3) для МНОП структур с сухим термическим SiO_2 .

к образцу. Его величина определяется случайными наиболее прозрачными проколами, встречающимися в области генерации.

Перечислим качественные предсказания, к которым приводит рассмотренная модель. Прежде всего отметим, что энергия активации генерации $\Delta E = G - \varepsilon_T$ растет с ростом T от $\Delta E \ll G$ до $\Delta E = G$. Кроме того, теория предсказывает, что с уменьшением напряжения на полупроводнике с понижением температуры, а также с уменьшением площади образцов могут стать существенными флуктуации генерационных токов от образца к образцу. Еще одно предсказание связано с возможностью подавления определяющих ток проколов за счет экранирования флуктуаций электрического поля дополнительными созданными носителями заряда. Последнего можно достичь, создавая при помощи облучения электрон-дырочные пары в диэлектрике. При этом в первую очередь должны исчезать наиболее прозрачные проколы, поскольку носители заряда наиболее эффективно затягиваются в области высокого электрического поля, создающие

эти проколы. Отсюда следует, в частности, что образцы с большими генерационными токами должны быть более чувствительны к облучению, чем образцы с малыми генерационными токами.

Наши экспериментальные данные подтверждают основные предсказания рассмотренной модели. Измерения производились на МОП (Al—Si*—SiO₂—Si) и МНОП (Al—Si*—Si₃N₄—SiO₂—Si) структурах с площадью управляющего электрода ~1 мм². На подложке из кремния КЭФ-4.5 выращивался сухой термический или пирогенетический окисел SiO₂ толщиной ~600 Å. В случае

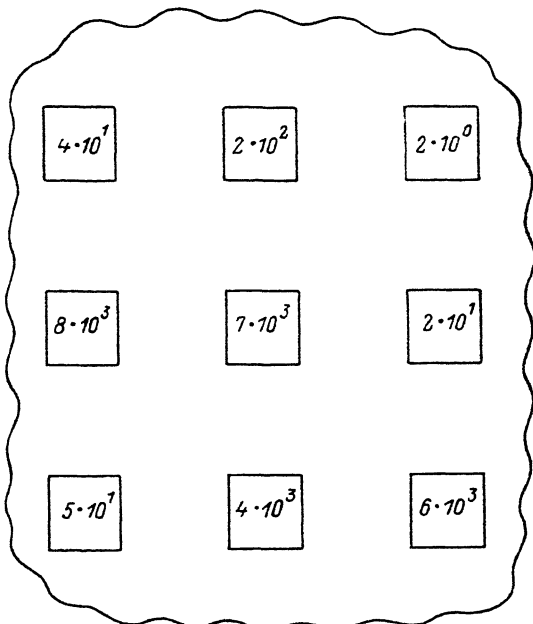


Рис. 3. Распределение времен релаксации МОП структур с сухим термическим SiO₂, выращенных на единой подложке.

Шаг сетки 10 мм, площадь структур 2 × 2 мм. Цифры обозначают время структуры в секундах.

МНОП структур на этот окисел осаждался нитрид кремния толщиной ~800 Å. Управляющим электродом служил слой легированного фосфором поликристаллического кремния с напыленным алюминием. Сформированные структуры подвергались длительному (~15 ч) отжигу при температуре 300 °С в атмосфере водорода. Облучение структур производилось γ-квантами от источника ⁶⁰Co в диапазоне доз 10³—10⁷ рад.

На рис. 2 показаны дозовые зависимости полного времени релаксации емкости τ_f и скорости поверхностной генерации s' при облучении γ-квантами одного из образцов. Величина s' извлекалась из временной зависимости релаксации емкости методом Цербаста [2]. Величины τ_f и s' приблизительно обратно пропор-

