

Простой контрпример для \mathcal{Z}_2 -классификации топологических изоляторов, основанной на соответствии объем–граница

С. Н. Молотков^{+*×1)}, М. И. Рыжкин⁺

⁺ Институт физики твердого тела РАН, 142432 Черноголовка, Россия

^{*} Академия криптографии РФ, 121552 Москва, Россия

[×] Факультет вычислительной математики и кибернетики, МГУ им. Ломоносова, 119991 Москва, Россия

Поступила в редакцию 20 апреля 2015 г.

После переработки 23 июня 2015 г.

Ранее была предложена так называемая \mathcal{Z}_2 -классификация топологических изоляторов, основанная на соответствии объем–граница (bulk-boundary correspondence), которая считается общепринятой и сводится к следующим утверждениям: 1) nontrivial \mathcal{Z}_2 invariants imply the existence of gapless surface states, 2) the \mathcal{Z}_2 invariants can be deduced from the topological structure of the Bloch wave functions of the bulk crystal in the Brillouin zone (L. Fu and C. L. Kane, Phys. Rev. B **76**, 045302 (2007)). В данной работе приводится простой контрпример для \mathcal{Z}_2 -классификации. Показано, что при одном и том же объеме, одной и той же пространственной симметрии полубесконечного кристалла и, соответственно, тривиальном значении \mathcal{Z}_2 -инварианта (тривиальном классе эквивалентности объемного гамильтониана) для 3→2D-системы на поверхностях могут существовать как топологически устойчивые, так и топологически неустойчивые поверхностные состояния. Более того, топологически устойчивые поверхностные состояния могут существовать как при тривиальном (поверхность Bi(111)), так и при нетривиальном (поверхность Sb(111)) значениях объемного \mathcal{Z}_2 -инварианта. Данные факты ставят под сомнение утверждение о том, что \mathcal{Z}_2 -классификация, основанная на соответствии объем–граница, отвечает за появление и топологическую устойчивость поверхностных состояний.

DOI: 10.7868/S0370274X15150102

Введение. Зонная теория твердого тела приводит к разделению трехмерных (3D) кристаллов на изоляторы (полупроводники) и металлы. Одноточечный спектр 3D кристаллической системы представляет собой набор энергетических зон $\varepsilon_n(\mathbf{k})$ (n – зонный индекс, $\mathbf{k} = (k_x, k_y, k_z)$ – трехмерный квазиимпульс). В объеме 3D-система является изолятором, если имеются два массива зон, $\varepsilon_n^c(\mathbf{k})$ и $\varepsilon_{n'}^v(\mathbf{k})$, такие, что энергии этих зон не перекрываются ни при каких значениях трехмерных квазиимпульсов \mathbf{k} и \mathbf{k}' ,

$$\varepsilon_G = \min_{n, n', \mathbf{k}, \mathbf{k}'} [\varepsilon_n^c(\mathbf{k}) - \varepsilon_{n'}^v(\mathbf{k}')] > 0, \quad (1)$$

где ε_G – запрещенная зона. Массив валентных зон $\varepsilon_n^v(\mathbf{k})$ заполнен, а массив зон проводимости $\varepsilon_n^c(\mathbf{k})$ пуст. В металлах и полуметаллах энергетический зазор между пустыми и заполненными состояниями отсутствует. Вид спектров $\varepsilon_n^c(\mathbf{k})$ и $\varepsilon_n^v(\mathbf{k})$ в разных точках трехмерной зоны Бриллюэна (ЗБ) диктуется неприводимыми спинорными представлениями пространственной группы волнового вектора в дан-

ной точке ЗБ. Учет инвариантности системы относительно инверсии времени может приводить к дополнительному принудительному вырождению уровней. Метод Херринга дает регулярный способ, который позволяет получить структуру энергетических зон с учетом инвариантности системы относительно инверсии времени [1, 2].

Поверхностные состояния в 3 → 2D-системах. Если разорвать бесконечный 3D-кристалл на две независимые половины, то возникнет полубесконечный 3D-кристалл с 2D-границей (3D→2D-система). В 3D→2D-системе периодичность сохраняется только вдоль границы. При разрыве 3D-кристалла на границе могут возникать поверхностные состояния, локализованные вблизи поверхности. Данные состояния возникают из объемных состояний 3D-системы. При этом общее число состояний при образовании границы в кристалле сохраняется. Отношение числа поверхностных состояний к числу состояний в объеме стремится к нулю. По этой причине основной массив объемных состояний в 3D→2D-системе при разрыве 3D-кристалла слабо возмущен. *Электронный спектр*

¹⁾ e-mail: sergei.molotkov@gmail.ru

всех зон $3D \rightarrow 2D$ -системы, включая объемные, имеет симметрию, диктуемую представлениями 17 поверхностных групп, а не 230 пространственными группами. Все состояния $3D \rightarrow 2D$ -системы описываются набором двумерных зон $\varepsilon_n^{\text{surf}}(\mathbf{k})$ (здесь $\mathbf{k} = (k_x, k_y)$ и координата z нормальна к границе). Поскольку число слабозмущенных состояний макроскопически велико, набор двумерных зон в полубесконечном кристалле $\varepsilon_n^{\text{surf}}(\mathbf{k})$ ($\mathbf{k} = (k_x, k_y)$) заполняет проекции объемных зон $3D$ -кристалла:

$$\varepsilon_n(k_x, k_y) = \sum_{k_z} \varepsilon_n(\mathbf{k}), \quad \mathbf{k} = (k_x, k_y, k_z), \quad (2)$$

где $\varepsilon_n(\mathbf{k})$ – спектр $3D$ -кристалла. Поверхностными состояниями называются те состояния $\varepsilon_n^{\text{surf}}(\mathbf{k})$, у которых энергия $\varepsilon_n^{\text{surf}}(\mathbf{k})$ ($\mathbf{k} = (k_x, k_y)$) хотя бы при некоторых значениях $\mathbf{k} = (k_x, k_y)$ находится в запрещенной зоне, вне проекций объемных зон (2).

