

## ПАМЯТИ ЛЕОНИДА ВЕНИАМИНОВИЧА КЕЛДЫША

## ОБЗОРЫ АКТУАЛЬНЫХ ПРОБЛЕМ

## Особенности туннельных процессов в полупроводниковых наноструктурах

П.И. Арсеев, В.Н. Манцевич, Н.С. Маслова, В.И. Панов

*Приведено описание экспериментальных данных, полученных методом сканирующей туннельной микроскопии/спектроскопии (СТМ/СТС), в которых наиболее ярко проявилось влияние неравновесных туннельных эффектов и реконструкции электронного спектра, а также дано их теоретическое объяснение, основанное на самосогласованном учёте неравновесного распределения электронов и изменения плотности электронных состояний в области туннельного контакта при протекании туннельного тока. Обсуждаются основные положения самосогласованной теории туннелирования, на которую могут опираться экспериментаторы в своих исследованиях и которая позволяет не только описывать многие эффекты, наблюдаемые в СТМ/СТС-экспериментах, но и даёт возможность предсказывать новые.*

**Ключевые слова:** туннелирование, сканирующая туннельная микроскопия/спектроскопия, локализованные состояния, диаграммная техника Келдыша, неравновесные процессы, межчастичное взаимодействие

PACS numbers: 05.60.Gg, 68.37.Ef, 73.40.Gk, **73.63. – b**

DOI: <https://doi.org/10.3367/UFNr.2017.01.038055>

### Содержание

1. Введение (1147).
2. Экспериментальные результаты (1148).
  - 2.1. Волны зарядовой плотности и индуцированная проводимость.
  - 2.2. Особенности электронной структуры чистых поверхностей полупроводников.
  - 2.3. Микроскопия и спектроскопия изолированных примесных состояний на поверхности полупроводников.
  - 2.4. Неравновесное взаимодействие примесных состояний.
  - 2.5. Неравновесные эффекты и многочастичное взаимодействие на d-орбиталях примесных атомов.
3. Особенности туннельных процессов в структурах пониженной размерности (1158).
4. Связанные состояния, отщеплённые от границ исходного спектра в результате туннельного взаимодействия (1159).
5. Появление индуцированной туннельной проводимости (1159).
6. Модель Андерсона для описания неравновесных примесных состояний с коррелированными электронами (1161).

7. Кулоновские корреляции экситонного типа и сингулярности туннельных характеристик (1163).
  8. Заключение (1167).
- Список литературы (1167).

*Посвящается Л.В. Келдышу — инициатору экспериментальных работ по сканирующей туннельной микроскопии/спектроскопии в МГУ им. М.В. Ломоносова и создателю неравновесной диаграммной техники, на основе которой появилось согласованное теоретическое описание туннельных процессов.*

### 1. Введение

Создание сканирующего туннельного микроскопа (СТМ) позволило использовать туннельный эффект для исследования атомной и электронной структуры поверхности и поверхностных наноструктур, а также дало возможность изучать локальные туннельные явления на уровне атомных масштабов. Уже в первых экспериментах по наблюдению реконструированной поверхности кремния "Si(111)-(7 × 7)" [1] было показано, что СТМ позволяет получать изображение поверхностных атомных структур с точностью до зарядового состояния отдельного атома. За многие годы применения и развития сканирующая туннельная микроскопия и спектроскопия (СТМ/СТС) стала одним из наиболее эффективных методов исследования локальных туннельных эффектов и изучения свойств поверхностных структур с атомным разрешением. Благодаря этому методу появилась возможность исследовать пространственное распределение примесных состояний, а также идентифицировать атомную примесь или дефект атомарного раз-

**П.И. Арсеев.** Физический институт им. П.Н. Лебедева РАН, Ленинский просп. 53, 119991 Москва, Российская Федерация; Национальный исследовательский университет "Высшая школа экономики", ул. Мясницкая 20, 101000 Москва, Российская Федерация  
E-mail: [ars@lpi.ru](mailto:ars@lpi.ru)

**В.Н. Манцевич, Н.С. Маслова, В.И. Панов.** Московский государственный университет им. М.В. Ломоносова, Ленинские горы 1, 119991 Москва, Российская Федерация  
E-mail: [vmantsev@gmail.com](mailto:vmantsev@gmail.com), [nsmaslova@gmail.com](mailto:nsmaslova@gmail.com), [panov@spmlab.phys.msu.ru](mailto:panov@spmlab.phys.msu.ru)

Статья поступила 12 декабря 2016 г.

мера по особенностям в спектре туннельной проводимости и его СТМ-изображения [2–7]. Вместе с тем особенности туннелирования в СТМ-контактах, масштабы которых сопоставимы с межатомными расстояниями, часто приводят к сильному искажению невозмущённой плотности состояний исследуемой системы. В равновесных условиях это может быть вызвано наличием в области контакта локализованных состояний (связанных, например, с обрывом периодической структуры на конце острия СТМ, примесными атомами или дефектами и др.) и их взаимодействием с состояниями непрерывного спектра в берегах контакта.

При СТМ/СТС-исследованиях даже отдельные локализованные состояния могут существенно влиять на получаемые в эксперименте СТМ-изображения и спектры туннельной проводимости, в отличие от экспериментов с макроскопическими туннельными контактами, для которых вклад в туннельный ток от локализованных состояний при их низкой плотности, как правило, пренебрежимо мал. При уменьшении размеров системы времена релаксации туннелирующих электронов становятся сопоставимыми с характерными временами туннелирования, поэтому распределение частиц становится существенно неравновесным. При наличии межчастичного взаимодействия такое распределение также приводит к изменению плотности состояний. По этим причинам исследование свойств поверхностных наносистем и локальных туннельных эффектов методом СТМ/СТС требует учёта различных процессов релаксации (рассеяние на примесях и неоднородностях на границе образца, электрон-фононное взаимодействие и т.д.). В результате часто возникают значительные трудности в интерпретации получаемых экспериментальных результатов при измерении локальной туннельной проводимости, а в ряде случаев и при анализе СТМ-изображений исследуемых наноструктур.

Таким образом, для адекватного описания результатов СТМ/СТС-измерений необходимо изменить стандартное представление о туннельных процессах. Становится неверным вывод о том, что туннельный ток всегда однозначно отражает исходную плотность состояний в исследуемом материале (электронной системе). В результате взаимодействия образца с близко расположенным остриём СТМ исходная плотность состояний искажается, появляются новые связанные состояния, играющие особо важную роль при изучении примесей, атомных дефектов, диэлектрических материалов (тонких плёнок), волн зарядовой плотности (ВЗП) на поверхности, поверхностных зон в полупроводниках и т.д. При этом может наблюдаться отличный от нуля туннельный ток при такой разности потенциалов, когда исходные зоны не перекрываются, а сам туннельный ток возникает за счёт присутствия в системе локализованных состояний.

Перечисленные особенности туннельных явлений в СТМ-контактах указывают на необходимость применения адекватной методики анализа результатов СТМ/СТС-исследований, основанной на современном представлении о туннельных процессах в наносистемах. Обычно при описании процессов туннелирования в СТМ/СТС используется стандартная формула для определения туннельного тока [8]:

$$I \sim 2\pi e \int d\omega T^2 v_k(\omega) v_p(\omega) [n_p(\omega) - n_k(\omega - eV)], \quad (1)$$

где  $T$  — амплитуда туннелирования,  $v_{k(p)}$  — плотность электронных состояний в берегах туннельного контакта,  $n_{k(p)}(\omega)$  — фермиевские функции распределения электронов в берегах контакта. Вместе с тем многочисленные СТМ/СТС экспериментальные исследования показали, что туннельные процессы в структурах нанометровых размеров и, следовательно, получаемые СТМ-изображения не всегда могут быть проанализированы в рамках стандартной теории туннельных явлений, основанной на использовании равновесной функции распределения туннелирующих частиц. Применение самосогласованной теории позволило наиболее полно описать процессы туннелирования в наносистемах с учётом релаксации.

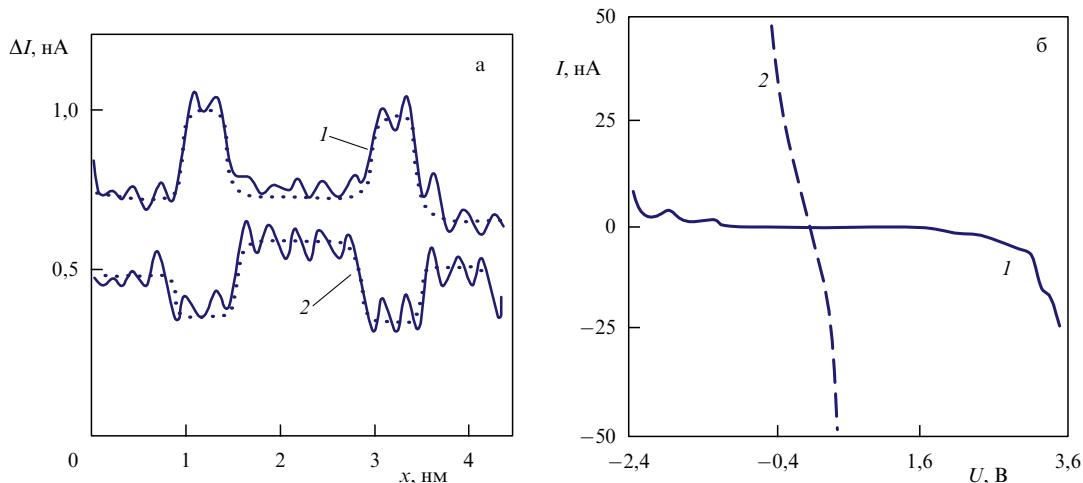
Релаксационные процессы в системах малых размеров уменьшают величину туннельного тока, что наблюдалось во многих экспериментах при понижении температуры и неизменных характеристиках туннельного контакта [9]. Наиболее общее выражение для туннельного тока должно учитывать различные механизмы релаксации в системе и возможные изменения чисел заполнения, а также влияние локализованных состояний. Эти вопросы будут рассмотрены в рамках теоретических моделей, описывающих туннельные процессы в контактах сверхмалых размеров.

В настоящем обзоре проводится анализ ряда ключевых результатов, полученных в СТМ/СТС-экспериментах, которые не находили полного объяснения в рамках стандартной теории туннелирования, основанной на использовании равновесной функции распределения туннелирующих частиц. Наблюдаемые в некоторых туннельных экспериментах эффекты анализируются исходя из современных представлений о туннелировании с учётом реконструкции спектра электронных состояний, влияния неравновесных процессов и наличия локализованных состояний. Представлены основные положения самосогласованной теории туннелирования, на которую могут опираться экспериментаторы в своих исследованиях и которая позволяет не только описывать многие эффекты, наблюдаемые в СТМ/СТС-экспериментах, но и даёт возможность предсказывать новые.

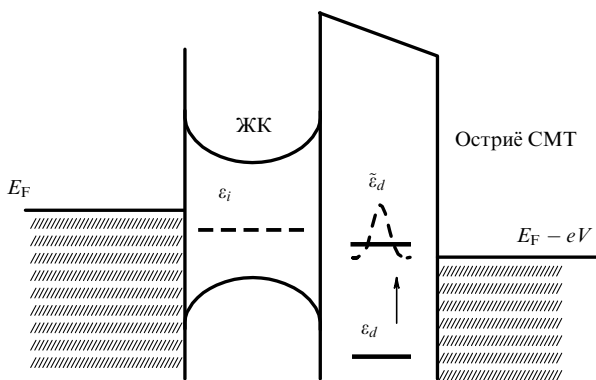
## 2. Экспериментальные результаты

### 2.1. Волны зарядовой плотности и индуцированная проводимость

Одним из экспериментальных результатов, полученных методом СТМ/СТС, который не удавалось однозначно объяснить в рамках традиционных представлений о процессах туннелирования, было наблюдение упорядоченной структуры поверхности молекулярной плёнки жидкокристаллического (ЖК) полимера [10]. Объектом исследования служил бислой ЖК-полимера гребнеобразного типа ПХА-10 толщиной 66 Å, нанесённый на атомно-гладкую проводящую подложку. На поверхности бездефектного участка молекулярной плёнки методом атомной силовой микроскопии (АСМ) было получено изображение рельефа с ярко выраженной двумерной решёткой со сторонами  $a \simeq b \simeq 11,9$  Å и длиной диагональю  $d \simeq 20$  Å. В то же время применение метода СТМ позволило выявить на токовых изображениях исследуемой поверхности периодическую сверхструктуру при разности потенциалов 5 мВ, соответствующую решётке с параметрами  $a^* \simeq b^* \simeq 25,5$  Å и  $d^* \simeq 20,2$  Å, превы-



**Рис. 1.** (а) Сечения изображений зарядовой сверхструктуры при смене знака туннельного напряжения (кривая 1 получена при  $V = +5$  мВ, кривая 2 при  $V = -5$  мВ). (б) Вольт-амперные характеристики при большом (1) и малом (2) расстояниях между остриём СТМ и поверхностью.  $x$  — расстояние вдоль поверхности.



**Рис. 2.** Схема туннельного перехода подложка–ЖК–барьер–остриё.  $E_F$  — уровень Ферми,  $\epsilon_i$  — резонансные уровни,  $\epsilon_d$  — энергия локализованного состояния острья,  $\tilde{\epsilon}_d$  — энергия связанного состояния.

шающими параметры решётки, полученной на АСМ-изображении. Токовое СТМ-изображение поверхности меняло фазу на противоположную при смене знака потенциала, приложенного к туннельному переходу (рис. 1а). Выпуклости и провалы на СТМ-изображении менялись местами, что можно было бы объяснить наличием на поверхности распределённого заряда типа волны зарядовой плотности (ВЗП). Кроме того, в изучаемой системе проводящая подложка–ЖК–остриё СТМ (рис. 2) были получены два типа вольт-амперных характеристик (ВАХ). При больших расстояниях между остриём и поверхностью на ВАХ наблюдалась широкая запрещённая область  $\sim 5$  эВ (рис. 1б), а уровень Ферми  $E_F$  лежал внутри запрещённой зоны.

При изменении приложенного напряжения в диапазоне от 1,5 В до 5 мВ, а также при уменьшении расстояния между остриём и поверхностью ЖК запрещённая область практически исчезала, а проводимость резко возрастала (не менее чем на два порядка). Детальный анализ особенностей электронной и атомной (молекулярной) структуры поверхности ЖК указывал на существование ВЗП на этой поверхности. Уровень Ферми при возникновении ВЗП находился внутри энергетической щели, ширина которой зависела от величины электрон-фононного

и межэлектронного взаимодействия, а минимальный период зарядовой сверхструктуры совпадал с длиной большей диагонали решётки, отвечающей АСМ-изображению рельефа. Вместе с тем требовалось прояснить условия, при которых возможно наблюдение ВЗП при комнатной температуре ( $kT = 25$  мэВ) и туннельном напряжении  $V \simeq 5$  мВ, а также объяснить резкое изменение проводимости системы. Условия наблюдения ВЗП требуют, чтобы величина энергетической щели превосходила  $kT$ , иначе тепловые флуктуации приводят к разрушению ВЗП. Следовательно, без локального взаимодействия электронных состояний проводящего острия СТМ и поверхности ЖК наблюдение ВЗП было бы невозможно.

Для объяснения наблюдаемых эффектов потребовалось привлечение теоретической модели, учитывающей наличие в области туннельного контакта локализованных состояний. В самом деле, на поверхности острия СТМ или ЖК могут существовать локализованные состояния с энергией  $\epsilon_d$ , лежащие значительно ниже  $E_F$ , обусловленные либо обрывом периодической структуры на конце острия, либо наличием дефектов и неоднородностей на поверхности. В результате сильной гибридизации электронных состояний острия и образца формируется связанное состояние с энергией  $\tilde{\epsilon}_d$ , лежащее в области энергетической щели, что вызывает изменение вида туннельной проводимости.

Более подробно самосогласованная теория туннелирования в наноструктурах с участием связанных локализованных состояний будет рассмотрена далее в разделе 5.

## 2.2. Особенности электронной структуры чистых поверхностей полупроводников

Метод СТМ/СТС используется преимущественно для изучения атомной и электронной структуры чистых поверхностей в условиях сверхвысокого вакуума, при котором наличие неконтролируемых адсорбированных слоёв не искажает исходную структуру поверхности. При комнатных температурах, когда релаксацией носителей можно пренебречь, а туннелирование осуществляется из заполненных состояний металлического острия СТМ, перестройка электронного спектра, как правило, мало влияет на СТМ-изображения периодических поверхност-

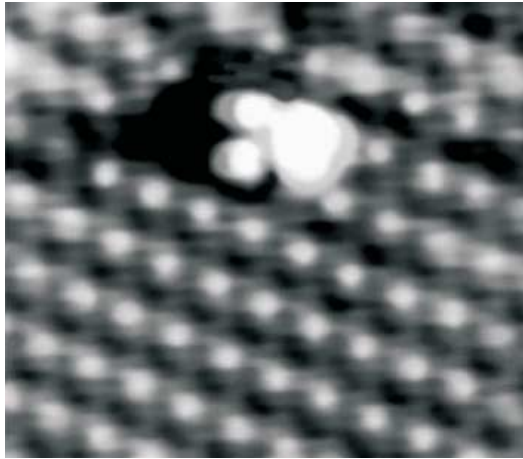


Рис. 3. STM-изображение подрешётки As и атомного кластера на поверхности InAs(110). Область сканирования  $4,4 \times 4,4$  нм. Туннельный ток 20 пА, напряжение на острие СТМ  $V = +0,5$  В.