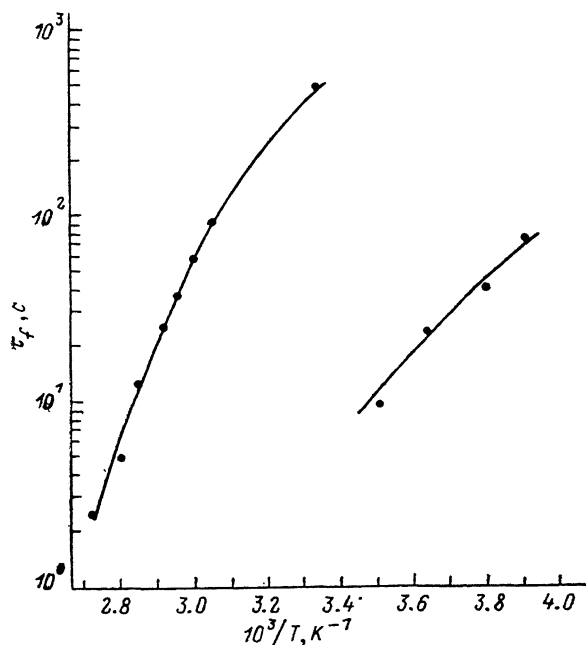


Рис. 4. Температурные зависимости времен релаксации для двух структур с сухим термическим SiO₂.

циональны друг другу. Затягивание релаксации при облучении мы интерпретируем как результат подавления проколов при экранировании областей сильного поля неравновесными носителями заряда. Как видно из рис. 2, при дальнейшем увеличении дозы облучения время релаксации начинает падать. Этот эффект сопровождается увеличением плотности поверхностных состояний и заряда в диэлектрике. Рост этих величин способствует, по нашему мнению,

ускорению темпа генерации (и соответствующему падению времени релаксации) как за счет увеличения числа состояний, участвующих в генерации, так и за счет увеличения флуктуаций электрического поля при накоплении заряда в диэлектрике.

Заметим, что аналогичная описанной дозовая зависимость τ_f наблюдалась в недавней работе [3]. Авторы [3, 4] сделали попытку объяснить полученные данные с точки зрения гетерирования дефектов в напряженной приграничной области и связанного с этим экранирования генерационно активных центров.

На рис. 3 представлено распределение измеренных нами времен релаксации в ансамбле образцов, приготовленных в одном технологическом цикле. Как видно, флуктуации τ_f достигают 3 порядков величины. Дозовые зависимости типа изображенной на рис. 2 относятся только к образцам с короткими начальными временами релаксации. Образцы с начальными значениями $\tau_f \geq 10^3$ с демонстрируют практически полную нечувствительность к облучению вплоть до доз $\sim 10^5$ рад.

Наконец, на рис. 4 показаны активационные зависимости релаксации нескольких образцов. Оцененные по этим данным энергии активации $\Delta E \approx \frac{G}{3} \cdot \frac{G}{4}$, где $G \approx 1.1$ эВ — ширина запрещенной зоны Si. Эти данные свидетельствуют также о некотором повышении энергии активации с ростом температуры.

Таким образом, полученные экспериментальные данные качественно согласуются с предсказаниями модели. Мы попытались произвести некоторые количественные оценки. Взяв на эмпирических основаниях $\epsilon_T \sim 0.7$ эВ, $\sigma_0 \sim 10^{11}$ см⁻², $L \sim 1$ мкм, а также $\epsilon_B \sim 0.01$ эВ, $a_B \sim 100$ Å, получаем из (8) $\bar{U}(\epsilon_T) \sim 1$ эВ. Следовательно, использованные нами напряжения $U \sim 10$ эВ, скорее, соответствуют области слабых флуктуаций. В то же время величина N [левая часть (16)] оказывается много меньше единицы при использованных значениях параметров и плотности поверхностных состояний $g(\epsilon_T) \sim 10^{12}$ см⁻²·эВ⁻¹. Таким образом, реализуется критерий малых площадей. Это согласуется с фактом наблюдения флуктуаций на рис. 3.

Мы благодарим М. Э. Райха за полезное обсуждение результатов работы.

Список литературы

- [1] Райх М. Э., Рузин И. М. // ЖЭТФ. 1987. Т. 92. В. 6. С. 2257—2276.
- [2] Kang J. S., Schroder D. K. // Phys. St. Sol. (a). 1985. V. 89. N 1. P. 13—43.
- [3] Борковская О. Ю., Дмитрук Н. А., Конакова Р. В., Литовченко В. Г., Маева О. И. ФТП. 1986. Т. 20. В. 9. С. 1640—1646.
- [4] Борковская О. Ю., Дмитрук Н. А., Литовченко В. Г., Мищук О. Н. // ФТП. 1988. Т. 23. В. 2. С. 207—212.

Ленинградский политехнический институт им. М. И. Калинина

Получена 22.06.1989
Принята к печати 5.07.1989