Безмассовые дираковские поверхностные состояния, \mathcal{Z}_2 -классификация, соответствие объем–граница. Поверхностные состояния на поверхностях полупроводников и металлов изучаются достаточно давно. В последнее десятилетие были обнаружены поверхностные состояния с безмассовым дираковским спектром в окрестности симметричных точек двумерной зоны Бриллюэна (см. обзоры [3, 4]). Все типы особенностей электронного спектра в $3 \rightarrow 2D$ -системах перечислены в [5, 6]. Поверхностные состояния возможны также в $2 \rightarrow 1D$ -системах (двумерные системы с одномерной границей или ленты с двумя границами). Такие $2 \rightarrow 1D$ -системы с поверхностными состояниями с коническим дираковским спектром были названы спин-холловскими изоляторами.²⁾ Классификация особенностей спектра в $2 \rightarrow 1D$ -системах приведена в [7].

Возникло понятие “топологический изолятор” (для $2 \rightarrow 1D$ -системы – спин-холловский изолятор). Под топологическим изолятором понимается такая $3 \rightarrow 2D$ (или $2 \rightarrow 1D$) система, которая в объеме является изолятором, а на поверхности имеет состояния, по которым возможна проводимость. Под топологической устойчивостью поверхностных состояний понимается следующее [3, 8, 9]: 1) топологически устойчивые конические состояния (topologically protected states) невозможно непрерывным изменением параметров гамильтониана удалить из запрещенной зоны в проекции объемного спектра

²⁾ В терминологии [3, 4, 8, 9], системы $3 \rightarrow 2D$ носят название трехмерных топологических изоляторов. Поэтому часто неясно, о чем идет речь: о поверхностных состояниях на границе $3D$ -системы или объемном спектре $3D$ -системы.

без захлопывания объемной запрещенной зоны; 2) топологически неустойчивые конические состояния могут быть удалены из запрещенной зоны без захлопывания последней, т.е. при сохранении системы изолятором.

В [3, 8, 9] была предложена так называемая \mathcal{Z}_2 -классификация, основанная на соответствии объем–граница (bulk-boundary correspondence). Затем она использовалась во многих работах (см. обзоры [3, 4, 8, 9] и ссылки там).

Согласно [3, 9] эти две ситуации различаются значением топологического инварианта: $\mathcal{Z}_2 = -1$ (нетривиальное значение) или $\mathcal{Z}_2 = 1$ (тривиальное значение). Для систем $2 \rightarrow 1D$ имеется один \mathcal{Z}_2 -инвариант. В первом случае система – топологический изолятор (спин-холловский), во втором – обычный. Для систем $3 \rightarrow 2D$ имеется четыре \mathcal{Z}_2 -инварианта. В зависимости от числа знаков всех четырех \mathcal{Z}_2 -инвариантов авторами вводится [9] подразделение на слабые (weak) и сильные (strong) топологические изоляторы в $3 \rightarrow 2D$ -случае. Значение \mathcal{Z}_2 -инварианта представляет собой пффафиан, вычисленный на объемных волновых функциях.

Кроме того, согласно классификации ([9], с. 045302-3): “The \mathcal{Z}_2 invariants can be deduced from the topological structure of the Bloch wave functions of the bulk crystal in the Brillouin zone”.

Дается и объяснение самого факта появления поверхностных состояний ([9], с. 045302-1): “Nontrivial \mathcal{Z}_2 invariants imply the existence of gapless surface states”.

Эти утверждения означают, что для предсказания факта появления поверхностных состояний и их топологической устойчивости достаточно знать только топологический инвариант, вычисленный на объемных волновых функциях. Если инвариант нетривиален, то должны иметь место топологически устойчивые поверхностные состояния. Если же он тривиален, то могут существовать только топологически неустойчивые поверхностные состояния.

Критика классификации, основанной на соответствии объем–граница. На наш взгляд, самоочевидно, что поверхностные состояния и их тип не определяются только объемными свойствами системы (напомним, что в систематике объемного спектра вообще нет поверхностных состояний), важны также граничные условия. Чтобы показать несостоятельность \mathcal{Z}_2 -классификации, достаточно одного контрпримера. Достаточно предъявить кристаллическую систему, у которой (естественно, при одинаковом объемном спектре, одинаковой симметрии полубесконечного кристалла и тривиальном значе-

нии инварианта), на одной границе поверхностные состояния будут топологически устойчивыми, а на другой – топологически неустойчивыми. Разные поверхности отвечают разным граничным условиям. Такой явно сконструированный пример будет приведен ниже. Иначе говоря, одни и те же объемные волновые функции и спектр дадут одно и то же значение \mathcal{Z}_2 , при котором на разных поверхностях будут реализовываться две принципиально разные ситуации. Известны и другие примеры (см. [10]), явно противоречащие \mathcal{Z}_2 -классификации. \mathcal{Z}_2 -классификация дает некоторую топологическую характеристику объемного спектра как такового, которая напрямую (один в один) не связана с появлением поверхностных состояний и их топологической устойчивостью. Кроме того, \mathcal{Z}_2 -классификация не воспроизводит другие типы особенностей, которые предсказывает симметрия. Соответствующий явный пример был приведен в [11]. Это указывает на неполноту классификации $3 \rightarrow 2D$ -систем, приведенную в [12].

Простой контрпример для \mathcal{Z}_2 -классификации. Сконструируем гамильтониан для конкретной 3D пространственной кристаллической группы, который будет удовлетворять всем требованиям пространственной симметрии и будет инвариантен относительно инверсии времени. Затем перейдем от 3D-системы к $3 \rightarrow 2D$ -системе. Данный пример не является абстрактным, а относится к элементам V группы таблицы Менделеева (As, Bi, Sb, Po [13]) и узкощелевым полупроводникам A^4B^6 (SnTe, PbTe и соединениям на их основе) [14]. На гетерограницах PbTe–SnTe были обнаружены конические безмассовые состояния [15], которые связаны с инверсным расположением по энергии p -орбиталей на краях валентной зоны и зоны проводимости. Эти состояния были предсказаны Волковым и Панкратовым еще задолго до появления понятия “топологический изолятор”. Элементы V группы и полупроводники A^4B^6 имеют общий генезис спектра: базовой является простая кубическая решетка (пра-фаза) с центрированными на каждом узле тремя p -орбиталями. Пра-фаза является металлом. Ферми-поверхность пра-фазы имеет плоские участки – вкладывается сама в себя при трансляции на определенные векторы обратной решетки. Это свойство спектра пра-фазы приводит к пайерлсовской неустойчивости последней – сдвигу атомных плоскостей в направлении [111] через одну. При этом на основной части поверхности Ферми открывается щель. Лишь небольшое перекрытие валентной зоны и зоны проводимости вдоль некоторых направлений ЗБ делает указанные системы полуме-

таллами. (Подробности происхождения спектра см. в [13, 14].)