ных структур. С понижением температуры возрастает роль релаксационных процессов, а скорость туннельных переходов может превышать скорость релаксации. В результате в области контакта возникает неравновесное распределение туннелирующих частиц, которое сильно меняет исходную плотность состояний и приводит к появлению неравновесных зарядов.

Подобный эффект наблюдался во многих экспериментах [9, 11]. Его детальное исследование было проведено, в частности, в работе [9], в которой методом СТМ/СТС изучена атомно-гладкая поверхность InAs(110) и область над атомным кластером, сформированным на этой поверхности (рис. 3). Измерения проводились при температуре жидкого гелия. Образцом служил монокристалл InAs с n-типом проводимости, допированный атомами Sn ( $5 \times 10^{17} \text{ см}^{-3}$ ).

В этих экспериментах над атомно-гладким участком поверхности наблюдалось увеличение ширины запрещённой зоны  $\Delta$  до 1,8 эВ по сравнению со значением для объёмного кристалла (0,43 эВ) (рис. 4). При этом уровень Ферми был смещён от края зоны проводимости в за-

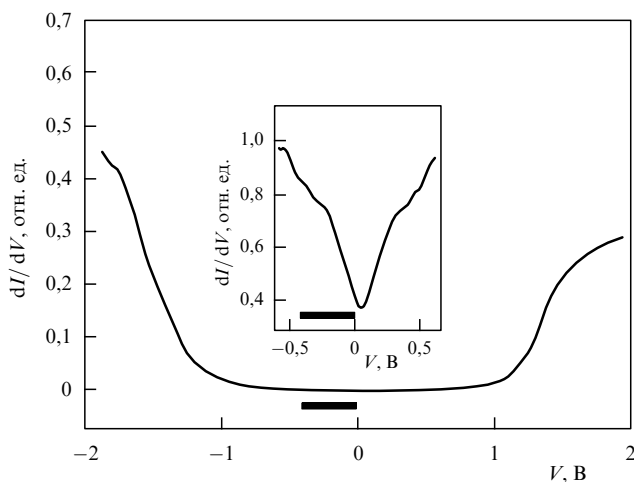


Рис. 4. Туннельная проводимость, измеренная над атомарно чистой поверхностью InAs(110) и над трёхатомным кластером (приведена на вставке). Толстой горизонтальной чертой отмечено положение ширины запрещённой зоны для объёмного кристалла, равной 0,43 эВ.

прещённую зону, несмотря на высокую степень легирования образцов. Измерения над атомным кластером выявили резкое уменьшение ширины запрещённой зоны по сравнению со значением, полученным над гладкой поверхностью. Для трёхатомного кластера она соответствовала 0,40 эВ. В другом эксперименте для шестиатомного кластера эта величина равнялась 0,55 эВ. Однако и в этих случаях уровень Ферми находился в запрещённой зоне (см. рис. 4).

Наблюдаемые экспериментальные результаты могут быть связаны с тем, что локализованный на острие неравновесный заряд вызывает дополнительный изгиб зон на поверхности полупроводника. Над кластером локализованный на острие СТМ заряд может уменьшаться или даже компенсироваться неравновесным зарядом противоположного знака. В результате измеряемая над кластером ширина щели уменьшается.

### 2.3. Микроскопия и спектроскопия изолированных примесных состояний на поверхности полупроводников

Особенности применения СТМ для изучения атомных примесей и дефектов были продемонстрированы при первых исследованиях зарядовых состояний донорных атомов Si на поверхности GaAs [12]. В этой работе СТМ-изображение положительно заряженной донорной примеси Si наблюдалось на глубине до двух нанометров под поверхностью. Изображения атомов, полученные при комнатных температурах, проявлялись в виде выпуклостей на поверхности радиусом  $\sim 2,5$  нм. Их высота зависела от напряжения, приложенного к туннельному переходу СТМ. При анализе этих изображений предполагалось наличие дополнительного изгиба зон, вызываемого кулоновским потенциалом зонда СТМ.

Возможности применения СТМ для изучения примесных состояний при низких температурах были продемонстрированы на примере исследования атомов Te [13]. Полученные в этой работе результаты указывали на необходимость учёта влияния неравновесных эффектов, связанных с наличием локализованных состояний. Исследовались монокристаллы GaAs, допированные атомами Te (концентрация  $5 \times 10^{17} \text{ см}^{-3}$ ). Теллур является донорной примесью замещения в подрешётке мышьяка, который теряет пять электронов, отдаваемых в валентную зону после связывания в кристаллической решётке GaAs. Его энергия ионизации составляет  $\sim 6$  мэВ, что соответствует формированию мелкого донорного уровня в запрещённой зоне GaAs, а теоретическая оценка радиуса локализации для внешнего s-электрона даёт величину  $\sim 7$  нм. Следовательно, изображение атома Te, замещающего атом As в узле решётки, должно быть сферически симметричным с радиусом, охватывающим  $\sim 15$  периодов решётки GaAs. СТМ-изображения атомов в верхних подповерхностных слоях наблюдались в виде пятен с радиусом локализации около 4 нм (рис. 5). Поскольку СТМ-изображение определяется внешними 5s-орбиталями Te, которые имеют дополнительный электрон по отношению к полупроводниковой матрице, а в изображение дают вклад нелокализованные электроны проводимости, размер изображения должен уменьшаться по мере увеличения глубины залегания атомов под поверхностью.

При положительных значениях потенциала, приложенного к зонду СТМ, изображения атомов окружены осцилляциями Фриделя, возникающими вследствие эк-

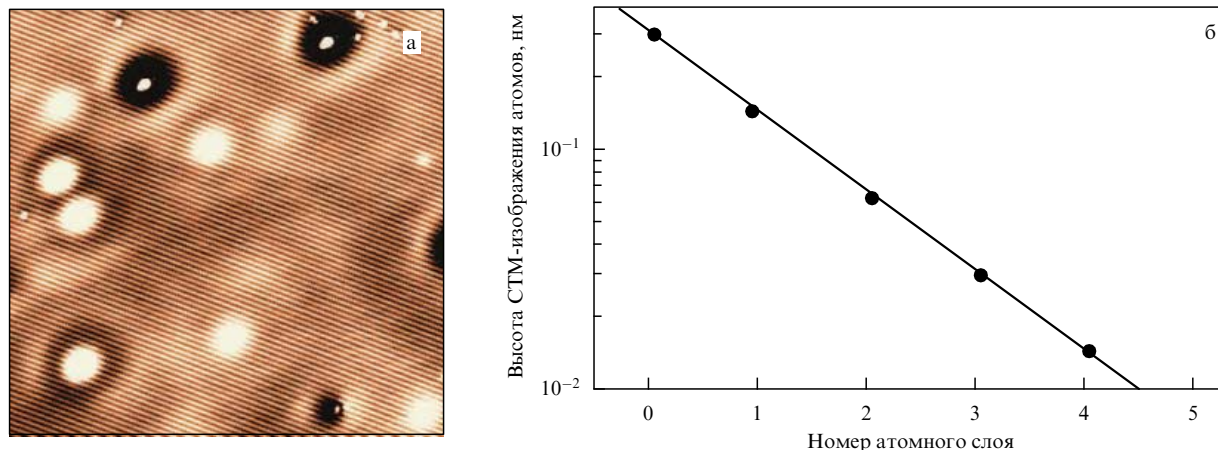


Рис. 5. (а) STM-изображения атомов Te на поверхности GaAs(110), потенциал зонда +1,5 В, туннельный ток 60 пА. Размер области сканирования  $40 \times 40$  нм. (б) Зависимость амплитуды изображений атомов Te от глубины их залегания.

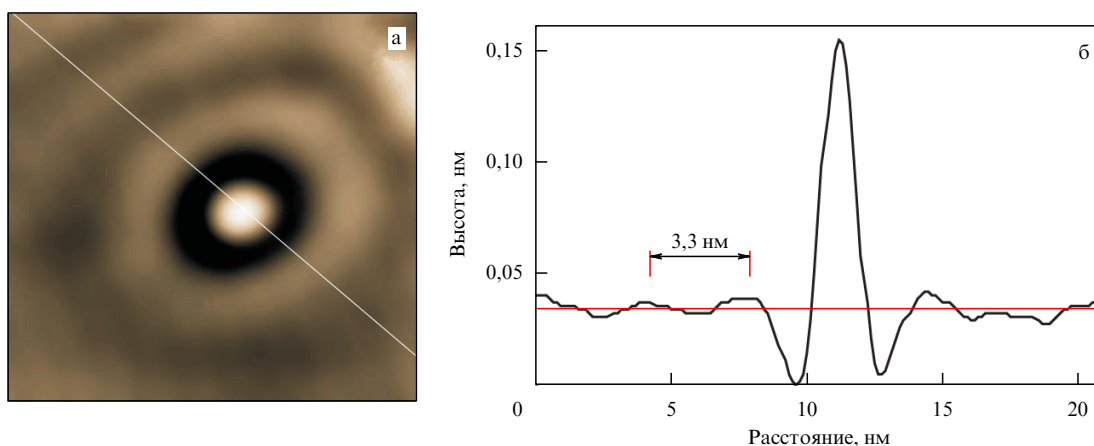


Рис. 6. (а) STM-изображение атома Te, полученное при напряжении на зонде STM +1,5 В. (б) Рельеф STM-изображения вдоль линии, показанной на рис. а.

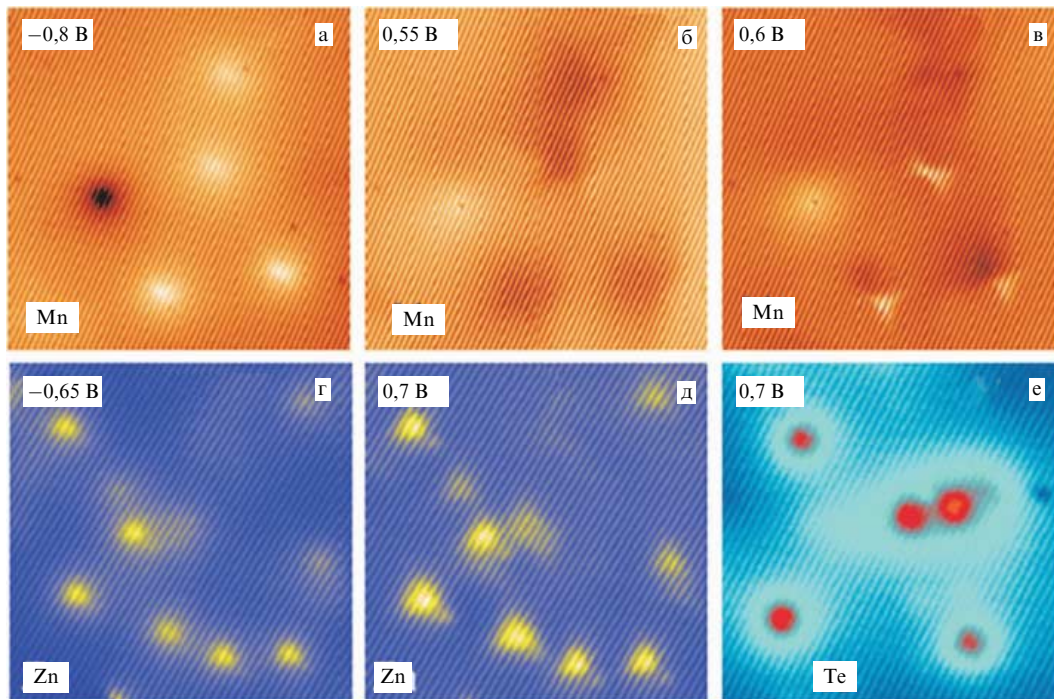
ранирования электронами заряда, локализованного на примесном атоме Te (рис. 5а).

На рисунке 6 показано STM-изображение атома Te в первом подповерхностном слое, которое окружено осцилляциями с периодом 3,3 нм, что не отвечает оценке, получаемой из стандартной модели экранирования для объёмного материала. На основании этой модели период осцилляций должен быть равен половине фермиевской длины волны  $\lambda_F/2 \approx 10,5$  нм. Однако влияние неравновесных эффектов и наличие локализованных состояний в области туннельного контакта модифицируют распределение электронной плотности вблизи примесного атома, вызывают изменение локального изгиба зон и изменяют период осцилляций Фриделя.

STM является единственным методом, позволяющим определять форму, симметрию и пространственную локализацию примесных состояний в полупроводниковой матрице. В качестве характерного примера на рис. 7 приведены STM-изображения примесных атомов Te, Zn, Mn на поверхности InAs(110) и в подповерхностных слоях для различных значений приложенного напряжения. Изображения различаются формой, симметрией и радиусом локализации примесного состояния. В частности, большинство изображений атомов Zn имеют фор-

му равнобедренного треугольника (рис. 7д) для незаполненных состояний и циркулярно симметричную форму для заполненных состояний (рис. 7г). Изображения атомов Te циркулярно симметричны (рис. 7е) независимо от приложенного напряжения и глубины залегания. STM-изображение атомов Mn изменяется сложным образом в зависимости от туннельного напряжения и от подповерхностного слоя, в котором расположен атом.

Амплитуда STM-изображений (при неизменном значении приложенного потенциала) для большинства изученных примесей экспоненциально уменьшается с увеличением глубины залегания примеси (см. рис. 5). При больших концентрациях примесей, когда среднее расстояние между ними не сильно превышает радиус локализации примесного состояния, такая зависимость подвержена влиянию объёмного распределенного заряда. В эксперименте удаётся получать изображения под поверхностью на глубине до десяти атомных слоёв, что позволяет использовать STM в качестве метода туннельной подповерхностной нанотомографии. Для сферически симметричных изображений примесных атомов (например, для атомов Te) можно определить зависимость радиуса локализации  $R$  примесного состояния от глубины подповерхностного слоя  $N$ , в котором расположен

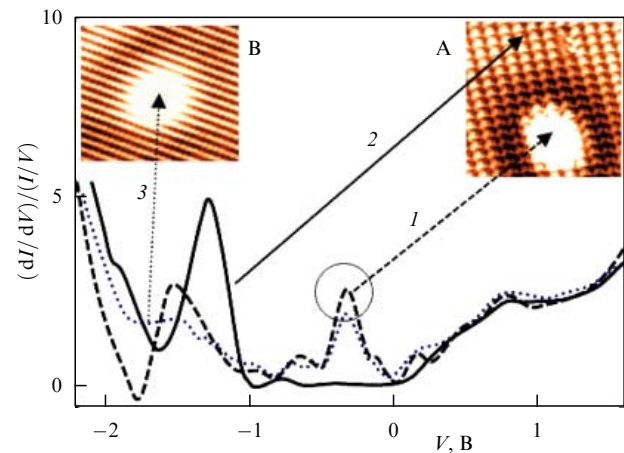


**Рис. 7.** STM-изображения примесных атомов на поверхности и в подповерхностных слоях InAs(110): (а–в) Mn, (г, д) Zn, (е) Te. Значения туннельного напряжения, при которых проводились измерения, указаны на рисунках. Размер кадра: (а–в)  $40 \times 40$  нм, (г, д)  $30 \times 30$  нм, (е)  $27 \times 27$  нм.

примесный атом. Однако и в этих случаях требуется учитывать взаимодействие локализованных состояний остря и примесного атома, а также влияние релаксационных процессов и неравновесных эффектов.

Из анализа STM-изображений следует, что различные типы атомных дефектов не всегда могут быть корректно идентифицированы путём простого сопоставления их изображений, которые зависят от параметров туннельного перехода. В этих случаях ключевые свойства дефектов, их различие или идентичность, а также особенности их взаимодействия могут быть определены на основании анализа спектров туннельной проводимости. На рисунке 8 приведены спектры нормированной туннельной проводимости, измеренные для примесных состояний Te, в первом и втором подповерхностных слоях. Ширина запрещённой зоны в этих измерениях зависела от относительного положения примесного атома и остря зонда STM. Границы запрещённой зоны вблизи примесного атома были сдвинуты на  $0,1-0,2$  эВ относительно значения, получаемого над чистой поверхностью. Следует отметить, что если поперечный размер туннельного контакта сопоставим или меньше радиуса локализации примесного состояния, то туннельный ток и туннельная проводимость в такой системе могут определяться в основном этим локализованным состоянием и его взаимодействием с STM-зондом.

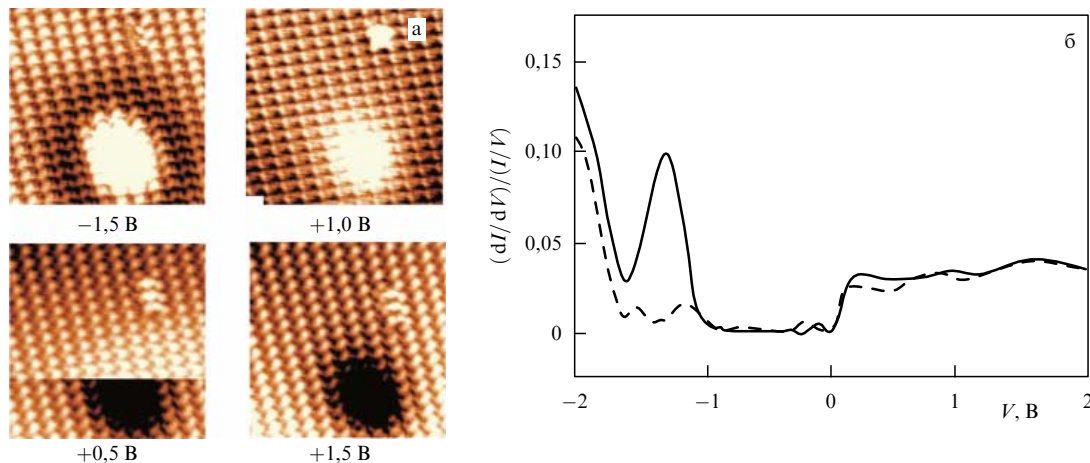
На рисунке 8 примесные атомы Te характеризуются пиками в спектре нормированной туннельной проводимости при напряжении  $\sim -0,7$  В, которые отсутствуют при измерениях над чистой поверхностью. Эти пики в основном сохраняют своё положение в спектрах туннельной проводимости при разных глубинах залегания атомов Te под поверхностью. Над чистой поверхностью (на расстоянии более 5 нм от примесного атома) в локальной туннельной проводимости формируется пик в диапазоне



**Рис. 8.** STM-изображение атомов Te в первом (А) и втором (В) подповерхностных слоях GaAs(110) и зависимость нормированной туннельной проводимости от напряжения на образце: 1 и 3 над атомами А и В соответственно, 2 над бездефектным участком поверхности. Кругом обведены пики, сохраняющие своё положение в спектрах туннельной проводимости для разных глубин залегания атомов Te.

напряжений от  $-1,5$  В до  $-1$  В, связанный с локализованным состоянием зонда. Положение этого пика и его амплитуда зависят от расстояния между примесным атомом и остриём зонда STM.

В спектрах туннельной проводимости примесных состояний на поверхности полупроводников проявляются общие закономерности. На графиках туннельных спектров для всех изученных примесей и для чистой поверхности полупроводников запрещённая зона сдвинута, и её ширина отличается от ширины объёмной зоны, как это было показано выше. Этот эффект связан с из-



**Рис. 9.** STM-изображение примесного атома Te (а) при различных напряжениях на образце. Скачкообразное изменение контраста изображения при +0,5 В соответствует скачкообразному изменению энергии локализованного состояния (б) на острие зонда STM. На рисунке б сплошной и штриховой линией обозначены измеренные спектры нормированной туннельной проводимости над бездефектным участком поверхности до и после скачка в контрасте изображения.

гибом зон, индуцированным зарядом, возникающим на локализованном состоянии примеси и/или на острие STM вследствие конечного времени релаксации туннелирующих электронов.

На кривых туннельной проводимости, измеренных над примесными атомами (независимо от знака их заряда), присутствуют пики в запрещённой зоне и на её краях. Такое поведение туннельной проводимости обусловлено кулоновским взаимодействием локализованных зарядов, которое изменяет начальное положение уровней примесного состояния относительно границ запрещённой зоны. Локальная плотность состояний в окрестности примесного атома существенно меняется и в глубине валентной зоны. В частности, для донорных примесей Te и Si наблюдается уменьшение плотности состояний при энергиях порядка 2 эВ, а для акцепторной примеси Zn плотность состояний в валентной зоне увеличивается.

Измерение туннельной проводимости непосредственно в области локализации атома примеси показало, что для каждого примесного состояния (по крайней мере, в полупроводниковой матрице) спектр локальной туннельной проводимости имеет свои специфические особенности, присущие только данному типу атомов. Это даёт возможность разработать методику STM/СТС-измерений примесного состава, опирающуюся на выявление в спектрах туннельной проводимости характерных деталей, отвечающих определённому типу примесного состояния, и позволяет идентифицировать примесные атомы по их спектральным особенностям [14].

Влияние локализованного состояния зонда STM на изображения примесных состояний и атомных дефектов ярко проявляется при наблюдении скачкообразного изменения контраста STM-изображений примесных атомов в процессе сканирования поверхности по растру [9]. Для этой цели был проведён эксперимент, в котором исследовалось влияние изменения зарядового состояния острия STM на изображения локализованных примесных состояний на поверхности полупроводника. На рисунке 9 показаны STM-изображения одного и того же примесного атома Te, полученные для разных напряжений, приложенных к образцу. При отрицательном потен-

циале на образце электроны, находящиеся в приповерхностной области, экранируют заряд примесного атома, индуцируя осцилляции Фриделя вокруг атома. При положительном потенциале на образце такие осцилляции отсутствуют.

В некоторый момент при сканировании вдоль поверхности контраст изображения примесного атома изменяется скачком. Светлое пятно в области локализации атома меняется на тёмное. При этом изображения поверхностных атомных рядов справа и слева от изображения примесного атома не претерпевают никакого скачка или разрыва. Дальнейшее увеличение положительного напряжения не меняет контраста изображения. После скачкообразного изменения контраст STM-изображения возвращается к первоначальному, если изменить знак приложенного напряжения. Эти факты указывают на то, что различия в контрасте STM-изображений не могут быть результатом "двойного" острия STM или любой другой механической нестабильности, а вызваны скачкообразным изменением локализованного заряда на острие.

Ключевое влияние локализованных состояний острия зонда на STM-изображения проиллюстрировано двумя кривыми нормированной туннельной проводимости на рис. 9б. Оба измерения проведены на атомно-гладких, бездефектных участках поверхности до и после скачкообразного изменения контраста. Пик в туннельной проводимости вблизи края запрещённой зоны при -1,2 В исчезает после скачка на STM-изображении. Это означает изменение энергии локализованного состояния острия STM, которое в свою очередь приводит к модификации локализованного на острие неравновесного заряда вследствие изменения вклада в туннельный ток неравновесных электронов. Изменение заряда на острие приводит к локальному изгибу зон на поверхности в области контакта и вызывает смену STM-изображения примесного состояния.

Прецизионные методы STM/СТС позволяют изучать детали формирования пространственной и энергетической структуры изолированных оборванных связей на поверхности и зарядовых состояний орбиталей примесных атомов в глубине полупроводника [15]. В работе [16]

получены СТМ-изображения пространственной локализации оборванной связи примесного атома Сг, возникающей при замещении As-узла в решётке InAs на поверхности, а также подповерхностной структуры зарядового состояния орбиталей примесных атомов Sn и Mn. В верхнем атомном слое InAs обнаружена область повышенной электронной плотности размером 0,5–0,6 нм вблизи атома Сг. Её форма соответствует пространственной локализации неспаренного электрона на гибридной орбитали оборванной связи атома Сг с отсутствующим (верхним) атомом In. Возмущающий потенциал, вносимый примесным атомом переходного металла, ограничен размерами одного периода кристаллической решётки. Обнаруженная вытянутая форма электронного облака и её зависимость от туннельного напряжения демонстрируют пространственную направленность, характерную для гибридной орбитали с участием d-электронного состояния.

Оценка характерной величины энергии кулоновского отталкивания локализованных электронов такого состояния даёт значение, сопоставимое с шириной запрещённой зоны, что приводит к наблюдаемому в спектрах туннельной проводимости сдвигу энергетических уровней глубоких примесных состояний.

#### 2.4. Неравновесное взаимодействие примесных состояний

Примесные состояния могут сильно изменять локальную электронную структуру на поверхности и границах раздела полупроводников. Если расстояние между примесями сопоставимо с их радиусом локализации, то возникают достаточно сильные корреляционные эффекты, которые изменяют электронные свойства такой системы и оказывают существенное влияние на её туннельную проводимость. Эти эффекты, обусловленные неравновесным взаимодействием примесных состояний, радикально влияют на туннельную проводимость и туннельный ток в области локализации каждого взаимодействующего атома при изменении приложенного потенциала. В такой системе туннелирование над каждым атомом (примесным состоянием) может неоднократно "включаться" или "выключаться" в зависимости от приложенного потенциала, увеличивая либо уменьшая величину туннельного тока, протекающего через каждое примесное состояние.

Первое наблюдение локальных эффектов неравновесного взаимодействия двух примесных состояний было проведено методом СТМ/СТС в работе [17]. В эксперименте было изучено пространственное распределение нормированной туннельной проводимости на поверхности GaAs(110) вблизи взаимодействующих примесных атомов кремния для различных напряжений на острие зонда СТМ. Для получения пространственного распределения туннельной проводимости вдоль всей поверхности исследуемой наноструктуры применялась методика дискретного измерения производной от туннельного тока по напряжению. Методика позволила измерить спектр туннельной проводимости на площади менее 0,1 нм<sup>2</sup>. В экспериментах контролируемым образом было получено расщепление электронных состояний с противоположными спинами до энергий порядка 1 эВ. Традиционными методами такое расщепление не может быть достигнуто даже в сверхсильных полях. Исследуемым образцом служил монокристалл GaAs, легирован-

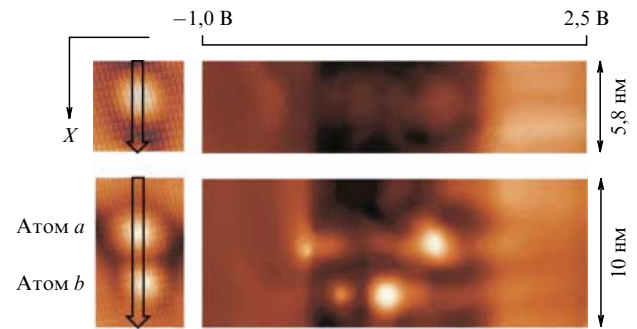


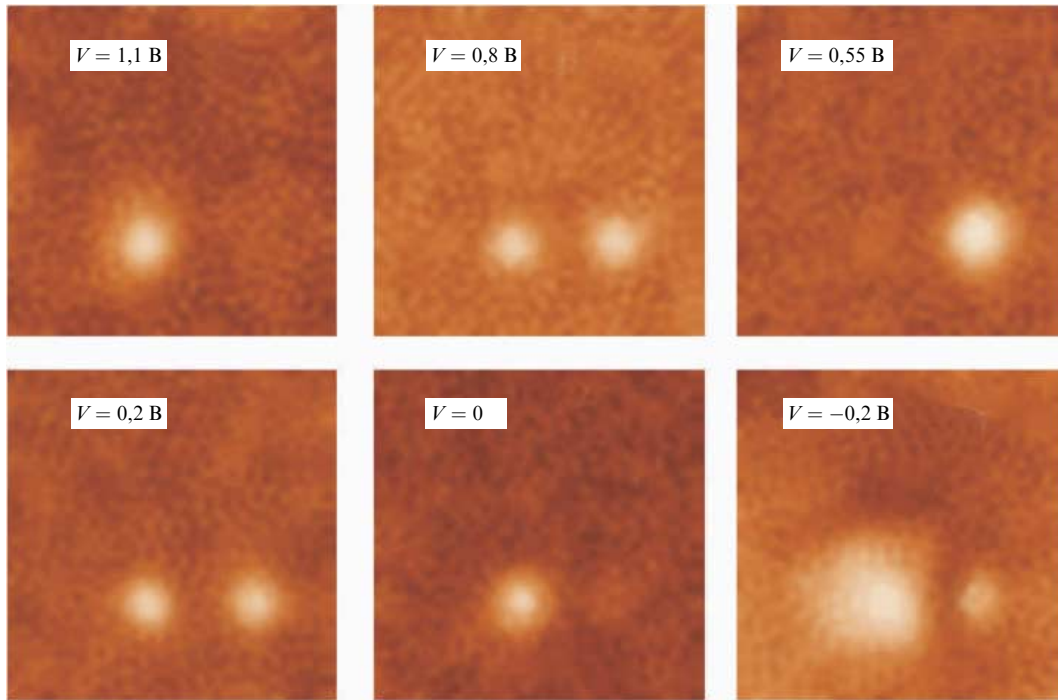
Рис. 10. В верхней части рисунка приведены СТМ-изображения изолированного атома (слева) и карта распределения нормированной локальной туннельной проводимости вблизи примесного атома, измеренная вдоль направления  $X$ , показанного стрелкой. В нижней части рисунка приведены СТМ-изображения взаимодействующих атомов  $a$  и  $b$  (слева) и карта распределения нормированной локальной туннельной проводимости вблизи примесных атомов, измеренная вдоль направления  $X$ .

ный взаимно компенсирующими примесями Si и Zn с концентрацией  $5 \times 10^{18}$  и  $2 \times 10^{19}$  см<sup>-3</sup> соответственно. После раскалывания кристалла на поверхности скола (110) был выбран участок с двумя идентичными примесными атомами Si, разделёнными расстоянием 3 нм, которое соответствует удвоенному радиусу локализации примесного состояния Si. Измерения проводились при температуре 4,2 К.

Взаимодействие примесных состояний изучалось путём измерения туннельной проводимости ( $dI/dV$ )/( $I/V$ ) на участке поверхности размером  $10 \times 10$  нм вблизи взаимодействующих атомов Si. Для сопоставления были проведены аналогичные измерения на изолированном атоме Si. На рисунке 10 приведена зависимость нормированной туннельной проводимости от приложенного напряжения в диапазоне от +2,5 до -1 В, измеренная вдоль направления  $X$ , указанного стрелками на топографическом СТМ-изображении для одного изолированного атома и для двух взаимодействующих атомов примеси  $a$  и  $b$ . Следует отметить, что вклад от других легирующих примесей вблизи исследуемых примесных атомов Si наблюдался лишь в виде распределённого неоднородного кулоновского потенциала в кристалле. Этот потенциал вызывал асимметрию начального состояния взаимодействующей пары атомов Si, наблюдаемую в различии СТМ/СТС-изображений этих атомов при нулевом напряжении на переходе.

В экспериментально наблюдаемом пространственном распределении локальной туннельной проводимости видно двукратное "включение" и "выключение" каждого состояния, связанного с атомами  $a$  и  $b$ , при изменении туннельного напряжения. В окрестности каждого из этих атомов после "включения" избыточная туннельная проводимость существует в интервале напряжений 0,6–0,7 В, что много больше ширины уровня локализованного состояния. Переход из одного состояния в другое для атомов  $a$  и  $b$  происходит при изменении напряжения в интервале порядка 0,15 В, что сопоставимо с шириной уровня энергии локализованного состояния. В то же время в окрестности изолированного примесного атома Si отсутствуют характерные особенности, которые отчётливо проявляются в области двух взаимодействующих атомов.





**Рис. 11.** Изображение нормированной туннельной проводимости на участке поверхности GaAs(110) размером  $10 \times 10$  нм вблизи взаимодействующих примесных атомов кремния, отделённых друг от друга расстоянием 3 нм, для ключевых напряжений на острие СТМ. Изображения взяты из рис. 10 для выбранных значений туннельного напряжения. Приведённые распределения туннельной проводимости не нормированы, поэтому контраст изображений иллюстрирует лишь относительное соотношение между минимальной и максимальной величиной плотности состояний.

На рисунке 11 изображено пространственное распределение нормированной туннельной проводимости вблизи взаимодействующих атомов  $a$  и  $b$  для напряжений на острие СТМ, соответствующих ключевым состояниям этих атомов. На последовательности этих изображений приведена зависимость от энергии областей взаимного перекрытия электронной плотности (получаемая кадр за кадром), из которой определяются моменты включения и выключения взаимодействия и симметрия электронных состояний. Величина перекрытия состояний зависит от приложенного потенциала и, следовательно, может управляться внешним электрическим полем.

Карта туннельной проводимости (см. рис. 11) позволяет проанализировать поведение локальной туннельной плотности состояний вблизи каждого атома и даёт возможность определить её связь с состояниями взаимодействующих электронов с данной энергией.

Объяснение этих эффектов можно дать в рамках модели, аналогичной модели Андерсона [18], в которой учитывается кулоновское взаимодействие локализованных электронов и влияние туннельного напряжения на положение энергетических уровней примесных атомов (см. далее раздел 6, рис. 14). Взаимодействие примесных состояний с состояниями непрерывного спектра предполагается не слишком малым — таким, что изменение энергии примесных состояний за счёт кулоновского взаимодействия локализованных зарядов определяется средними числами заполнения при фиксированном значении приложенного напряжения.

Эта ситуация отличается от эффектов, обусловленных кулоновской блокадой, связанной с дискретностью чисел заполнения локализованных состояний. Действи-

тельно, кулоновская блокада при наличии как внутриатомного кулоновского взаимодействия, так и межатомного вызывает появление пиков в туннельной проводимости при изменении напряжения на контакте, ширина которых порядка ширины уровня локализованного состояния (и не превышает 0,1 эВ). Однако наличие туннельного взаимодействия между примесными атомами может приводить к перераспределению локализованного заряда между атомами и к эффекту затягивания уровней энергии локализованных состояний друг относительно друга в некотором интервале приложенного напряжения. В результате для адекватного описания полученных экспериментальных данных необходимо учесть следующие факторы.

Появление избыточной локальной туннельной проводимости в окрестности примесного атома наблюдается, если  $|\varepsilon_{a(b)}(V) - E_F| < \Gamma$  ( $\Gamma$  — уширение локализованного уровня атома,  $\varepsilon_{a(b)}$  — уровни энергии примесных атомов). При этом если приложенное напряжение превосходит ширину уровня  $|eV| > \Gamma$ , то локальная туннельная проводимость может возрастать как при прохождении энергии локализованного состояния через уровень Ферми образца  $E_{SF}$ , так и через уровень Ферми острия  $E_{TF} = E_{SF} - eV$ .