Поскольку для нашего контрпримера нужен только спектр, а не волновые функции, удобно воспользоваться методом сильной связи. Данный метод прост и описан практически в любом стандартном учебнике по теории твердого тела. Это позволяет всем желающим легко проверить каждый шаг приводимых ниже вычислений. Рассмотрим простую кубическую решетку, в каждом узле которой центрированы три орбитали (p_x, p_y, p_z). С учетом спина имеется 6 состояний на атом. Число электронов на атом равно 3. Таким образом, орбитали заполнены наполовину. Данный спектр является металлическим. Поскольку мы не боремся за описание тонких деталей электронного спектра, например таких, как эффективные массы электронов и дырок, для того чтобы не загромождать выкладки, можно ограничиться интегралами перекрытия ближайших соседей. Все требования кристаллической симметрии при этом соблюдаются.

Ниже для иллюстративных расчетов мы будем использовать такие параметры гамильтониана, при которых вообще нет пересечения объемной валентной зоны уровня Ферми ни по одному направлению. Таким образом, обсуждаемые примеры будут относиться к классу (0;0,0,0) (согласно [9]), т.е. к тривиальному случаю. Более того, если не учитывать ромбоэдрическое искажение, то все направления между объемными точками L и T станут эквивалентными. Тем не менее при разных сколах объемного кристалла возникают как топологически устойчивые, так и топологически неустойчивые поверхностные состояния.

Вернемся к конструированию гамильтониана. В базисе p -орбиталей гамильтониан параметризуется двумя независимыми константами. Спектр состоит из трех вырожденных по спину зон (постоянная решетка $a = 1$, $\mathbf{k} = (k_x, k_y, k_z)$):

$$\begin{aligned} \hat{H}^0(\mathbf{k}) &= \hat{\xi}(\mathbf{k}), \\ \xi_{xx}(\mathbf{k}) &= \xi_0 \cos(k_x) + \xi_1 [\cos(k_y) + \cos(k_z)]. \end{aligned} \quad (3)$$

Другие матричные элементы получаются циклической перестановкой индексов:

$$\begin{aligned} \xi_0 &= 2\langle p_x(0,0,0) | \mathcal{H} | p_x(1,0,0) \rangle, \\ \xi_1 &= 2\langle p_x(0,0,0) | \mathcal{H} | p_x(0,1,0) \rangle, \end{aligned} \quad (4)$$

где \mathcal{H} – гамильтониан пра-фазы. Матрица спин-орбитального взаимодействия в базисе p -орбиталей имеет вид

$$\hat{\Delta} = i\lambda \begin{pmatrix} 0 & \sigma_x & -\sigma_y \\ -\sigma_z & 0 & \sigma_x \\ \sigma_y & -\sigma_x & 0 \end{pmatrix}. \quad (5)$$

Спектр гамильтониана пра-фазы $\hat{H}^0(\mathbf{k})$, состоящий из трех зон, $\varepsilon_{1,2,3}(\mathbf{k})$, обладает свойством $\varepsilon_{1,2,3}(\mathbf{k}) = -\varepsilon_{1,2,3}(\mathbf{k} + \mathbf{Q})$ (имеет плоские участки (совмещается сам с собой при трансляции на один из векторов \mathbf{Q}), что приводит при половинном заполнении зон к пайерсовской неустойчивости (сдвигу атомных плоскостей через одну) пра-фазы вдоль одного из кристаллографических направлений – $[111]$ (где $\mathbf{Q} = (1, 1, -1), (1, -1, 1), (-1, 1, 1)$ – один из векторов). Непосредственной проверкой можно убедиться в том, что гамильтониан $\hat{H}(\mathbf{k}) = \hat{H}^0(\mathbf{k}) + \hat{\Delta}$ удовлетворяет всем требованиям кристаллической симметрии и инвариантен относительно инверсии времени, а его спектр описывается спинорными представлениями пространственной группы (см. [13])³.

Пусть пайерсовский сдвиг атомных плоскостей через одну происходит в кристаллографическом направлении $[111]$. Он приведет к открытию пайерсовской щели на поверхности Ферми и превращению металлического спектра в диэлектрический. Данный сдвиг вызовет изменение интегралов перекрытия на ближайших соседних атомах. При этом симметрия диктует появление трех независимых матричных элементов [13, 14]. Матричные элементы имеют вид

$$U_{xx} = i\{u_1 \sin(k_x) + u_2[\sin(k_y) + \sin(k_z)]\}, \quad (6)$$

$$U_{xy} = iu_3[\sin(k_y) + \sin(k_z)],$$

где

$$u_1 = \pm 2\langle p_x(0, 0, 0) | \mathcal{H}_{\text{Pierls}} | p_x(\pm 1, 0, 0) \rangle, \quad (7)$$

$$u_2 = \pm 2\langle p_x(0, 0, 0) | \mathcal{H}_{\text{Pierls}} | p_y(0, \pm 1, 0) \rangle,$$

$$u_3 = \pm 2\langle p_x(0, 0, 0) | \mathcal{H}_{\text{Pierls}} | p_y(\pm 1, 0, 0) \rangle.$$

Остальные элементы получаются циклической перестановкой индексов. Здесь $\mathcal{H}_{\text{Pierls}}$ – изменение гамильтониана \mathcal{H} пра-фазы за счет пайерсовского сдвига. Этого достаточно для наших целей. Учет перекрытия вторых соседей и малого ромбоэдрического искажения решетки нужен лишь для тонких деталей спектра. Нам он не потребуется. Расчет спектров поверхностных состояний при разрыве по короткой

³Если спин-орбитальное взаимодействие достаточно велико, то оно сильно искажает плоские участки поверхности Ферми [16] и система остается в кубической фазе. В природе имеется элемент, который при нормальных условиях кристаллизуется в простую кубическую решетку. Это α -Ро, открытый М. Складовской-Кюри.