Кулоновское взаимодействие локализованных электронов может описываться в самосогласованном приближении среднего поля в модели Андерсона. Это приближение хорошо применимо, когда энергия кулоновского взаимодействия локализованных электронов на примесных уровнях  $U_{a(b)}$  не сильно превышает критическое значение  $U_{cr}$ . Таким образом, кулоновское взаимодействие локализованных электронов в некотором интервале значений приложенного напряжения может приводить к

возникновению двух различных по энергии состояний для электронов с противоположными спинами на отдельном примесном атоме.

Наличие взаимодействия между атомами приводит к более сложному закону распределения локализованных зарядов в области контакта СТМ. Происходит значительное перераспределение локализованного заряда между атомами так, что возможна стабилизация уровней энергии одного из атомов вблизи уровня Ферми образца или острия и затягивание энергетических уровней друг относительно друга.

Учёт приведённых факторов позволяет объяснить наблюдаемые эффекты и даёт возможность построить теоретические модели, предсказывающие нетривиальные эффекты взаимодействия локализованных состояний, образуемых атомами, атомными дефектами, квантовыми точками и др., которые будут рассмотрены в разделе 6. В этих экспериментах впервые продемонстрирована практическая возможность управления электронной структурой системы взаимодействующих атомов. В такой системе для каждого взаимодействующего атома резонансный канал туннельного тока может неоднократно включаться или выключаться в зависимости от приложенного потенциала. Эти эффекты могут стать основой для реализации новых элементов полупроводниковой нанoeлектроники, основанной на использовании изолированных и взаимодействующих примесных атомов в полупроводниковой матрице.

### 2.5. Неравновесные эффекты и многочастичное взаимодействие на d-орбиталях примесных атомов

Появление неравновесных зарядов в области туннельного контакта, который содержит примесные атомы, может сильно изменять исходную локальную плотность состояний и спектры туннельной проводимости при наличии кулоновского взаимодействия. Такие изменения возможны и в результате изменения туннельных амплитуд, обусловленного резким включением кулоновского потенциала, возникающего за счёт присутствия неравновесных чисел заполнения на примесном атоме. Эта ситуация аналогична эффекту возникновения степенных сингулярностей в спектрах рентгеновского поглощения металлов [19]. Сингулярные особенности могут проявляться в СТМ-экспериментах, если скорость туннельных переходов с уровня энергии глубокого примесного атома в металлическое острие зонда СТМ значительно превосходит скорость релаксации неравновесного распределения электронных чисел заполнения на примесном атоме, и уровень энергии примесного состояния оказывается опустошённым. В результате происходит резкое включение кулоновского взаимодействия между заряженной примесью и электронами проводимости, приводящее к изменению величины туннельной амплитуды.

Отклик электронов проводимости в металлическом острие зонда СТМ на резкое включение кулоновского потенциала принципиально отличается от одночастичной картины. В этом случае можно ожидать формирования степенных сингулярностей на вольт-амперных характеристиках, когда величина приложенного напряжения приближается к величине энергии уровня примесного атома. Впервые попытка описать возникновение сингулярности на вольт-амперных характеристиках была предпринята в работе [20], однако авторы ограничились

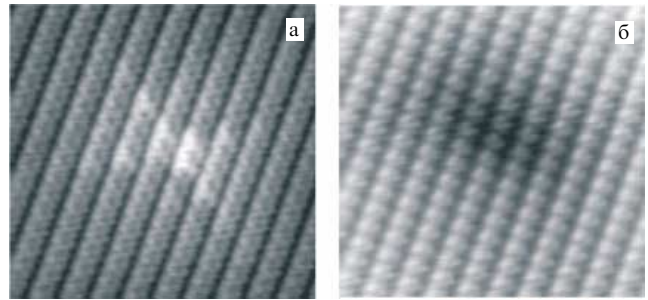


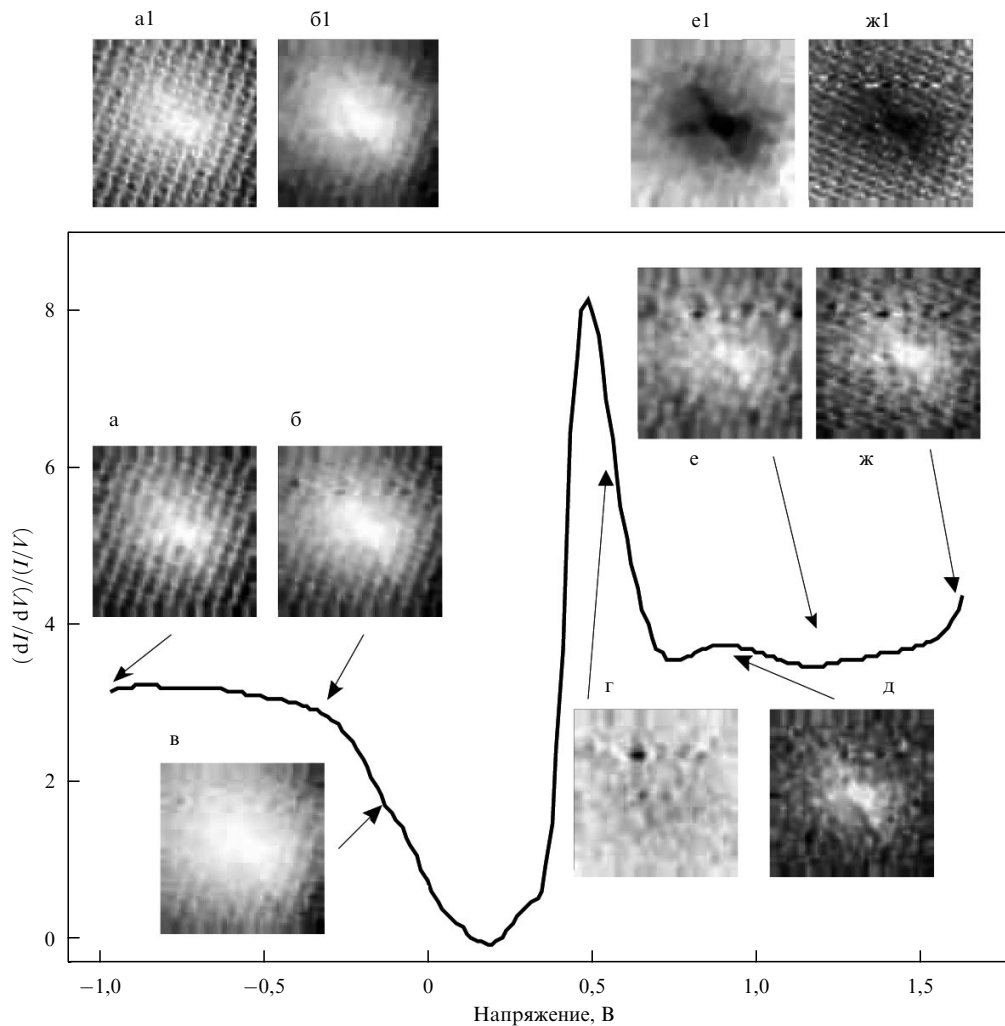
Рис. 12. СТМ-изображение примесного состояния Mn на поверхности InAs(110) размером  $5 \times 5$  нм при напряжении на образце  $V_s = -0,5$  В (а) и  $V_s = +0,7$  В (б).

качественными рассуждениями о перенормировке туннельной амплитуды, используя известный результат [19], что привело к ошибочным заключениям, о чём будет сказано подробнее в разделе 7.

Похожие эффекты наблюдались при СТМ-исследовании примесных состояний Cr и Mn на поверхности InAs(110) [21]. Сингулярные особенности наиболее ярко проявились при низкотемпературных СТМ/СТС-исследованиях примесных атомов Mn. На рисунке 12 представлены СТМ-изображения рельефа поверхности вблизи допирующего атома Mn при различных значениях туннельного напряжения. Для отрицательного напряжения  $V_s$  на образце в области локализации примесного состояния Mn наблюдалось светлое изображение (выпуклость) крестообразной формы, которое переходило в тёмное овальное изображение (провал) при положительном напряжении  $V_s$ .

Зависимость нормированной туннельной проводимости от напряжения и токовые СТМ-изображения в области локализации примесного состояния Mn показаны на рис. 13. При отрицательном напряжении на образце в диапазоне  $-1 < V_s < -0,5$  В сохраняется крестообразное изображение, которое соответствует электронной плотности d-орбитали, гибридизированной с состояниями решётки InAs для примесного атома Mn в валентной зоне (рис. 13а–в). При малых напряжениях это примесное состояние имеет заполнение, близкое к однократному. Когда приложенное напряжение приближается к значению, соответствующему энергии акцепторного уровня  $\varepsilon_d$ , в диапазоне напряжений  $|eV - \varepsilon_d| < \Gamma$  туннельная проводимость возрастает в области локализации d-орбитали. Параметр  $\Gamma$  соответствует уширению уровня энергии  $\varepsilon_d$ , вызываемому взаимодействием с состояниями непрерывного спектра валентной зоны полупроводника и состояниями СТМ-зонда. В результате в области локализации примесного атома Mn туннельный ток и локальная туннельная проводимость возрастают за счёт увеличения локальной плотности состояний в полупроводнике. Эта плотность состояний непосредственно отражает несферическую симметрию локализованного примесного d-уровня. Дополнительное усиление туннельного тока и проводимости возникает из-за включения неравновесного кулоновского взаимодействия между примесью и носителями заряда в области туннельного контакта.

Кулоновское взаимодействие между положительно заряженным примесным атомом и электронами проводимости при напряжениях, близких к энергии локализо-



**Рис. 13.** Нормированная туннельная проводимость и токовые СТМ-изображения в области локализации d-орбитали примесного состояния Mn. Вставки а–ж соответствуют пространственному распределению нормированной туннельной проводимости при различных значениях напряжения смещения. Напряжения смещения, соответствующие вставкам, отмечены стрелками. На вставках а1–ж1 приведены изображения пространственного распределения тока  $I(V)$  для напряжений, соответствующих вставкам (а, б, е, ж), расположенным непосредственно под токовыми СТМ-изображениями.

ванного состояния, приводит к следующей зависимости туннельного тока от напряжения:

$$I(V) \sim \frac{\gamma_k \gamma_p}{\gamma} \left[ \frac{D^2}{(eV - \epsilon_d)^2 + \gamma^2} \right]^{|W|v}, \quad (2)$$

где  $W$  — среднее значение кулоновского потенциала примесного атома,  $\gamma = \gamma_k + \gamma_p$  ( $\gamma_{k(p)}$  — скорости туннельных переходов в берега туннельного контакта),  $v$  — плотность электронных состояний в острие зонда СТМ, а  $D$  — ширина зоны проводимости в металле. Таким образом, можно ожидать дополнительного возрастания величины туннельного тока в окрестности примесного атома, когда величина напряжения на контакте приближается к пороговому значению  $eV = \epsilon_d$ .

На рисунке 13 показано возрастание туннельной проводимости в окрестности примесного атома, когда величина приложенного к туннельному контакту напряжения приближается к значению энергии локализованного состояния. Очевидно, что пик в туннельной проводимости в окрестности примесного уровня не связан с эффектом Кондо, поскольку напряжение на туннельном

переходе ( $\sim 400$  мэВ) превышает ширину уровня энергии примесного атома ( $\sim 100$  мэВ). Вблизи нулевых значений разности потенциалов на туннельном контакте крестообразная структура локальной плотности примесного состояния становится неразрешимой на СТМ-изображениях участка поверхности, показанного на рис. 13. Для значений напряжения в диапазонах  $-1 < V < -0,3$  В и  $0,5 < V < 1,5$  В светлая особенность отчётливо присутствует в пространственном распределении локальной туннельной проводимости в окрестности примесного атома. С другой стороны, необходимо отметить, что в пространственном распределении туннельного тока в диапазоне напряжений  $0,5 < V < 1,5$  В примесный атом имеет вид тёмной особенности, что соответствует локальному уменьшению величины тока.

Таким образом, кулоновское взаимодействие между примесным состоянием и электронами проводимости в состояниях непрерывного спектра острия СТМ значительно изменяет туннельные амплитуды и приводит к формированию особенностей в туннельном токе и в туннельной проводимости, когда величина приложенного напряжения приближается к величине энергии при-

месного атома. При этом вольт-амперные характеристики могут иметь асимметричную форму из-за различных зависимостей фазового фактора от приложенного напряжения: ниже или выше порогового. Следовательно, изменяя расстояние между остриём зонда СТМ и поверхностью, можно варьировать форму особенностей на вольт-амперных характеристиках.

### 3. Особенности туннельных процессов в структурах пониженной размерности

Результаты современных экспериментальных исследований туннельных процессов в структурах с ограниченной геометрией, в мезоскопических системах и тем более в структурах нанометровых размеров не могут быть проанализированы в рамках стандартной теории туннельных явлений, основанной на использовании равновесной функции распределения туннелирующих частиц. При этом неявно предполагается, что скорость релаксации неравновесных частиц намного выше скорости туннелирования. Для адекватного описания процессов туннелирования в таких структурах и контактах СТМ требуется учёт процессов релаксации, перестройки электронного спектра в берегах контакта и влияния локализованных состояний.

Общий подход к описанию туннельных процессов может быть основан на использовании диаграммной техники Келдыша, позволяющей самосогласованным образом учесть как перенормировку исходного спектра системы за счёт туннельных процессов, так и релаксацию неравновесных частиц [22].

Этот подход связан с использованием модельного гамильтониана системы следующего вида:

$$\hat{H} = \hat{H}_0 + \hat{H}_{\text{imp}} + \hat{H}_{\text{tun}} + \hat{H}_{\text{int}}, \quad (3)$$

где гамильтониан  $\hat{H}_0$  описывает изолированные берега туннельного контакта:

$$\hat{H}_0 = \sum_{\mathbf{k} \in \text{L}, \sigma} \varepsilon_{\mathbf{k}} c_{\mathbf{k}\sigma}^+ c_{\mathbf{k}\sigma} + \sum_{\mathbf{p} \in \text{R}, \sigma} (\varepsilon_{\mathbf{p}} - eV) c_{\mathbf{p}\sigma}^+ c_{\mathbf{p}\sigma}, \quad (4)$$

$c_{\mathbf{k}(p)\sigma}^+ / c_{\mathbf{k}(p)\sigma}$  — операторы рождения/уничтожения для невзаимодействующих электронов с импульсами  $\mathbf{k}(\mathbf{p})$  и энергиями  $\varepsilon_{\mathbf{k}(p)}$  в левом L (образец) и правом R (остриё зонда СТМ) берегах туннельного контакта соответственно.

Гамильтониан  $\hat{H}_{\text{imp}}$  соответствует примесным состояниям, локализованным в области туннельного контакта, и учитывает кулоновское взаимодействие локализованных электронов:

$$\hat{H}_{\text{imp}} = \sum_{d, \sigma} \varepsilon_d d_{\sigma}^+ d_{\sigma} + U \sum_{d, \sigma} n_d^{\sigma} n_d^{-\sigma}. \quad (5)$$

В этом выражении  $n_d^{\sigma} = d_{\sigma}^+ d_{\sigma}$ , оператор  $d_{\sigma}^{\pm} (d_{\sigma})$  соответствует рождению/уничтожению электрона со спином  $\sigma$  на примесном уровне,  $U$  — величина кулоновского взаимодействия электронов на уровне энергии локализованного состояния. Туннельные процессы через промежуточное состояние между берегами контакта в общем случае описываются гамильтонианом:

$$\hat{H}_{\text{tun}} = \sum_{\mathbf{k}, d, \sigma} T_{\mathbf{k}d} (d_{\sigma}^+ c_{\mathbf{k}\sigma} + c_{\mathbf{k}\sigma}^+ d_{\sigma}) + \sum_{\mathbf{p}, d, \sigma} T_{\mathbf{p}d} (d_{\sigma}^+ c_{\mathbf{p}\sigma} + c_{\mathbf{p}\sigma}^+ d_{\sigma}). \quad (6)$$

Обычно предполагается, что туннельные амплитуды  $T_{k(p)d}$  не зависят от импульса и описывают переходы электронов из одного берега туннельного контакта в другой через промежуточные локализованные состояния. Гамильтониан  $\hat{H}_{\text{int}}$  описывает наличие в рассматриваемой системе различного рода взаимодействий, таких как взаимодействие с термостатом, электрон-фононное взаимодействие, взаимодействие локализованных электронов с электронами проводимости в берегах туннельного контакта и т.д.

Для начала рассмотрим модельную систему, описываемую гамильтонианом (3), в котором отсутствует слабое  $\hat{H}_{\text{int}}$ , а берега туннельного контакта находятся в термодинамическом равновесии. Выражение для туннельного тока можно записать с использованием диаграммной техники Келдыша для неравновесных процессов:

$$\begin{aligned} I(V) &= i \sum_{\mathbf{k}, \sigma, d} T_{\mathbf{k}d} (c_{\mathbf{k}\sigma}^+ d_{\sigma} - d_{\sigma}^+ c_{\mathbf{k}\sigma}) = 2 \text{Re} \sum_{\mathbf{k}, \sigma, d} T_{\mathbf{k}d} G_{\mathbf{k}d}^{\sigma <} = \\ &= -2 \text{Re} \sum_{\mathbf{p}, \sigma, d} T_{\mathbf{p}d} G_{\mathbf{p}d}^{\sigma <}. \end{aligned} \quad (7)$$

Последнее равенство является следствием уравнения непрерывности. Функции Грина в диаграммной технике Келдыша определяются стандартным образом из уравнения Дайсона [23]. В действительности для нахождения туннельного тока достаточно найти только функции  $G_{dd}^{\leq}$  и  $G_{dd}^A$ , которые связаны соотношением

$$G_{dd}^{\leq}(\omega) = (-i) 2n_d(\omega) \text{Im} G_{dd}^A(\omega). \quad (8)$$

При этом туннельный ток определяется выражением

$$I(V) = 2i\gamma_k \int (iG_{dd}^{\leq} - 2n_k \text{Im} G_{dd}^A) \frac{d\omega}{2\pi}, \quad (9)$$

где  $\gamma_{k(p)d} = v_{L(R)}^0 \pi T_{k(p)d}^2$ ,  $G^A$  — опережающая функция Грина. В рамках данного подхода наличие в системе кулоновского взаимодействия локализованных электронов приводит к самосогласованному изменению запаздывающей функции Грина  $G_{dd}^R$ .

Для корректного описания туннельных явлений во многих случаях надо самосогласованным образом учитывать как перенормировку электронного спектра, так и конечную скорость релаксации неравновесных электронов в берегах контакта. Для модельного учёта релаксации можно ввести в гамильтониан дополнительное взаимодействие с термостатом:

$$\hat{H}_{\text{int}} = \sum_{\mathbf{k}\mathbf{p}'\sigma} g(\mathbf{k} - \mathbf{p}') c_{\mathbf{k}\sigma}^+ b_{\mathbf{p}'\sigma} + \text{h.c.}, \quad (10)$$

где

$$g(\mathbf{k} - \mathbf{p}') = \frac{1}{L^d} \int g(r) \exp [i(\mathbf{k} - \mathbf{p}')\mathbf{r}] \mathbf{r}, \quad (11)$$

$g(r)$  — эффективный матричный элемент взаимодействия с термостатом,  $g(r) \rightarrow 0$  в области туннельного барьера и на его границах,  $b_{\mathbf{p}'\sigma}^+$  — оператор рождения электрона термостата в состоянии  $(\mathbf{p}'\sigma)$ ,  $L$  — характерный размер системы,  $d$  — размерность системы. Термостат находится в состоянии равновесия, которое не меняется под действием возмущения  $g(r)$ .

Теперь в уравнениях для функций Грина в технике Келдыша появляется дополнительная собственно энер-

гетическая часть  $\hat{\Sigma}$ , обусловленная взаимодействием с термостатом:

$$\Sigma_{\mathbf{k}\mathbf{k}'}^{\alpha\beta} = \sum_{\mathbf{p}'} g(\mathbf{k} - \mathbf{p}') g^+(\mathbf{k}' - \mathbf{p}') G_{\mathbf{p}'\mathbf{p}'}^{\alpha\beta}(\omega). \quad (12)$$

Разумно считать, что термостат представляет собой систему случайно расположенных в точках  $\mathbf{R}_i$  рассеивателей (или  $g(r)$  — случайная функция). При этом

$$g(\mathbf{k} - \mathbf{p}) g(\mathbf{k}' - \mathbf{p}) = N^{-1} \sum_i g_i^2 \exp[i(\mathbf{k} - \mathbf{k}')\mathbf{R}_i]. \quad (13)$$

Для состояний непрерывного спектра при хаотическом распределении таких центров

$$\Sigma_{\mathbf{k}\mathbf{k}'}^{\alpha\beta} = g_i^2 \bar{n} \delta_{\mathbf{k}\mathbf{k}'} \sum_{\mathbf{p}} G_{\mathbf{p}}^{\alpha\beta}(\omega), \quad (14)$$

где  $\bar{n} = N_i/N$  — средняя концентрация рассеивателей. При этом

$$\frac{\partial n_k}{\partial t} + iT_k \int d\omega (G_{dk}^<(\omega) - G_{kd}^<(\omega)) = -2\Gamma_k(n_k - n_p^0), \quad (15)$$

$$\Gamma_k = g_i^2 \pi \bar{n} \nu_{p'}(\omega). \quad (16)$$

Однако если интересоваться вкладом в ток от связанных локализованных вблизи микроконтактов состояний, то можно для модели термостата выбрать другой предельный случай: локальное подключение термостата. В этом случае  $\hat{\Sigma}_{kk'}$  не зависит от  $k$  и  $k'$ , т.е.

$$\Sigma_{\mathbf{k}\mathbf{k}'}^{\alpha\beta} = g^2 \sum_{\mathbf{p}} G_{\mathbf{p}\mathbf{p}}^{\alpha\beta}(\omega), \quad \Sigma_{\mathbf{k}\mathbf{k}'}^{\mathbf{R}} = i\pi g^2 \nu_{\mathbf{p}}(\omega) = i\Gamma_{\mathbf{k}}(\omega). \quad (17)$$

Мы остановимся более подробно на последней модели.

В рамках данной модели можно найти явный вид запаздывающей функции Грина  $\hat{G}^{\mathbf{R}}$ . Например:

$$G_{dd}^{\mathbf{R}}(\omega) = \frac{1 - i\Gamma_k N(\omega)}{Z(\omega)}, \quad (18)$$

$$Z(\omega) = (1 - i\Gamma_k N(\omega))(\omega - \varepsilon_d + i\gamma_p(\omega)) - T_k^2 N(\omega), \quad (19)$$

где

$$N(\omega) = \sum_{\mathbf{k}} G_{\mathbf{k}\mathbf{k}}^{\mathbf{R}0}(\omega), \quad (20)$$

$$\Sigma_{dd}^{\mathbf{r}} = i\pi |T_p|^2 \nu_{p'}(\omega) = i\gamma_p(\omega), \quad (21)$$

где  $\nu_{p'}(\omega)$  — плотность электронных состояний  $p'$ .

Из явного вида функций Грина можно определить условия формирования связанных состояний, отщеплённых от границ исходного спектра за счёт туннельного взаимодействия [24].

#### 4. Связанные состояния, отщеплённые от границ исходного спектра в результате туннельного взаимодействия

Спектр исследуемой системы определяется полюсами функции Грина  $G_{dd}^{\mathbf{R}}$ . Если уширения электронных состояний  $\gamma_p$ ,  $\Gamma_k$  малы, то энергия коллективного связанного состояния, лежащего в запрещённой зоне, определяется уравнением

$$\omega - \varepsilon_d - T_k^2 N(\omega) = 0 \quad (22)$$

или

$$\omega - \varepsilon_d - T_k^2 \int d\varepsilon \frac{\nu_0(\varepsilon)}{\omega - \varepsilon} = 0, \quad (23)$$

где  $\nu_0(\varepsilon)$  — затравочная плотность состояний непрерывного спектра. Уравнение (23) имеет чисто действительное решение в области  $\omega > \varepsilon_v$  и  $\omega < \varepsilon_0$ , где  $\varepsilon_0$  и  $\varepsilon_v$  — соответственно нижняя и верхняя границы спектра  $\varepsilon_k$ . Этому значению  $\omega$  соответствует локализованное состояние, лежащее в запрещённой зоне. Энергия этого состояния зависит от величины  $T_k$  и конкретного вида затравочной плотности состояний  $\nu_0(\varepsilon)$ .

Если использовать приближённое выражение для плотности состояний двумерной электронной системы  $\nu_0(\varepsilon) = (\varepsilon_v - \varepsilon_0)^{-1}$ , то в области  $\omega > \varepsilon_v$

$$\omega = \tilde{\varepsilon}_d = \omega_0 + \varepsilon_v, \quad (24)$$

$$\omega_0 = (\varepsilon_v - \varepsilon_0) \exp\left[-\frac{(\varepsilon_v - \varepsilon_0)(\tilde{\varepsilon}_d - \varepsilon_d)}{T_k^2}\right]. \quad (25)$$

При квазиодномерном характере  $\nu_0(\varepsilon)$  вблизи границы спектра

$$\tilde{\varepsilon}_d = \varepsilon_v + \left(\frac{T_k^2}{\varepsilon_v - \varepsilon_d}\right)^2 W_1^{-1}, \quad (26)$$

где  $W_1$  — эффективная ширина квазиодномерной зоны. Нетрудно видеть, что амплитуда волновой функции связанного состояния  $\psi_0$  непосредственно в области контакта определяется вычетом функции  $G^{\mathbf{R}}(\omega, 0) = \sum_{kk'} G_{kk'}^{\mathbf{R}}(\omega)$  в полюсе  $\omega = \tilde{\varepsilon}_d$ .

Это даёт

$$a^D |\psi_0|^2 = \left[ \int d\varepsilon \nu_0(\varepsilon) (\tilde{\varepsilon}_d - \varepsilon)^{-1} \right]^2 \left[ \int d\varepsilon \nu_0(\varepsilon) (\tilde{\varepsilon}_d - \varepsilon)^{-2} \right]^{-1}, \quad (27)$$

где  $a$  — постоянная решётки,  $D$  — пространственная размерность системы. В случае двумерной зоны

$$|\psi_0|^2 = (\varepsilon_v - \varepsilon_d)^2 W_2 \omega_0 T_k^{-4} a^{-D}, \quad (28)$$

где  $W_2 = \varepsilon_v - \varepsilon_0$  — ширина  $2D$  зоны.

Оказывается, что, хотя экспоненциальное уменьшение хвостов волновой функции происходит на характерной длине  $R_0 \approx a(W/\omega_0)^{1/D}$ , размер эффективной области локализации определяется не  $R_0$ , а значением  $(|\psi_0|^{-2})^{1/D}$ .

#### 5. Появление индуцированной туннельной проводимости

Наличие отщеплённых связанных состояний может приводить к появлению индуцированной туннельной проводимости. В стационарном случае, интересуясь вкладом в туннельный ток от связанного состояния, можно считать, что

$$G_{\mathbf{k}\mathbf{k}'}^<(\omega) = -2in_k(\omega) \text{Im} G_{\mathbf{k}\mathbf{k}'}^{\mathbf{R}}(\omega), \quad (29)$$

$$G_{dd}^<(\omega) = -2in_d(\omega) \text{Im} G_{dd}^{\mathbf{R}}(\omega), \quad (30)$$

где  $n_k(\omega)$  и  $n_d(\omega)$  — числа заполнения для состояний  $k$  и  $d$  соответственно.

Туннельный ток определяется выражением

$$I = -2 \int \frac{d\omega}{2\pi} \gamma_p(\omega) \operatorname{Im} G_{dd}^R(\omega) (n_d(\omega) - n_p^0(\omega)) = \\ = 2 \int \frac{d\omega}{2\pi} \Gamma_k(\omega) \operatorname{Im} G^R(\omega, 0) (n_k(\omega) - n_p^0(\omega)), \quad (31)$$

где

$$\operatorname{Im} G^R(\omega, 0) = \operatorname{Im} \sum_{\mathbf{k}\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'}^R(\omega), \quad (32)$$

$n_p^0(\omega)$  — фермиевская функция распределения для состояний  $p'$  с энергией  $\omega$ ; распределения  $n_p^0(\omega)$  и  $n_{p'}^0(\omega)$  различаются положением уровня Ферми.

Для конкретного случая двумерной системы электронов, учитывая, что при  $\omega \sim \tilde{\varepsilon}_d$

$$\sum_{\mathbf{k}\mathbf{k}'} \operatorname{Im} G_{\mathbf{k}\mathbf{k}'}^R(\omega) = \operatorname{Im} G_{dd}^R(\omega) (\varepsilon_v - \varepsilon_d)^2 T_k^{-2}, \quad (33)$$

получим следующее выражение для туннельного тока в приближении  $T_k^2 v_0 \gg \gamma_p, \Gamma_k$ :

$$I \sim \frac{2\gamma_p \Gamma_k (n_p^0(\tilde{\varepsilon}_d) - n_p^0(\tilde{\varepsilon}_d))}{[\Gamma_d T_k^2 (\varepsilon_v - \varepsilon_d)^{-2} + \Gamma_k] Z'(\tilde{\varepsilon}_d)}, \quad (34)$$

где  $Z'(\tilde{\varepsilon}_d)^{-1} = (T_k v_0)^{-2} \exp[-(\varepsilon_v - \varepsilon_d) \varepsilon_v T_k^{-2}]$ .

Значения величин  $\gamma_p$  и  $\Gamma_k$ , входящих в формулу (34), определяются при  $\omega \sim \tilde{\varepsilon}_d$ . Если  $\tilde{\varepsilon}_d$  лежит между  $E_F$  и  $E_F - eV$ , то туннельный ток отличен от нуля благодаря появлению нового коллективного состояния  $\tilde{\varepsilon}_d$ .

Напомним ещё раз, что положение  $\tilde{\varepsilon}_d$  относительно границы зоны зависит от приложенного напряжения  $V$ , так как разность всех характерных энергий изменяется дополнительно на величину  $eV$ . Следует отметить, что даже в отсутствие локализованных состояний в области контакта наличие конечной скорости релаксации неравновесных носителей приводит к модификации выражения (1) для туннельного тока. При этом формула для туннельного тока, учитывающая процессы релаксации носителей, приобретает вид [25]:

$$I \sim 2\pi e \int d\omega \frac{T^2 v_k(\omega) v_p(\omega) [n_p(\omega) - n_k(\omega - eV)] \Gamma_k \Gamma_p}{T^2 v_p \Gamma_p + T^2 v_k \Gamma_k + \Gamma_p \Gamma_k}. \quad (35)$$

Здесь  $\Gamma_{k(p)}$  — скорости релаксации носителей в берегах контакта,  $T$  — амплитуда туннельных переходов между берегами контакта. При  $\Gamma_{k(p)} \gg T^2$  это выражение переходит в стандартную формулу (1). Релаксация в системе уменьшает величину туннельного тока, а изменение локальной плотности состояний и возникновение неравновесного локального распределения электронов в области туннельного контакта могут значительно исказить его спектральные характеристики. Следовательно, ток при наличии релаксации определяется самым слабым звеном в цепи.

Можно модифицировать полученные результаты для модели термостата с хаотически распределёнными рассеивателями. Отметим, что для такой модели термостата

$$\Sigma_{\mathbf{k}\mathbf{k}'}^{z\beta} = g_i^2 \bar{n} \delta_{\mathbf{k}\mathbf{k}'} \sum_p G_p^{z\beta}(\omega), \quad (36)$$

где  $\bar{n}$  — среднее число рассеивателей, приходящихся на одну элементарную ячейку.

В отличие от формул, в которых туннельный ток определяется скоростью релаксации локальной электронной плотности:

$$I \sim \int d\omega (n_k - n_p^0) \operatorname{Im} G^R(\omega, 0) v_p(\omega) g^2, \quad (37)$$

теперь туннельный ток задаёт скорость изменения полного числа частиц:

$$I \sim \int d\omega v_p(\omega) g^2 \bar{n} (n_k - n_p^0) \operatorname{Im} \sum_{\mathbf{k}} G_{\mathbf{k}\mathbf{k}}^R(\omega). \quad (38)$$

Оказывается, что выражения (37) и (38) вблизи энергии отщеплённого уровня различаются на величину  $\bar{n} |\psi_0|^{-2}$ , где  $\psi_0$  — амплитуда волновой функции связанного состояния в точке  $r=0$ . Следовательно, если концентрация релаксационных центров  $\bar{n}$  такова, что внутри области локализации связанного состояния  $S \sim |\psi_0|^{-2}$  находится в среднем один такой центр, то оба выражения, определяющие туннельный ток, просто совпадают.

Таким образом, при анализе современных экспериментов по СТМ/СТС необходимо существенно изменить привычное описание туннельных контактов [26]. Становится неверным представление о том, что величина туннельного тока отражает исходную плотность состояний в исследуемом материале. Изменение локальной плотности состояний и появление неравновесного распределения электронов возможны и в отсутствие локализованных состояний. В этом случае в гамильтониане (3) отсутствуют слагаемые  $\hat{H}_{\text{imp}}$  и  $\hat{H}_{\text{int}}$ , а туннельные процессы между берегами контакта описываются гамильтонианом  $\hat{H}_{\text{tun}}$ :

$$\hat{H}_{\text{tun}} = \sum_{\mathbf{k}, p, \sigma} T (c_{\mathbf{k}\sigma}^+ c_{p\sigma} + c_{p\sigma}^+ c_{\mathbf{k}\sigma}). \quad (39)$$

Изменение локальной плотности в берегах контакта  $v_{L(R)}$  определяется с помощью диаграммной техники Келдыша следующим образом [27]:

$$v_{L(R)} = \frac{1}{\pi} \operatorname{Im} \sum_{\mathbf{k}, \mathbf{k}' \in L(R)} G_{\mathbf{k}, \mathbf{k}'}^A = \frac{v_{L(R)}^0}{1 + T^2 v_L^0 v_R^0}. \quad (40)$$

Локально неравновесное распределение электронов может быть найдено из выражения:

$$\sum_{\mathbf{k}, \mathbf{k}' \in L} G_{\mathbf{k}, \mathbf{k}'}^< = \frac{2iv_L^0}{(1 + T^2 v_L^0 v_R^0)^2} (n_k(\omega) + T^2 v_L^0 v_R^0 n_p(\omega)) = \\ = 2\pi i n_L v_L, \quad (41)$$

где  $v_L^0$  и  $v_R^0$  — невозмущённые плотности состояний в левом и правом берегах туннельного контакта,  $n_{k(p)}(\omega)$  — равновесные фермиевские функции распределения электронов в берегах туннельного контакта,  $v_L$  и  $v_R$  — локальные плотности состояний в области контакта с учётом туннельных переходов.