связи с учетом всех деталей спектра для Bi, Sb, As может быть найден в работе [10]. Окончательно объемный (3D) гамильтониан имеет вид

$$\hat{H}(\mathbf{k}) = \begin{pmatrix} \hat{\xi}(\mathbf{k}) + \hat{\Delta} & i\hat{u}(\mathbf{k}) \\ -i\hat{u}(\mathbf{k}) & \hat{\xi}(\mathbf{k} + \mathbf{Q}) + \hat{\Delta} \end{pmatrix}. \quad (8)$$

При этом его размерность 12×12 . После удвоения периода вдоль направления $[111]$ в элементарной ячейке имеется два атома. Число атомных состояний $2 \times 3 \times 2 = 12$. На трех p -орбиталях находится шесть электронов. Зоны заполнены наполовину. Заполненные состояния отделены от пустых запрещенной зоной.

Гамильтониан полубесконечного кристалла с разными поверхностями. Разрыв 3D-кристалла перпендикулярно направлению $[111]$ можно устроить двумя способами (рис. 1): по короткой или по длинной

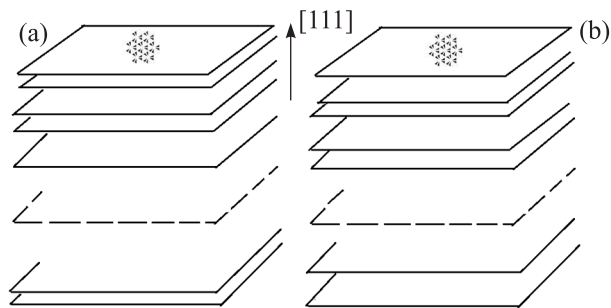


Рис. 1. Два варианта скола 3D-кристалла: по длинной связи (a) и по короткой связи (b)

связи. Симметрия полубесконечного кристалла с поверхностью будет одной и той же. Она относится к 15-й группе поверхностных групп [5, 6].

Для построения гамильтониана полубесконечного кристалла надо исключить соответствующие интегралы перекрытия. Для этого удобно преобразовать систему координат: перейти к системе, в которой одна из осей направлена вдоль направления $[111]$. Преобразование системы координат в прямой решетке индуцирует соответствующее преобразование в обратной решетке, которое нам и понадобится

$$\begin{aligned} k_x &= \frac{k'_x}{\sqrt{6}} - \frac{k'_y}{\sqrt{2}} + \frac{k_\perp}{\sqrt{3}}, \\ k_y &= \frac{k'_x}{\sqrt{6}} + \frac{k'_y}{\sqrt{2}} + \frac{k_\perp}{\sqrt{3}}, \\ k_z &= -\sqrt{\frac{2}{3}}k'_x + \frac{k_\perp}{\sqrt{3}}. \end{aligned} \quad (9)$$

Матричные элементы перекрытия в повернутой системе координат между соседними плоскостями (111) преобразуются к виду

$$\begin{aligned}\xi_{x\perp}^\pm(k'_x, k'_y) &= \frac{1}{2}[(\xi_0 \pm u_1)e^{ik^-} + (\xi_0 \pm u_2)(e^{ik^+} + e^{ik_{\parallel}})], \\ \xi_{y\perp}^\pm(k'_x, k'_y) &= \frac{1}{2}[(\xi_0 \pm u_1)e^{ik^+} + (\xi_0 \pm u_2)(e^{ik^-} + e^{ik_{\parallel}})], \\ \xi_{z\perp}^\pm(k'_x, k'_y) &= \frac{1}{2}[(\xi_0 \pm u_1)e^{ik_{\parallel}} + (\xi_0 \pm u_2)(e^{ik^+} + e^{ik^-})].\end{aligned}\quad (10)$$

Здесь введены обозначения $k_{\parallel} = 2k'_y/\sqrt{6}$, $k_{\pm} = k'_x/\sqrt{2} \pm k'_y/\sqrt{6}$. Осталось преобразовать недиагональные по индексам p -орбиталей матричные элементы (6). Имеем

$$u_{xy\perp}^\pm(k'_x, k'_y) = \pm \frac{1}{2}[u_3(e^{ik^+} + e^{ik^-})],$$

$$u_{xz\perp}^\pm(k'_x, k'_y) = \pm \frac{1}{2}[u_3(e^{ik^-} + e^{ik_{\parallel}})], \quad (11)$$

$$u_{yz\perp}^\pm(k'_x, k'_y) = \pm \frac{1}{2}[u_3(e^{ik^+} + e^{ik_{\parallel}})].$$

В такой записи обрыв кристалла организуется исключением интегралов перекрытия между слоями в направлении [111]. Выбор разрыва поверхности по короткой или длинной связи осуществляется выбором знака "±" в матричных элементах, ответственных за пайерлсовский сдвиг, в $\hat{E}_{\perp}^+(k'_x, k'_y)$ или $\hat{E}_{\perp}^-(k'_x, k'_y)$ в двух первых и двух последних строках матрицы (12). После этого возникает кристалл с двумя одинаковыми поверхностями (рис. 1; это стандартная процедура при расчете поверхностных состояний). Гамильтониан кристалла (slab) – периодический в направлениях k'_x, k'_y , $\hat{E}_{\parallel}(k'_x, k'_y) = \hat{\Delta}$,