В результате выражение для туннельного тока преобразуется к виду

$$I = \int d\omega \frac{4T^2 v_L^0 v_R^0}{(1 + T^2 v_L^0 v_R^0)^2} [n_k(\omega) - n_p(\omega)]. \quad (42)$$

Таким образом, вследствие взаимодействия образца с близко расположенным остриём в процессе туннелирования исходная плотность состояний искажается, а в ряде случаев появляются новые связанные состояния, играющие особо важную роль в случае диэлектриков, ВЗП на поверхности, поверхностных зон в полупроводниках и т.д. При этом может наблюдаться отличный от нуля ток при такой разности потенциалов, когда исходные зоны не перекрываются, появляется вклад в ток от примесей с уровнями энергии, лежащими глубоко под уровнем Ферми. В частности, в низкоразмерных системах может появляться индуцированный туннельный ток при напряжениях, соответствующих энергии запрещённой зоны, за счёт отщеплённых от границ спектра локализованных состояний (см. формулу (34)).

Тот факт, что положение связанного состояния сложным образом зависит от величины туннельного матричного элемента  $T$ , должен приводить к неэкспоненциальной (а иногда и немонотонной) зависимости туннельного тока от расстояния до образца в области потенциалов, соответствующих энергии таких состояний. Данная модель туннельных процессов с участием отщеплённых локализованных состояний позволяет объяснить многие экспериментальные результаты, например, резкое увеличение туннельной проводимости при наблюдении ВЗП на поверхности ЖК с уменьшением расстояния между поверхностью образца и остриём СТМ, описанное в разделе 2.1.

## 6. Модель Андерсона для описания неравновесных примесных состояний с коррелированными электронами

В данном разделе проанализированы экспериментальные спектры туннельной проводимости, обусловленные неравновесным взаимодействием примесных состояний. В СТС-экспериментах, описанных в разделе 2.4, исследовалось поведение туннельной проводимости в широком диапазоне напряжений от +2,5 до -2 В (типичное значение температуры Кондо не превышает 1 мэВ), поэтому влияние эффекта Кондо на формирование особенностей туннельных характеристик пренебрежимо мало. Исследуемая система находится в слабом магнитном поле  $\mathbf{V}$ , так что величина расщепления уровней не превосходит их уширения.

Кулоновское взаимодействие локализованных на примесных атомах электронов учитывается самосогласованным образом в приближении среднего поля. Оказывается, что при изменении туннельного напряжения на контакте различие по энергии для электронов с противоположными спинами на каждом примесном атоме может сильно увеличиваться и происходит резкое увеличение спиновой асимметрии: переход от "парамагнитного" режима к "магнитному". При дальнейшем увеличении приложенного напряжения возможен и обратный переход от "магнитного" состояния к "парамагнитному". Кроме того, выявлены условия, при которых взаимодействие между примесными атомами приводит к усилению "магнитного" режима в неравновесном состоянии. Взаимодействие между примесными атомами также приводит к перераспределению неравновесных зарядов, локализованных вблизи примесей, поэтому становится возможной стабилизация примесных уровней в окрестности уровня Ферми любого из берегов контакта и взаимное

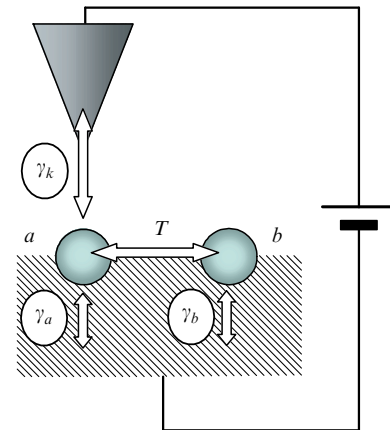


Рис. 14. Схематическое изображение туннельной структуры при наличии взаимодействующих примесей.

сближение энергетических уровней взаимодействующих примесей в определённом интервале значений приложенного внешнего напряжения.

Если вблизи поверхности полупроводника находятся два примесных атома  $a$  и  $b$  и остриё СТМ расположено над атомом  $a$  (рис. 14), то туннельные переходы между локализованными состояниями атома  $a$  и остриём СТМ определяются скоростью туннелирования  $\gamma_k$ . Скорости переходов между локализованными состояниями на примесных атомах  $a$  и  $b$  и состояниями непрерывного спектра полупроводника характеризуются соответствующими кинетическими параметрами  $\gamma_a$  и  $\gamma_b$ .

Кинетические параметры  $\gamma_a$ ,  $\gamma_b$ ,  $\gamma_k$  определяются из гамильтониана исследуемой структуры. Взаимодействие между примесными атомами приводит к непосредственному "перескоку" электрона с одного атома на другой с соответствующей амплитудой перехода  $T$ . Энергии примесных атомов  $a$  и  $b$  при нулевом напряжении на контакте отличаются, поскольку атом  $a$  проявляется в СТС-спектрах при  $V = 0$ , а атом  $b$  — нет (см. раздел 2.4, рис. 11). Отличие в положении уровней локализованных состояний может быть вызвано как неоднородным потенциалом кристаллической решётки полупроводника, так и расположением атомов в разных приповерхностных слоях полупроводниковой матрицы. Различие в энергии локализованных состояний примесных атомов может приводить к различным значениям энергии внутреннего кулоновского взаимодействия.

Амплитуда электронного перехода между примесными атомами  $T$  не превосходит скоростей релаксации  $\gamma_a$  и  $\gamma_b$ , возникающих вследствие гибридизации локализованных состояний с состояниями непрерывного спектра полупроводниковой матрицы, поскольку, согласно экспериментальным результатам, расстояние между исследуемыми примесями составляет приблизительно 3 нм. Это предположение означает, что из-за уширения энергетических уровней их равновесное расщепление за счёт взаимодействия не проявляется в плотности состояний исследуемой системы. Но кинетические процессы, благодаря такому взаимодействию, приводящему к перераспределению заряда между примесными атомами, сильно модифицируются. Наличие слабого магнитного поля не приводит к разрешимому в эксперименте расщеплению уровней энергии электронов с противоположными спинами, поскольку величина расщепления не пре-

восходит ширину уровней ( $2\mu B < \Gamma$ ) при малых напряжениях на туннельном контакте ( $\mu$  — магнитная проницаемость).

Рассмотрим особенности туннельных характеристик над примесными атомами в режиме переменной валентности, для которого параметры системы удовлетворяют условиям:  $\varepsilon/\Gamma \sim 1$ ,  $U \gg \varepsilon, \Gamma$ ,  $2\mu B < \Gamma$ . В этом режиме кулоновское взаимодействие локализованных электронов можно самосогласованно учитывать в приближении среднего поля:

$$\tilde{\varepsilon}_a^\sigma = \varepsilon_a^\sigma + \alpha V + U_a \langle n_a^{-\sigma} \rangle, \quad \tilde{\varepsilon}_b^\sigma = \varepsilon_b^\sigma + \beta V + U_b \langle n_b^{-\sigma} \rangle, \quad (43)$$

$$\varepsilon_{a(b)}^{\pm\sigma} = \varepsilon_{a(b)} \mp \mu B. \quad (44)$$

Начальное положение уровня зависит от приложенного напряжения как вследствие существования в области контакта внешнего электрического поля (член  $\alpha V$ , где  $\alpha < 1$ , приближённо учитывает влияние внешнего поля), так и за счёт кулоновского взаимодействия локализованных электронов, величина которого определяется неравновесным распределением электронной плотности при фиксированном напряжении на контакте.

Числа заполнения примесных состояний являются неравновесными и определяются самосогласованным образом из системы кинетических уравнений. Взаимодействие между примесными атомами приводит к зависимости неравновесных чисел заполнения одного примесного атома ( $n_a^{\pm\sigma}$ ) от соответствующих чисел заполнения другого атома ( $n_b^{\pm\sigma}$ ).

Гамильтониан исследуемой системы имеет вид (3), в котором часть  $\hat{H}_{\text{imp}}$  соответствует примесным состояниям, и учитывает кулоновское взаимодействие локализованных электронов:

$$\hat{H}_{\text{imp}} = \sum_{\sigma} \varepsilon_a^{\sigma} a_{\sigma}^{\dagger} a_{\sigma} + \frac{U_a}{2} \sum_{\sigma} n_{\sigma}^a n_{-\sigma}^a + \sum_{\sigma} \varepsilon_b^{\sigma} b_{\sigma}^{\dagger} b_{\sigma} + \frac{U_b}{2} \sum_{\sigma} n_{\sigma}^b n_{-\sigma}^b. \quad (45)$$

В этом выражении  $n_{\sigma}^a = a_{\sigma}^{\dagger} a_{\sigma}$ , оператор  $a_{\sigma}$  уничтожает электрон со спином  $\sigma$  на примесном атоме  $a$ ,  $n_{\sigma}^b = b_{\sigma}^{\dagger} b_{\sigma}$ ,  $b_{\sigma}$  уничтожает электрон на атоме  $b$  со спином  $\sigma$ ,  $\varepsilon_a$ ,  $\varepsilon_b$  — уровни энергии примесных атомов  $a$  и  $b$ , которые в общем случае зависят от приложенного напряжения.

$\hat{H}_{\text{tun}}$  отвечает за туннельные переходы между примесными состояниями и каждым из берегов контакта и имеет вид

$$\hat{H}_{\text{tun}} = \sum_{\mathbf{p}, \sigma} T_{\mathbf{p}, a} (c_{\mathbf{p}, \sigma}^{\dagger} a_{\sigma} + \text{h.c.}) + \sum_{\mathbf{p}, \sigma} T_{\mathbf{p}, b} (c_{\mathbf{p}, \sigma}^{\dagger} b_{\sigma} + \text{h.c.}) + \sum_{\mathbf{k}, \sigma} T_{\mathbf{k}, a} (c_{\mathbf{k}, \sigma}^{\dagger} a_{\sigma} + \text{h.c.}). \quad (46)$$

Взаимодействие между примесями включено в

$$\hat{H}_{\text{int}} = \sum_{\sigma} T (a_{\sigma}^{\dagger} b_{\sigma} + b_{\sigma}^{\dagger} a_{\sigma}). \quad (47)$$

Функции Грина  $G_{\mathbf{a}\mathbf{k}}^{\sigma <}$ , определяющие величину туннельного тока, могут быть найдены из системы кинетических уравнений [23].

Как отмечалось выше, при использовании диаграммной техники Келдыша собственно энергетическая часть  $\hat{\Sigma}$  может включать различные типы взаимодействия, но в интересующем нас случае разумно считать, что  $\hat{\Sigma}$  опре-

деляется только туннельными переходами между примесными состояниями и берегами контакта и непосредственным взаимодействием между примесями. Использование этого приближения вполне оправдано, поскольку самое сильное взаимодействие в рассматриваемом случае — кулоновское отталкивание  $U$  — включено в невозмущённую примесную функцию Грина  $G_0$ .

Кинетические коэффициенты, которые определяют туннельные процессы, могут быть представлены в виде

$$\gamma_k(\omega) = \pi |T_{ka}|^2 v_{\mathbf{k}}(\omega), \quad \gamma_b(\omega) = \pi |T_{pb}|^2 v_{\mathbf{p}}(\omega), \quad (48)$$

$$\gamma_a(\omega) = \pi |T_{pa}|^2 v_{\mathbf{p}}(\omega), \quad \Gamma = \gamma_a + \gamma_b + \gamma_k.$$

Ключевым моментом данного подхода является условие самосогласования, которому должны удовлетворять неравновесные электронные числа заполнения на примесных атомах  $n_a^{\sigma}$  и  $n_b^{\sigma}$ :

$$n_a^{\sigma} = \frac{1}{\pi} \int d\omega n_a^{\sigma}(\omega) \text{Im} G_{aa}^{\Lambda}(\omega), \quad n_b^{\sigma} = \frac{1}{\pi} \int d\omega n_b^{\sigma}(\omega) \text{Im} G_{bb}^{\Lambda}(\omega). \quad (49)$$

Неравновесные числа заполнения  $n_a^{\sigma}(\omega)$  и  $n_b^{\sigma}(\omega)$  определяются из системы кинетических уравнений, подробный вывод которых можно найти в работе [17]. Полная электронная плотность для каждого из атомов примесного комплекса может быть определена из системы уравнений:

$$n_a^{\sigma} = n_p^{\sigma}(a) + \frac{\gamma_k(\gamma_b + \eta_{\sigma})(n_k^{\sigma}(a) - n_p^{\sigma}(a)) + \gamma_b \Gamma_{ab}}{(\gamma_k + \gamma_a)(\gamma_b + \eta_{\sigma}) + \gamma_b \eta_{\sigma}}, \quad (50)$$

$$n_b^{\sigma} = n_p^{\sigma}(b) + \frac{\gamma_k \eta_{\sigma}(n_k^{\sigma}(a) - n_p^{\sigma}(a)) - (\gamma_a + \gamma_k) \Gamma_{ab}}{(\gamma_k + \gamma_a)(\gamma_b + \eta_{\sigma}) + \gamma_b \eta_{\sigma}},$$

где

$$\eta^{\sigma} = T^2 \text{Im} R_{ab}^{\sigma},$$

$$\Gamma_{ab} = T \gamma_k \int \text{Im} \left( R_{ab} \sum_{\sigma} G_{ab}^{\text{R}\sigma}(\omega) \right) (n_k(\omega) - n_p(\omega)) d\omega,$$

$$n_{p(k)}^{\sigma}(a) = \frac{1}{\pi} \int d\omega n_{p(k)}^0(\omega) \text{Im} G_{aa}^{\Lambda-\sigma}(\omega) \quad (51)$$

и

$$R_{ab}^{-1} = G_{0a}^{\text{R}-1} - G_{0b}^{\text{A}-1} - i\Gamma. \quad (52)$$

Это система нелинейных уравнений для определения  $n_{a(b)}^{\sigma}$ , поскольку энергии  $\varepsilon_{a(b)}^{\sigma}$  и функции  $G_{ab}^{\text{R}\sigma}$ ,  $G_{aa}^{\text{R}\sigma}$ ,  $R_{ab}$  зависят от электронных чисел заполнения с противоположным спином.

В равновесном случае  $V = 0$  и, следовательно:  $n_a^{\sigma} = n_p^{\sigma}(a) = n_k^{\sigma}(a)$ .

Подставляя в выражение для туннельного тока найденные из кинетических уравнений неравновесные числа заполнения на примеси, получаем окончательное выражение для туннельного тока:

$$I(V) = e \frac{\gamma_k(\gamma_b + \eta_{\sigma})}{(\gamma_a + \gamma_k)(\gamma_b + \eta_{\sigma}) + \gamma_b \eta_{\sigma}} \times \int \left[ \left( \gamma_a + \frac{\gamma_b \eta_{\sigma}}{\gamma_b + \eta_{\sigma}} \right) \text{Im} G_{a,a}^{\text{R}\sigma}(\omega, V) + T \frac{\gamma_b \gamma_k}{\gamma_b + \eta_{\sigma}} \text{Im} (R_{ab}^{\sigma} G_{ab}^{\text{A}\sigma}(\omega, V)) \right] \times (n_p^0(\omega) - n_k^0(\omega - eV)) d\omega, \quad (53)$$



в котором для каждого фиксированного значения напряжения на контакте самосогласованно определённые неравновесные числа заполнения  $n_a^\sigma, n_b^\sigma$  входят в функции  $G_{aa}^R, G_{ab}^R, R_{ab}$ .

Как и ожидалось, туннельный ток зависит только от разности функций распределения электронов в берегах контакта. Первое слагаемое в выражении для туннельного тока описывает перенормировку скорости релаксации неравновесной электронной плотности на примесном атоме  $a$  за счёт взаимодействия с соседним примесным атомом  $b$ :

$$\gamma_a \longrightarrow \gamma_a + \frac{\gamma_b \eta_\sigma}{\gamma_b + \eta_\sigma}. \quad (54)$$

Если взаимодействие между примесными атомами отсутствует,  $T = 0, \eta = 0$ , то полученное выражение для тока (53) просто описывает ток через примесное локализованное состояние [23].

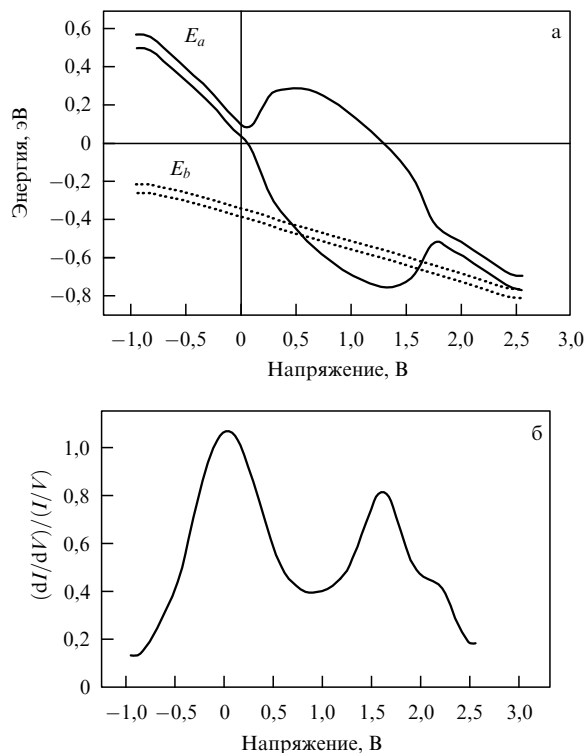
Второе слагаемое из выражения (53) отвечает за перераспределение заряда между взаимодействующими примесными атомами. Вследствие этого перераспределения заряда туннельная проводимость, полученная из (53), уже не просто пропорциональна плотности состояний на примеси.