$$\hat{H}_{\text{surf}}(k'_x, k'_y) = \begin{pmatrix} \hat{E}_{\parallel}(k'_x, k'_y) & \hat{E}_{\perp}^+(k'_x, k'_y) & & & & \\ \hat{E}_{\perp}^{+*}(k'_x, k'_y) & \hat{E}_{1\parallel}(k'_x, k'_y) & \hat{E}_{\perp}^-(k'_x, k'_y) & & & \\ & \hat{E}_{\perp}^{-*}(k'_x, k'_y) & \hat{E}_{1\parallel}(k'_x, k'_y) & & & \\ & & & \dots & & \\ & & & & \hat{E}_{\parallel}(k'_x, k'_y) & \hat{E}_{\perp}^+(k'_x, k'_y) \\ & & & & \hat{E}_{\perp}^{+*}(k'_x, k'_y) & \hat{E}_{\parallel}(k'_x, k'_y) \end{pmatrix}, \quad (12)$$

$$\hat{E}_{\perp}^{\pm}(k'_x, k'_y) = \begin{pmatrix} \hat{u}^{\pm}(k'_x, k'_y) & 0 \\ 0 & \hat{u}^{\pm}(k'_x, k'_y) \end{pmatrix}, \quad \hat{u}_{\perp}^{\pm}(k'_x, k'_y) = \begin{pmatrix} u_{xx}^{\pm}(k'_x, k'_y) & u_{xy}^{\pm}(k'_x, k'_y) & u_{xz}^{\pm}(k'_x, k'_y) \\ u_{yx}^{\pm}(k'_x, k'_y) & u_{yy}^{\pm}(k'_x, k'_y) & u_{yz}^{\pm}(k'_x, k'_y) \\ u_{zx}^{\pm}(k'_x, k'_y) & u_{zy}^{\pm}(k'_x, k'_y) & u_{zz}^{\pm}(k'_x, k'_y) \end{pmatrix}. \quad (13)$$

Результаты.⁴⁾ Еще раз подчеркнем, что при двух способах разрыва кристалла симметрия 3→2D-системы одинакова. Одинаковы также все параметры 3D объемного гамильтониана, а значит, и величина \mathcal{Z}_2 -инварианта. Две поверхности удалены друг от друга так, чтобы поверхностные состояния внутри запрещенной зоны на противоположных поверхностях не перекрывались.

Полное число состояний при разрыве кристалла сохраняется. Появление поверхностных состояний вне проекций объемного спектра приводит к уменьшению числа состояний в проекциях объемных зон.

Число состояний в объемной валентной зоне (N_v) и зоне проводимости (N_c) есть $N_c = N_v = 2 \cdot 3 \cdot N_{\text{layer}}$ (N_{layer} – число слоев, см. рис. 1). Расчет изменения числа состояний в объемных зонах производится прямым подсчетом собственных значений гамильтониана (12). При обоих способах разрыва кристалла (по длинной или короткой связи) поверхностные состояния возникают на обеих границах. Число топологически устойчивых состояний на одной границе (рис. 2) равно двум (по одному состоянию на каждую ветвь конуса). Это приводит к уменьшению на одно состояние в объемной валентной зоне и на одно

⁴⁾ Обычно в экспериментах с элементами V группы при приготовлении поверхности скол происходит по длинной связи. Однако можно приготовить и второй тип поверхности (скол по короткой связи). Имеются экспериментальные данные ARPES, указывающие на наличие двух типов состояний на разных поверхностях (см. частное сообщение А.М. Ионина, С.И. Божко, которое будет опубликовано отдельно).

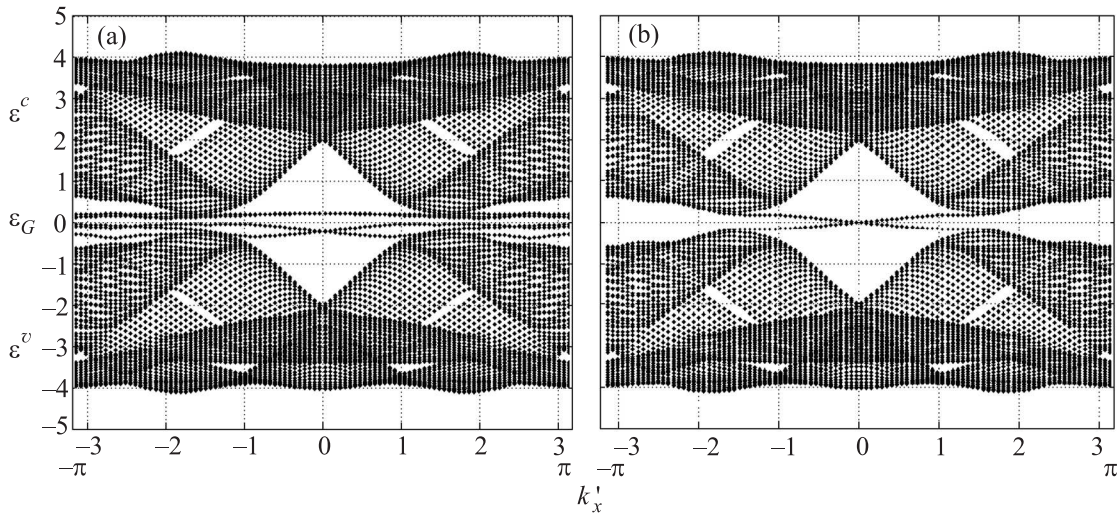


Рис. 2. Два типа поверхностных состояний: скол по длинной связи, топологически неустойчивые (а), скол по короткой связи, топологически устойчивые (b). В случае а нижняя коническая ветвь состояний отщепляется от валентной зоны, а верхняя (с почти плоским раствором конуса) – от зоны проводимости. В случае b коническая ветвь спектра отщепляется сразу от двух зон – валентной и зоны проводимости. При расчетах использовались следующие параметры (в эВ): $\xi_0 = 3.606$, $\xi_1 = 0.0$, $u_1 = -0.7$, $u_2 = -0.07$, $u_3 = 0.7$, $\lambda = 0.2143$, которые близки к значениям для Sb. Число слоев равно 40. Данные параметры взяты произвольными. Спектр приведен только вдоль одного симметричного направления (Г–М). Вид спектра вдоль других симметричных направлений двумерной зоны Бриллюэна (Г–К) имеет аналогичный вид

состояние в объемной зоне проводимости. С учетом двух границ уменьшение числа состояний удваивается. Происхождение поверхностных состояний сразу из двух зон не позволяет выдавить их в одну из зон, не захлопывая щель.