Рассмотренная система с двумя взаимодействующими андерсоновскими примесями является простейшим примером многоканальной туннельной структуры с взаимодействием между различными каналами туннелирования, которое может приводить к интерференционным эффектам, что сильно усложняет исследование туннельных процессов в таких системах.

В результате самосогласованного анализа предложенной модели можно выявить различные режимы поведения туннельной проводимости в окрестности взаимодействующих примесных атомов в широком диапазоне изменения напряжения на контакте. При численных расчётах коэффициенты  $\alpha$  и  $\beta$ , описывающие изменение энергии примесных уровней за счёт внешнего поля, принимают значения  $\alpha \simeq 0,3$  и  $\beta \simeq 0,1$ .

Если  $\gamma_k \ll \gamma_a, \gamma_b$ , то в определённой области значений напряжения на контакте разница энергий для электронов с противоположными спинами существенно превосходит  $\Gamma$ , один из примесных атомов (атом  $a$ ) может находиться в "магнитном" состоянии. При изменении приложенного напряжения могут происходить переходы от "парамагнитного" режима к "магнитному" и наоборот (рис. 15а). Такие переходы приводят к двукратному включению и выключению атома  $a$  в спектре пространственного распределения туннельной проводимости при изменении туннельного напряжения (рис. 15б), что наблюдалось в экспериментах (см. раздел 2.4).

Кроме того, уровни энергии примесных атомов оказываются стабилизированными в окрестности уровня Ферми одного из берегов контакта (острия СТМ или образца), в то время как напряжение на контакте меняется в широком интервале, сравнимом с величиной кулоновского взаимодействия  $U$ . В спектре туннельной проводимости, полученном над одним из примесных атомов, возникают два пика, ширина которых сопоставима с величиной кулоновского взаимодействия и сильно превышает туннельную ширину уровней локализованных состояний. Это согласуется с экспериментальными данными, приведёнными в разделе 2.4.



**Рис. 15.** Два anomalно широких отчётливых пика в спектре туннельной проводимости, хорошо согласующихся с экспериментальными кривыми СТС. (а) Зависимость энергии атомов  $a$  и  $b$  от приложенного напряжения. Типичные значения параметров (эВ):  $\varepsilon_a^0 = -0,25, \varepsilon_b^0 = -0,5, U_a = 1,6, U_b = 0,5, \gamma_a = 0,2, \gamma_b = 0,2, \gamma_k = 0,05, \varepsilon_a = \varepsilon_a^0 - 0,3, \varepsilon_b = \varepsilon_b^0 - 0,1; T = 0,2$ . Сплошные линии соответствуют  $\varepsilon_a^\sigma$  и  $\varepsilon_a^{-\sigma}$ . Пунктирные линии соответствуют среднему значению  $\varepsilon_b = (1/2)(\varepsilon_b^\sigma + \varepsilon_b^{-\sigma})$ , потому что атом  $b$  близок к парамагнитному состоянию для этого набора параметров. (б) Зависимость нормированной туннельной проводимости от приложенного напряжения.

## 7. Кулоновские корреляции экситонного типа и сингулярности туннельных характеристик

В настоящее время уже довольно хорошо понято, что в туннельных контактах малых размеров кулоновское взаимодействие может сильно менять наблюдаемую плотность состояний за счёт появления неравновесного заряда в области контакта. Даже эффекты, описание которых может быть построено, как в предыдущем разделе, на основе приближения среднего поля, являются необычными и интересными [28–31].

Помимо этого, существует интересная возможность перенормировки самой туннельной амплитуды за счёт кулоновского взаимодействия. Такое изменение не может быть сведено к модификации локальной плотности состояний и распределения электронов. В этом разделе будет показано, что учёт вершинных поправок к туннельной амплитуде приводит к появлению сингулярностей на вольт-амперных характеристиках [32], аналогичных сингулярностям на краю рентгеновского поглощения в металлах [33]. Эти сингулярности должны проявляться при протекании туннельного тока через глубокие примесные состояния на поверхности полупроводника. Наиболее благоприятными для развития сингулярности являются такие условия, при которых

скорость туннелирования  $\gamma_k$  в металлический электрод (например, в иглу туннельного микроскопа) становится больше, чем скорость релаксации неравновесного заряда на примеси  $\gamma_p$ .

Как будет показано ниже, для  $\gamma_k \gg \gamma_p$  примесный уровень становится почти пустым при напряжении на туннельном контакте, превосходящем энергию уровня примесного состояния. В каком-то смысле можно сказать, что в этот момент скачком "включается" потенциал примесного иона. По аналогии со случаем скачкообразного (по времени) включения потенциала ионного остова при выбивании электрона гамма-излучением можно ожидать появления степенных сингулярностей на вольт-амперных характеристиках. Насколько эта аналогия действительно существует, может быть выяснено только на основе вычислений в рамках неравновесной диаграммной техники. Такое предположение было сделано в работе Ларкина и Матвеева [20], однако авторы ограничились только качественными рассуждениями о перенормировке туннельной амплитуды, что привело к ряду ошибочных выводов.

Рассмотрим туннельную систему типа полупроводник–примесное состояние–металл, описываемую гамильтонианом (3). Для простоты мы остановимся на случае "одноэлектронной нейтральной" примеси: в равновесных условиях из-за кулоновского взаимодействия состояние примеси однократно заполнено. При достаточно больших туннельных константах кулоновское взаимодействие на самой примеси можно не принимать во внимание, если нас интересует проводимость при напряжениях, близких к положению уровня энергии примеси  $\varepsilon_d$ . В этих условиях неравновесное заполнение уровня становится малым, что и подавляет внутриатомные кулоновские эффекты. Заметим также, что режим Кондо заведомо разрушается при таких больших напряжениях [34, 35] и не даёт особенностей на туннельных характеристиках.

Самое существенное, что в этой задаче учитывается кулоновское взаимодействие электрона (или, вернее, дырки) на примеси с электронами проводимости в металле:

$$\hat{H}_{\text{int}} = \sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} W_{\mathbf{k}\mathbf{k}'} c_{\mathbf{k}\sigma}^+ c_{\mathbf{k}'\sigma'} (1 - c_{d\sigma'}^+ c_{d\sigma'}). \quad (55)$$

Гамильтониан  $H_{\text{int}}$  описывает рассеяние электронов проводимости на кулоновском потенциале иона, который перестаёт быть скомпенсированным, если электрон переходит с локализованного состояния в берега контакта. Поскольку при выбранных условиях режим Кондо не формируется, достаточно рассмотреть ток в первом порядке по туннельной амплитуде  $T_{kd}$ . Кулоновское взаимодействие с положительно заряженной дыркой на примеси не приводит к перевороту спина электронов, поэтому можно рассматривать перенормировку туннельной амплитуды для каждого спина независимо. Для упрощения всех выражений мы будем заменять точное кулоновское взаимодействие константой, равной средней величине экранированного кулоновского потенциала, что соответствует учёту s-рассеяния электронов проводимости точечным потенциалом  $W_{\mathbf{k}\mathbf{k}'} = W$ .

Вычисления тока в рамках неравновесной диаграммной техники можно начать с выражения, записанного в терминах функции Грина  $G^<$  (в последующих формулах

заряд электрона  $e = 1$ ):

$$I(V) = \text{Im} J(V), \quad J(V) = i \sum_{\mathbf{k}, \sigma} \int d\omega T_{kd} G_{\mathbf{k}\sigma}^{\sigma<}. \quad (56)$$

Здесь введена туннельная "функция отклика"  $J(V)$ , так как именно эта величина перенормируется под действием кулоновского взаимодействия.

Чтобы не рассматривать отдельно влияние постоянной части кулоновского взаимодействия от ионного остова, проще перейти к дырочному представлению для электронов на примеси:  $d_{\sigma'} = c_{d\sigma'}^+$ . Тогда кулоновское взаимодействие есть просто

$$\hat{H}_{\text{int}} = W \sum_{\mathbf{k}\mathbf{k}'\sigma\sigma'} c_{\mathbf{k}\sigma}^+ c_{\mathbf{k}'\sigma'} d_{\sigma'}^+ d_{\sigma'} \quad (57)$$

с константой  $W < 0$ . Мы не будем вводить специального обозначения для дырочных функций Грина, понимая под  $G_{dd}$  и  $G_{kd}$  функции с дырочными операторами.

Если пренебречь кулоновским взаимодействием в низшем порядке по  $T_{kd}$ , то немедленно получается обычное выражение для туннельного тока в терминах дырочных функций Грина:

$$J^0(V) = i \sum_{\mathbf{k}, \sigma} \int d\omega T_{kd}^2 (G_{\mathbf{k}\sigma}^{\sigma<}(\omega) G_{dd}^{\sigma\text{R}}(-\omega) + G_{\mathbf{k}\sigma}^{\sigma\text{R}}(\omega) G_{dd}^{\sigma>}(-\omega)). \quad (58)$$

Подставляя соответствующие выражения для келдышевских функций Грина, после суммирования по импульсам  $k$  получаем:

$$J^0(V) = \gamma_k \int d\omega \left[ \frac{n_k^0(\omega)}{\omega + eV - \varepsilon_d - i\gamma} - \frac{-i\gamma n_d(\omega)}{(\omega + eV - \varepsilon_d)^2 + \gamma^2} \right], \quad (59)$$

где появилась стандартная скорость туннельных переходов  $\gamma_k = T_{kd}^2 v \pi$ , а  $v$  — плотность состояний в металлической игле. Кинетический параметр  $\gamma = \gamma_k + \gamma_p$ , где  $\gamma_p$  соответствует скорости релаксации электронного распределения на локализованном состоянии. В данной системе эта скорость релаксации также определяется туннельными переходами, но в объёмные состояния полупроводника:  $\gamma_p = T_{pd}^2 v_p \pi$ . (В общем случае в параметр релаксации  $\gamma_p$  могут быть включены и другие типы релаксационных процессов.)

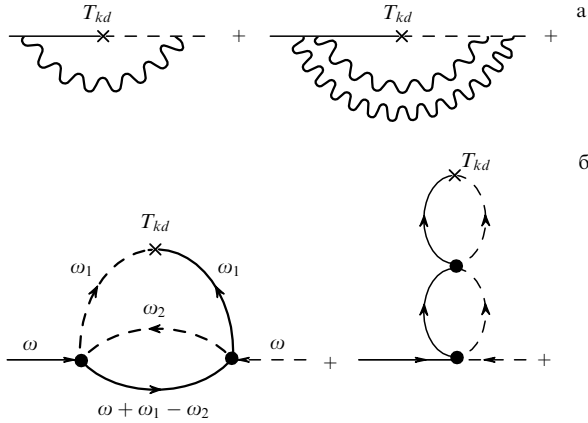
Неравновесные числа заполнения  $n_d(\omega)$  в формуле (59) — электронные, определяющиеся из кинетических уравнений для функции  $G^<$ . Для низких температур из уравнения (59) получается:

$$J^0(V) = \gamma_k \left[ \ln X - i \left( \frac{\gamma_k}{\gamma} \arccot \left( \frac{\varepsilon_d - eV}{\gamma} \right) + \frac{\gamma_p}{\gamma} \arccot \left( \frac{\varepsilon_d}{\gamma} \right) \right) \right], \quad (60)$$

где через  $X$  обозначено

$$X = \frac{\varepsilon_d - eV + i\gamma}{D}, \quad (61)$$

$D$  — ширина зоны проводимости в металле (или, точнее, величина, обратно пропорциональная плотности состояний  $v$ ). Выделив из  $\ln X$  мнимую часть, можно воспроизвести из (56), (59), (60) обычные формулы для туннель-



**Рис. 16.** Кулоновские поправки к туннельной амплитуде  $T_{kd}$ . Сплошными линиями показаны функции Грина электронов в металле  $G_k$ , штриховыми — на примеси  $G_d$ . (а) Лестничное приближение, (б) паркетные диаграммы (кулоновские волнистые линии заменены для компактности чёрными кружками в вершинах взаимодействия).

ного тока:

$$J^0(V) = \gamma_k \ln |X| + i \frac{\gamma_k \gamma_p}{\gamma} \left[ \operatorname{arccot} \left( \frac{\varepsilon_d - eV}{\gamma} \right) - \operatorname{arccot} \left( \frac{\varepsilon_d}{\gamma} \right) \right]. \quad (62)$$

Эти выражения были записаны в несколько нестандартном виде для связи с дальнейшим изложением, так как величина  $J^0(V)$  является базисным элементом для высших порядков теории возмущений.

Рассмотрим, к чему приводит влияние кулоновского взаимодействия на туннельную амплитуду [36]. Появляющиеся вершинные поправки не всегда являются малыми, и ответ, полученный на основании многочастичного подхода, сильно отличается от одноэлектронной картины вблизи пороговых напряжений (когда уровень Ферми металла пересекает примесный уровень). Поправки к туннельной амплитуде показаны на рис. 16. Легко убедиться, что первая поправка имеет логарифмическую расходимость вблизи пороговых значений напряжения  $eV = \varepsilon_d$ , обрезанную конечными скоростями релаксации и туннельных переходов:

$$J^1(V) = i \sum_{\mathbf{k}, \sigma} \int d\omega T_{kd} (G_{\mathbf{k}\mathbf{k}}^{\sigma<}(\omega) G_{dd}^{\sigma\mathbf{R}}(-\omega) T_{kd}^{1++} - G_{\mathbf{k}\mathbf{k}}^{\sigma\mathbf{R}}(\omega) G_{dd}^{\sigma>}(-\omega) T_{kd}^{1--}). \quad (63)$$

Эффективные туннельные матричные элементы меняются из-за кулоновского взаимодействия следующим образом:

$$T_{kd}^{1--} = -iT_{kd} W \sum_{\mathbf{k}, \sigma} \int d\omega_1 (G_{\mathbf{k}\mathbf{k}}^{\sigma<}(\omega_1) G_{dd}^{\sigma\mathbf{R}}(-\omega_1) + G_{\mathbf{k}\mathbf{k}}^{\sigma\mathbf{R}}(\omega_1) G_{dd}^{\sigma>}(-\omega_1)). \quad (64)$$

Как мы видели при выводе формул (58), (59), первая комбинация функций Грина под интегралом:  $G_{\mathbf{k}\mathbf{k}}^{\sigma<} G_{dd}^{\sigma\mathbf{R}}$  даёт логарифмически большой вклад при  $|\varepsilon_d - eV| \leq \gamma$ . Поэтому, даже если кулоновское взаимодействие само по себе не является сильным:  $Wv \ll 1$ , то параметр  $Wv \ln(\gamma^2/D^2)$  может быть порядка или больше единицы. Туннельная вершина  $T_{kd}^{1++}$  отличается от  $T_{kd}^{1--}$

только знаком, так что можно записать:

$$T_{kd}^{1++} = -T_{kd}^{1--} = -T_{kd} L, \quad (65)$$

где величина  $L$  выражается как

$$L = Wv [\ln |X| + i\Phi], \quad (66)$$

$$\Phi = \frac{\gamma_p}{\gamma} \left[ \operatorname{arccot} \left( \frac{\varepsilon_d - eV}{\gamma} \right) - \operatorname{arccot} \left( \frac{\varepsilon_d}{\gamma} \right) \right].$$

Возникновение особенностей в туннельной проводимости проще всего описывается в лестничном приближении, показанном на рис. 16а.

Лестничные диаграммы могут быть просуммированы точно, так как для каждой пары гринских функций интегрирование по частотам выполняется независимо. Тогда

$$J^{ladd}(V) = J^0(V) [1 + L + L^2 + \dots] = \frac{J^0(V)}{1 - L}. \quad (67)$$

Используя формулы (60), (66), получаем для тока следующую зависимость от напряжения:

$$\begin{aligned} I(V) &= \operatorname{Im} J^{ladd}(V) = \gamma_k \operatorname{Im} \left[ \frac{\ln |X| + i\Phi}{1 - Wv(\ln |X| + i\Phi)} \right] = \\ &= \frac{\gamma_k \gamma_p}{\gamma} \left[ \operatorname{arccot} \left( \frac{\varepsilon_d - eV}{\gamma} \right) - \operatorname{arccot} \left( \frac{\varepsilon_d}{\gamma} \right) \right] \times \\ &\times \left\{ \left( 1 - Wv \ln \left[ \frac{(\varepsilon_d - eV)^2 + \gamma^2}{D^2} \right] \right)^2 + \right. \\ &\left. + (Wv)^2 \frac{\gamma^2}{\gamma^2} \left[ \operatorname{arccot} \left( \frac{\varepsilon_d - eV}{\gamma} \right) - \operatorname{arccot} \left( \frac{\varepsilon_d}{\gamma} \right) \right]^2 \right\}^{-1}. \quad (68) \end{aligned}$$

На обычную ступенчатую зависимость тока от напряжения, связанную с прохождением локализованного уровня, накладывается пик, дающий дополнительное возрастание тока вблизи порогового напряжения  $eV = \varepsilon_d$ . Известно, что если логарифмический параметр  $Wv \ln(\gamma^2/D^2)$  становится слишком большим, то лестничное приближение "переоценивает роль" возмущения, приводя к слишком сильным сингулярностям. Так, для  $Wv \ln(\gamma^2/D^2) \geq 1$  формула (68) давала бы два острых пика, симметрично расположенных по напряжению вблизи  $\varepsilon_d$ . Такое поведение является артефактом лестничного приближения, хотя оно правильно указывает на наличие дополнительного пика вблизи порогового напряжения.