Для топологически неустойчивых состояний (рис. 2а) состояния обеих ветвей конуса происходят только из состояний одной из зон (либо валентной, либо зоны проводимости). При этом уменьшение числа состояний в объемной зоне при образовании поверхностного состояния равно двум на каждый конус. Поскольку коническое поверхностное состояние генетически происходит только из состояний одной из объемных зон, то можно выдавить данное состояние в ту зону, из которой оно произошло, не затрагивая состояния другой зоны.

Почему состояния рис. 2b топологически устойчивы? Иными словами, почему они не могут быть удалены из щели так, чтобы система осталась изолятором? Если бы можно было, например, уменьшая до нуля связь на одной из границ, вытолкнуть пару наполовину заполненных состояний в одну из объемных зон, например в валентную, то число состояний в этой объемной зоне увеличилось бы на два. При этом одно состояние оказалось бы заполненным, а другое – пустым. Система стала бы металлом с проводимостью по объемным состояниям, что эквивалентно

захлопыванию щели. При попытке выталкивания в зону проводимости ситуация была бы аналогичной. Поэтому данная ситуация принципиально отличается от ситуации рис. 2а, когда поверхностные состояния либо полностью заполнены, либо полностью пусты (в зависимости от их происхождения).

Поясним физическую причину топологической устойчивости. Невозможно выдавить состояния в объемный спектр только на одной границе, деформируя любое конечное число интегралов перекрытия. В противном случае возникает ситуация, описанная выше. Для удаления состояний в объемный спектр требуется сдвинуть все плоскости в кристалле, заменив короткие связи на длинные. В этом случае на обеих поверхностях если и возникнут состояния, то первого типа (рис. 2а). Однако при замене коротких связей на длинные и наоборот (этого можно добиться, сдвигая плоскости через одну, т.е. меняя пайерлсовскую деформацию – знак пайерлсовского матричного элемента; при такой смене знака он обязательно пройдет через нуль) атомы пройдут через положения, когда все связи одинаковы (пайерлсовское искажение отсутствует). Подобная ситуация отвечает отсутствию диэлектрической щели, т.е. при такой деформации систему нельзя сохранить изолятором. Это и есть ситуация захлопывания щели. После такой деформации обрыв поверхности кристалла будет

отвечать обрыву по длинной связи, т.е. мы приходим к ситуации рис. 2а. Аналогичные рассуждения работают и в одномерном случае (см. ниже).

Согласно \mathcal{Z}_2 -классификации [9] висмут (Bi) и сурьма (Sb) относятся к разным топологическим классам. Висмут является тривиальным изолятором (набор инвариантов $(0 : 0, 0, 0)$), а Sb относится к нетривиальному объемному топологическому классу с набором инвариантов $(1 : 1, 1, 1)$. Согласно идеологии [9] на поверхности Sb(111) должны находиться топологически устойчивые конические поверхностные состояния в щели объемного спектра, происходящие из валентной зоны и зоны проводимости. И наоборот, Bi не должен иметь топологически устойчивых поверхностных состояний, происходящих одновременно из валентной зоны и зоны проводимости. Расчеты поверхностных состояний на поверхностях Bi(111) и Sb(111) при сколе по короткой связи (рис. 1) были проведены еще в 1988 г. (см. [10], рис. 3 и 4) для реальных параметров интегралов перекрытия, аккуратно подобранных для описания объемных спектров Bi и Sb [13].

Однако топологически устойчивые конические состояния, происходящие одновременно из двух зон, появляются как на поверхности Sb(111), так и на поверхности Bi(111). Данное обстоятельство является явным контрпримером против \mathcal{Z}_2 -классификации, основанной на соответствии объем-граница.

Поверхностные состояния в $1 \rightarrow 0D$ -системе.

Написание данного пункта было инициировано замечанием рецензента. Отметим, что в [17] рассмотрена одномерная модель Шокли. Ниже мы приведем нашу модель к модельной одномерной ситуации. В этом упрощенном случае выводы [17] переходят в наши результаты один в один. Приведем результаты для альтернированной одномерной цепочки (рис. 3) с нуль-мерной поверхностью. В данной ситуации при одном и том же объеме возможно как возникновение топологически устойчивых поверхностных состояний (при сколе цепочки по короткой связи), так и их отсутствие (при сколе по длинной связи). Топологическая устойчивость состояний при обрыве кристалла по короткой связи устанавливается аналогично предыдущей $3 \rightarrow 2D$ -ситуации посредством подсчета изменения числа состояний в валентной зоне и зоне проводимости.

Для того чтобы перейти к $1 \rightarrow 0D$ -системе, достаточно в гамильтониане (12) положить все интегралы перекрытия, включая константу спин-орбитального взаимодействия, равными нулю, кроме ξ_0 и u_1 (изменение ξ_0 за счет пайерлсовского сдвига плоскостей через одну). Тогда, как видно из (12), ситуация ста-

новится одномерной, т.е. объем эквивалентен одномерной цепочке с пайерлсовским сдвигом атомов через один и соответствующим изменением интегралов перекрытия (см. рис. 3, нижняя половина). Возможно образование поверхности при обрыве по длинной связи (рис. 3а) или по короткой связи (рис. 3б). При обрыве по длинной связи поверхностные состояния не возникают. При обрыве по короткой связи возникают топологически устойчивые состояния, происходящие из валентной зоны и зоны проводимости одновременно, что проверяется подсчетом изменения числа состояний. При этом из-за электрон-дырочной симметрии состояния возникают строго в середине запрещенной зоны.

Объемный спектр в модели Шокли имеет вид $\varepsilon(k) = \pm \sqrt{|t(k)|^2}$, $t(k) = t_1 + t_2 e^{ik}$, где k – квазиимпульс вдоль цепочки. Объемный спектр в одномерной пайерлсовской цепочке есть $\varepsilon(k) = \pm \sqrt{\xi_0^2 \cos^2(k) + u^2 \sin^2(k)}$. Две указанные модели одинаковы. Соответствие устанавливается очевидным образом: $2t_1 \leftrightarrow \xi_0 + u$, $2t_2 \leftrightarrow \xi_0 - u$, при условии, что $t_1 > t_2$, а ξ_0 и u одного знака. Поэтому обрыв по “длинной связи” (исключение t_2 , см. рис. 3) не приведет к появлению поверхностных состояний. Обрыв же цепочки по “короткой связи” (исключение t_1 , см. рис. 3) будет приводить к появлению топологически устойчивых состояний строго в середине запрещенной зоны. Результаты [17] находятся в полном соответствии с нашими выводами. Условие появления поверхностных состояний при обрыве по короткой связи является условием Шокли. (Shockley Criterion: *In the 1D tight-binding model with alternating tunneling amplitudes, the edge state exists if the bond of the greater magnitude is broken at the boundary, $|t_1/t_2| > 1$*). Это условие есть, по сути, граничное условие в том смысле, что обрыв по короткой или по длинной связи регулируется данным отношением. Подчеркнем, что оно возникает при решении граничной задачи полубесконечного кристалла – кристалла с границей, которая может быть двух типов: разрыв по короткой связи (t_1) и разрыв по длинной (t_2).

Граничное условие может быть переписано в эквивалентном виде через число вращения (winding number), что и было сделано в [17] (формула (20)):

$$W = \frac{1}{2\pi i} \int_0^{2\pi} dk \frac{d}{dk} \ln t(k).$$
 Далее (см. формулу (22) в [17]), если число вращения $W = 1$ ($t_1 > t_2$ – обрыв по короткой связи), то поверхностные состояния существуют. Если $W = 0$ ($t_1 < t_2$ – обрыв по длинной связи), то поверхностные состояния отсутствуют. Значения инварианта привязаны к типу разрыва

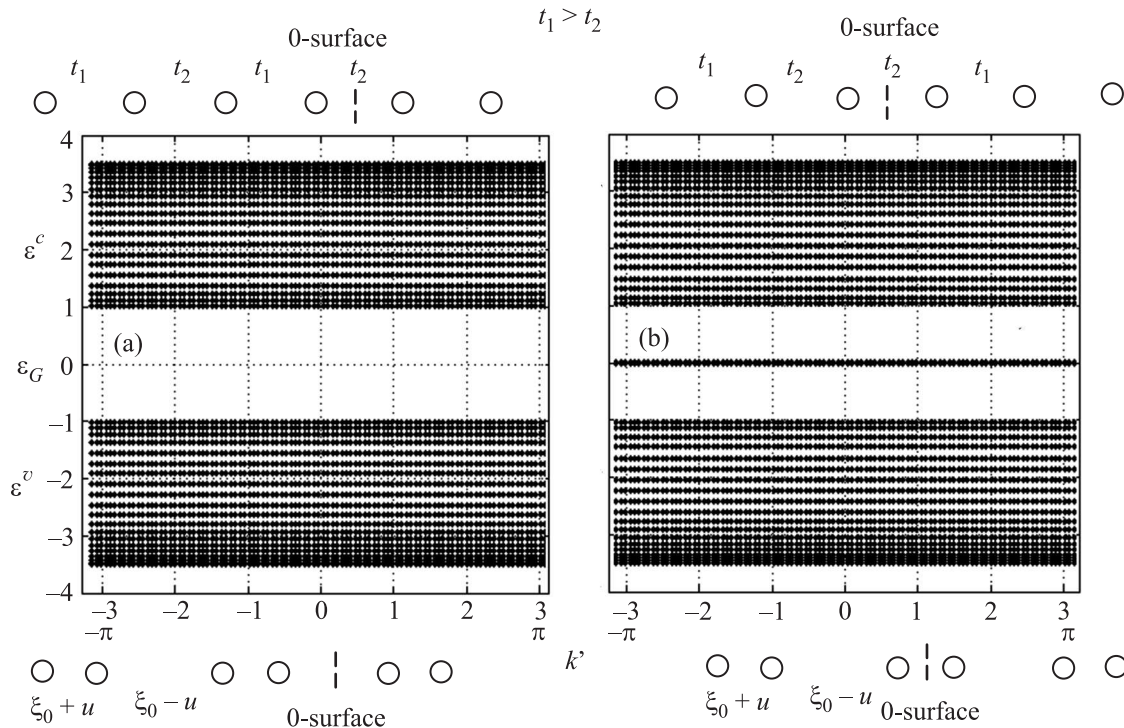


Рис. 3. (а) – Обрыв пайерлсовской цепочки по длинной связи. (б) – Обрыв пайерлсовской цепочки по короткой связи. Параметры $\xi_0 = 3.5$, $u_1 = 1.0$. Остальные параметры равны нулю. Число слоев 40

цепочки при образовании поверхности. Инвариант, хотя он и выражается через параметры перекрытия (в методе сильной связи других параметров нет, поэтому граничные условия будут выражаться через объемные матричные элементы метода сильной связи), явно содержит в себе информацию о типе разрыва кристалла: по короткой ($t_1 > t_2$) или по длинной ($t_1 < t_2$) связи. Значение инварианта привязано к типу поверхности. Поэтому имеется два разных его значения для разных типов поверхности. Любой инвариант, вычисленный только для данного объема, будет иметь только одно значение и выражаться через те же интегралы перекрытия. Одно, поскольку объем не несет в себе никакой информации о типе разрыва поверхности. Например, вычисляя значение W и обнаруживая, что $W = 1$ (при этом автоматически $t_1 > t_2$), затем делая утверждение о возникновении поверхностных состояний, при этом не говорить, по какой связи происходит разрыв цепочки, абсурдно. Разрывая цепочку по длинной связи, можно увидеть, что поверхностные состояния не образуются, и наоборот. Данный инвариант относится к цепочке с разрывом, и его значение зависит от типа разрыва.

Заключение. Критика Z_2 -классификации, основанной на соответствии объем–граница, не является самоцелью настоящей работы. Ее цель – обратить внимание на то, что хотя данная классификация ши-

роко цитируется, она не отвечает на вопрос о топологической устойчивости поверхностных состояний. Тем не менее интересно понять, в каком месте происходит сбой в логических рассуждениях. Приведем выдержки из обзора [3] (аналогичные утверждения встречаются и в других работах авторов обзора).

“...The difference, $N_R - N_L$, between the number of right and left moving modes, however, cannot change and is determined by the topological structure of the bulk states. This is summarized by the bulk-boundary correspondence, ...” ([3], с. 3049).

“...Depending on the details of the Hamiltonian near the edge there may or may not be states bound to the edge inside the gap. If they are present, however, then Kramers’ theorem requires they are twofold degenerate at the T invariant momenta $k_x = 0$ and π/a which is the as $-\pi/a$...” ([3], с. 3050).

Эти два утверждения, расположенные на двух соседних страницах, явно противоречат друг другу. Достаточно только объема (в котором нет поверхностных состояний) или граница тоже нужна? Объемный спектр, как отмечалось выше, не содержит в своей систематике поверхностных состояний. Все состояния, включая слабо возмущенные объемные и поверхностные, классифицируются по представлениям 7 пространственных кристаллических групп лент (2D-системы с одномерной границей или макроско-

пически широкие 2D-ленты с двумя одномерными границами). Объемный же спектр самой 2D-системы без границ классифицируется по представлениям 17 плоских кристаллических групп (2D). Конические особенности имеют место и в 2→1D- и 2D-системах (в последних есть и другие особенности [6]). Если 2D-система без границ – изолятор, то в запрещенной зоне нет никаких поверхностных состояний. При этом конические особенности, если они есть, лежат в области объемного спектра, а не в щели. Аналогичная ситуация имеет место и для 3D 3→2D-систем (трехмерные кристаллы с 2D-границей, см. приведенный выше контрпример). Путаница происходит из-за того, что авторы в обсуждениях неявно переходят от ситуации 2D к 2→1D (3D к 3→2D). Спектр, приведенный на рис. 3 (см. рис. 3 в обзоре [3]), является уже спектром 2→1D-системы, включая объемные заштрихованные состояния.

В некоторых работах также указывалось на важность учета граничных условий при рассмотрении появления поверхностных состояний и их типа. На примерах частных моделей [18] выясняется, что объемный топологический инвариант, основанный на соответствии объем–граница, не отвечает за появление поверхностных состояний, а отвечает лишь за тип состояний. Для нетривиального значения инварианта должно иметься *нечетное число пересечений* уровня Ферми (топологически устойчивые состояния типа рис. 2b), для тривиального – *четное* (неустойчивые состояния типа рис. 2a). Однако в общем случае и это не так. Приведенный пример показывает, что даже при тривиальном значении инварианта на разных гранях возможны состояния обоих типов (как с четным, так и с нечетным числом пересечений уровня Ферми). Кроме того, отметим пример с поверхностью Sb(111) (нетривиальный топологический изолятор) и Bi(111) (тривиальный изолятор), где при обрыве по короткой связи возникают состояния одного типа с нечетным числом пересечений уровня Ферми [10]. Инварианты различаются, а тип состояний одинаковый (с нечетным числом пересечений уровня Ферми). **Таким образом, как и утверждалось выше, Z_2 -классификация, основанная на соответствии объем–граница, не имеет отношения ни к появлению краевых состояний, ни к их типу.**

Во избежание недоразумений отметим, что классификация [19], основанная на первой гомотопической группе, относится ко всем состояниям (объем-

ным и поверхностным) в 2→1D-системах, а не в 2D-системах. При этом данная классификация только констатирует факт наличия двух типов состояний (типов соединений спектра в особых точках Γ и X одномерной ЗБ, которые имеют место без обращения к топологическим аргументам) и никак не связывает объемный спектр 2D с наличием или отсутствием поверхностных состояний в 2→1D-системе.

Авторы благодарят А. Ионову и С. Божко за предоставленную возможность ознакомиться с экспериментальными данными до публикации. Авторы также выражают благодарность неизвестному рецензенту за ссылку на работу [18].

1. C. Herring, Phys. Rev. **52**, 361 (1937).
2. Г. Л. Бир, Г. Е. Пикус, *Симметрия и деформационные эффекты в полупроводниках*, Наука, М. (1972).
3. M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
4. X.-L. Qi and Sh.-Ch. Zhang, Rev. Mod. Phys. **83**, 1057 (2011).
5. S. V. Meshkov and S. N. Molotkov, Surface Sci. **240**, 263 (1990).
6. С. Н. Молотков, Письма в ЖЭТФ **90**, 339 (2009).
7. С. Н. Молотков, Письма в ЖЭТФ **94**, 306 (2011).
8. C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 146802 (2005).
9. L. Fu and C. L. Kane, Phys. Rev. **76**, 045302 (2007); J. C. Y. Teo, L. Fu, and C. L. Kane, Phys. Rev. **78**, 045426 (2008).
10. С. Н. Молотков, В. В. Татарский, Поверхность **5**, 17 (1988).
11. С. Н. Молотков, Т. А. Потапова, Письма в ЖЭТФ **98**, 336 (2013).
12. R.-J. Slager, A. Mesaros, V. Juriycirc, and J. Zaanen, Nat. Phys. **9**, 98 (2013).
13. Б. А. Волков, Л. А. Фальковский, ЖЭТФ **85**, 2135 (1983); Е. А. Дорофеев, Л. А. Фальковский, ЖЭТФ **87**, 2202 (1984); Л. А. Фальковский, Письма в ЖЭТФ **45**, 291 (1987).
14. Б. А. Волков, О. А. Панкратов, ЖЭТФ **75**, 1362 (1978).
15. Б. А. Волков, О. А. Панкратов, Письма в ЖЭТФ **42**, 145 (1985).
16. С. Н. Молотков, В. В. Татарский, ФТТ **29**, 330 (1987).
17. S. S. Pershoguba and V. M. Yakovenko, Phys. Rev. B **86**, 075304 (2012).
18. L. Isaev, Y. H. Moon, and G. Ortiz, Phys. Rev. B **84**, 075444 (2011).
19. С. Н. Молотков, Письма в ЖЭТФ **98**, 466 (2013).