Для диаграммных рядов теории возмущений такого типа, как в этой задаче, можно улучшить лестничное приближение, суммируя так называемые паркетные графики. Дело в том, что при наличии большого логарифмического параметра лестничные диаграммы являются не единственным классом "сингулярных" диаграмм, т.е. содержащих максимальную степень логарифма. Если посмотреть на первую диаграмму на рис. 16б, то можно заметить, что появляется новая "петля" из функций Грина  $G_k G_d$ , интеграл в которой содержит большой логарифм при малой "суммарной энергии"  $(\omega + \omega_1)$ . Но существенный интервал интегрирования по  $\omega$  и  $\omega_1$  — это как раз область малых частот  $\omega$ . Именно эта область вносит наибольший вклад в логарифмический фактор  $L$  при интегрировании других пар функций  $G_{\mathbf{k}\mathbf{k}}^{\sigma<} G_{dd}^{\sigma\mathbf{R}}$ . Тем

самым центральная петля также привносит логарифмический фактор в общий результат.

Метод заключается в том, чтобы оставить в  $n$ -м порядке теории возмущений наиболее расходящиеся члены, пропорциональные  $(Wv)^n L^{n+1}$ , с максимально возможной для данного порядка степенью большого логарифма. Такие члены описываются рядом "паркетных" диаграмм, получающихся последовательной подстановкой двух типов петель вместо простых кулоновских вершин, как это сделано для построения двух диаграмм второго порядка на рис. 16б из диаграммы первого порядка. Паркетные диаграммы описывают рассеяние электронов в металле на кулоновском потенциале примеси в двух "наиболее сингулярных" каналах. Для такого ряда диаграмм можно написать приближённые интегральные уравнения, которые являются некоторым обобщением уравнения Бете–Солпитера. Интегральные уравнения могут быть решены с логарифмической точностью [37, 38].

С паркетной ситуацией "в чистом виде" мы имеем дело, когда очень малы числа заполнения локального уровня:  $n_d \ll 1$ , т.е. когда  $\gamma_p \ll \gamma_k$ . Тогда весь диаграммный ряд содержит только петли  $(++)$ , последовательно вставляемые на место вершин взаимодействия. В каждой такой петле остаётся только комбинация функций Грина  $G_k^< G_d^R$ , дающая при интегрировании логарифмический фактор  $L$  (66). Суммирование паркетного ряда диаграмм приводит к выражению [37]:

$$J(V) = \gamma_k \frac{1 - \exp(-2L)}{2Wv}. \quad (69)$$

Ток определяется мнимой частью  $J(V)$ , поэтому ответ главного приближения, содержащий только действительную часть логарифмических интегралов, является недостаточным. Из вида выражений (59) ясно, что во все интегралы, содержащие фермиевское распределение  $n_k$ , мнимая часть  $i\gamma$  входит аддитивно с энергией  $\varepsilon_d$ . Поэтому та же сумма диаграмм приведёт с логарифмической точностью к ответу (69), в котором  $L = Wv \ln X$  ( $X = (\varepsilon_d - eV + i\gamma)/D$ ).

Этот результат обобщается и на случай произвольных значений  $n_d$ . Самыми старшими по логарифмическому параметру являются вклады в мнимую часть  $J(V)$ , когда только в одной паре функций учитывается слагаемое  $G_k^R G_d^>$ . При этом из формул (62), (66) легко понять, что мнимая часть, пропорциональная  $n_d$ , входит во все выражения точно так же, как и мнимая часть основного логарифмического интеграла. Это позволяет в первом порядке по отношению мнимой части к логарифмически большой действительной части написать из формулы (69) ответ для  $\text{Im} J(V)$ :

$$I(V) = \text{Im} J(V) = \gamma_k \left[ \frac{\partial}{\partial \ln X} \frac{1 - \exp(-2L)}{2Wv} \right] \Phi. \quad (70)$$

В этом выражении  $L$  уже не содержит мнимой части, а только действительный логарифм (см. (66)). Подставляя явные выражения для  $L$  и  $\Phi$ , получаем окончательное выражение для тока:

$$I(V) = \frac{\gamma_k \gamma_p}{\gamma} \left[ \frac{D^2}{(eV - \varepsilon_d)^2 + \gamma^2} \right]^{-Wv} \times \left[ \text{arccot} \left( \frac{\varepsilon_d - eV}{\gamma} \right) - \text{arccot} \left( \frac{\varepsilon_d}{\gamma} \right) \right], \quad (71)$$

где  $W < 0$ . Паркетное приближение является более точным, чем лестничное, позволяя описать поведение вольт-амперных характеристик вблизи порога. Заметим, однако, что при учёте мнимых частей функций Грина, который необходимо делать в туннельных задачах, ограничение паркетным рядом диаграмм является приближением, так как, строго говоря, мы пренебрегаем рядом диаграмм, имеющих тот же порядок величины по степеням  $\ln |X|$ . Тем не менее и лестничное приближение, и паркетный ряд диаграмм однозначно указывают на появление пика на вольт-амперных характеристиках, связанного с кулоновскими корреляционными эффектами. Полученная формула (71) позволяет оценить величину дополнительного пика и общее поведение туннельных характеристик.

Помимо рассмотренного случая, когда уход электрона с уровня примеси создаёт положительно заряженную дырку, вполне осуществима и ситуация, когда при другой полярности напряжения электрон после туннелирования будет заполнять состояние на примеси, создавая дополнительный отталкивательный потенциал для электронов в металле. Тогда выражение (71) также будет определять туннельный ток, но знак кулоновского взаимодействия должен быть изменён на противоположный:  $W > 0$ . Вблизи порога кулоновские корреляции приводят к подавлению тока, зависимость тока от напряжения имеет степенной характер.

Примеры туннельных характеристик, рассчитанные для этих двух возможных случаев, показаны на рис. 17.

Заметим, что в работе [39] было высказано правильное предположение о возможном появлении такой особенности. Однако попытка использовать язык коэффициента прохождения, найденного для равновесной ситуа-

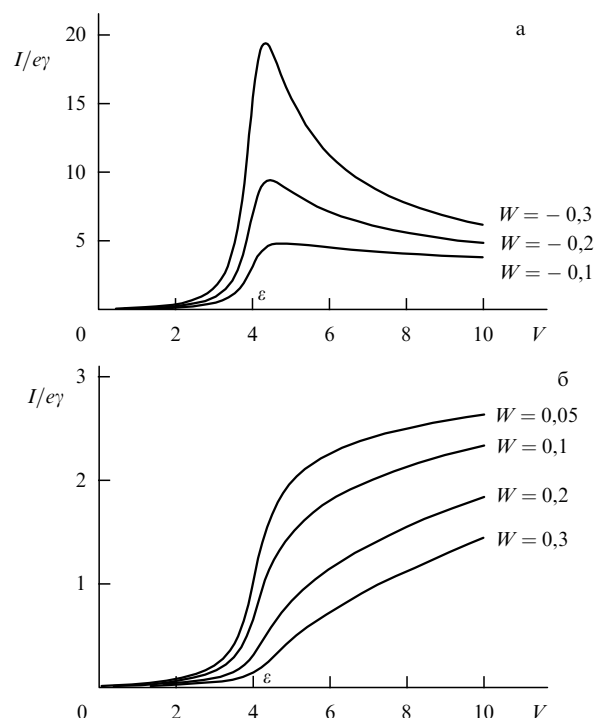


Рис. 17. Вольт-амперные характеристики для различной величины кулоновского взаимодействия. Ток отложен в безразмерных единицах  $I/e\gamma$ . (а)  $w = Wv < 0$ ,  $\varepsilon_d = 0,4$  эВ, (б)  $w = Wv > 0$ ,  $\varepsilon_d + U = 0,4$  эВ,  $\gamma_i/\gamma = 3$ ,  $\varepsilon_d/\gamma = 40$ .

ции, привела к совершенно неправильной формуле для тока при произвольных  $\gamma_k, \gamma_p$ . Также был сделан ошибочный вывод о том, что паркетное приближение наиболее адекватно случаю полного заполнения,  $n_d \simeq 1$ , в то время как анализ неравновесной ситуации показывает, что простой паркетный ряд воспроизводится в противоположном пределе,  $n_d \ll 1$ . Возможность существования подавления проводимости вблизи порога при другой полярности напряжения также была упущена из вида.

При экспериментальном исследовании поверхности GaAs, легированного атомами Mn, было обнаружено [33], что туннельные характеристики, измеренные в точке нахождения примеси, существенно отличаются от характеристик над остальной частью поверхности и соответствуют поведению, описываемому формулой (71).

Результаты расчётов показаны на рис. 17. Основная особенность заключается в том, что сама туннельная проводимость вблизи атома примеси подавлена при положительных напряжениях 0,5–1,5 В, а нормированная дифференциальная проводимость имеет пик. Именно такое необычное поведение описывает формула (71) при положительном знаке кулоновского взаимодействия.

Сингулярное поведение вольт-амперных характеристик может возникать при туннелировании через систему двух связанных квантовых точек при наличии кулоновского взаимодействия зарядов, локализованных в разных точках. При этом, благодаря взаимодействию локализованных электронов с непрерывным спектром в берегах контакта, уровни энергии электронов в квантовых точках приобретают конечную ширину. Поведение вольт-амперных характеристик при изменении напряжения на контакте описывается формулами, аналогичными (71), в которых роль ширины зоны  $v^{-1}$  играет туннельная ширина уровня  $\gamma$ .

## 8. Заключение

В данном обзоре приведено описание экспериментальных данных, полученных методом сканирующей туннельной микроскопии/спектроскопии, в которых наиболее ярко проявилось влияние неравновесных туннельных эффектов и реконструкции электронного спектра, а также дано их теоретическое объяснение, основанное на самоогласованном учёте неравновесного распределения электронов и изменения электронной плотности состояний в области туннельного контакта при протекании туннельного тока. Изложенная теория объясняет появление индуцированной туннельной проводимости в низкоразмерных структурах и структурах нанометровых размеров за счёт возникновения связанных состояний в запрещённой области спектра и при конечной скорости релаксации неравновесных электронов.

Представленный в обзоре теоретический подход позволяет понять причины экспериментально обнаруженной аномальной ширины запрещённой зоны в полупроводниках, обусловленной появлением неравновесных зарядов в области туннельного контакта. Изложенный теоретический анализ объясняет появление пиков туннельной проводимости вблизи границ запрещённой зоны полупроводника, положение которых зависит от геометрии эксперимента и слабо меняется при изменении начальных энергий примесных уровней. Возникновение пиков туннельной проводимости в запрещённой зоне

полупроводника сопровождается уменьшением плотности состояний в области разрешённого спектра по сравнению с её невозмущённым значением. Теоретический анализ неравновесных и интерференционных эффектов при туннелировании через взаимодействующие примесные атомы в модели Андерсона получил экспериментальное подтверждение при СТС-исследованиях взаимодействующих примесей на поверхности полупроводников.

При экспериментальном исследовании примесных атомов на поверхности полупроводников с глубокими уровнями энергии методом сканирующей туннельной микроскопии/спектроскопии была обнаружена немонотонная зависимость туннельного тока от напряжения на туннельном контакте. В соответствии с предложенной теоретической моделью такое поведение может быть объяснено формированием особенностей на вольт-амперных характеристиках, обусловленных перенормировкой туннельных амплитуд за счёт кулоновского взаимодействия экситонного типа для некоторых видов примесных атомов.

Работа выполнена при поддержке Российского научного фонда, грант 16-12-00072.

## Список литературы

1. Binning G et al. *Phys. Rev. Lett.* **50** 120 (1983)
2. Marczinowski F et al. *Phys. Rev. Lett.* **99** 157202 (2007)
3. Feenstra R M *Phys. Rev. Lett.* **63** 1412 (1989)
4. Madhavan V et al. *Phys. Rev. B* **64** 165412 (2001)
5. Morgenstern M et al. *Phys. Status Solidi B* **210** 845 (1998)
6. Mahieu G et al. *Phys. Rev. Lett.* **94** 026407 (2005)
7. Sullivan J M et al. *Phys. Rev. B* **68** 235324 (2003)
8. Wolf E L *Principles of Electron Tunneling Spectroscopy* (New York: Oxford Univ. Press, 1985); Пер. на русск. яз.: Вольф Е Л *Принципы электронной туннельной спектроскопии* (Киев: Наукова думка, 1990)
9. Маслова Н С и др. *Письма в ЖЭТФ* **67** 130 (1998); Maslova N S et al. *JETP Lett.* **67** 146 (1998)
10. Maslova N S et al. *Phys. Status Solidi A* **131** 35 (1992)
11. Maslova N S et al. *Solid State Commun.* **95** 507 (1995)
12. Zheng J F et al. *Phys. Rev. Lett.* **72** 1490 (1994)
13. Depuydt A et al. *Phys. Rev. B* **60** 2619 (1999)
14. Маслова Н С, Панов В И, Савинов С В *УФН* **170** 575 (2000); Maslova N S, Panov V I, Savinov S V *Phys. Usp.* **43** 531 (2000)
15. Marczinowski F et al. *Phys. Rev. B* **77** 115318 (2008)
16. Depuydt A et al. *Appl. Phys. A* **66** S171 (1998)
17. Arseev P I, Maslova N S, Panov V I, Savinov S V *ЖЭТФ* **121** 225 (2002); *JETP* **94** 191 (2002)
18. Anderson P W *Phys. Rev.* **124** 41 (1961)
19. Mahan G D *Phys. Rev.* **163** 612 (1967)
20. Matveev K A, Larkin A I *Phys. Rev. B* **46** 15337 (1992)
21. Arseyev P I, Maslova N S, Panov V I, Savinov S V *Письма в ЖЭТФ* **76** 345 (2002); *JETP Lett.* **76** 287 (2002)
22. Келдыш Л В *ЖЭТФ* **47** 1515 (1964); Keldysh L V *Sov. Phys. JETP* **20** 1018 (1965)
23. Арсеев П И *УФН* **185** 1271 (2015); Arseev P I *Phys. Usp.* **58** 1159 (2015)
24. Arseyev P I, Volkov B A *Solid State Commun.* **78** 373 (1991)
25. Арсеев П И, Маслова Н С *ЖЭТФ* **102** 1056 (1992); Arseev P I, Maslova N S *JETP* **75** 575 (1992)
26. Arseev P I, Maslova N S, Mantsevich V N *ЖЭТФ* **142** 156 (2012); *JETP* **115** 141 (2012)
27. Maslova N S, Arseyev P I, Mantsevich V N *Solid State Commun.* **241** 20 (2016)
28. Bardeen J *Phys. Rev. Lett.* **6** 57 (1961)
29. Feenstra R M, Fein A P *Phys. Rev. B* **32** 1394(R) (1985)

30. Zheng Z F, Salmeron M B, Weber E R *Appl. Phys. Lett.* **64** 1836 (1994)
31. Lengel G et al. *J. Vac. Sci. Technol. B* **11** 1472 (1993)
32. Renner Ch, Fischer O *Phys. Rev. B* **51** 9208 (1995)
33. Arseyev P I, Maslova N S *Solid State Commun.* **108** 717 (1998)
34. Domke C et al. *J. Vac. Sci. Technol. B* **16** 2825 (1998)
35. Pacheroová O et al. *Czech. J. Phys.* **49** 1621 (1999)
36. Арсеев П И и др. *Письма в ЖЭТФ* **72** 819 (2000); Arseev P I et al. *JETP Lett.* **72** 565 (2000)
37. Arseyev P I, Maslova N S, Savinov S V *Письма в ЖЭТФ* **68** 299 (1998); *JETP Lett.* **68** 320 (1998)
38. Beenakker C W J *Phys. Rev. B* **44** 1646 (1991)
39. Kaminski A, Glazman L I *Phys. Rev. B* **61** 15927 (2000)

### Tunneling features in semiconductor nanostructures

P.I. Arseev<sup>(1,2)</sup>, V.N. Mantsevich<sup>(3)</sup>, N.S. Maslova<sup>(3)</sup>, V.I. Panov<sup>(3)</sup>

<sup>(1)</sup> *Lebedev Physical Institute, Russian Academy of Sciences, Leninskii prosp. 53, 119991 Moscow, Russian Federation*

<sup>(2)</sup> *National Research University "Higher School of Economics", ul. Myasnitskaya 20, 101000 Moscow, Russian Federation*

<sup>(3)</sup> *Lomonosov Moscow State University,*

*Leninskie gory 1, 119991 Moscow, Russian Federation*

*E-mail: (1,2)ars@lpi.ru, (3)vmantsev@gmail.com, (3)nsmaslova@gmail.com, (3)panov@spmlab.phys.msu.ru*

The most telling scanning tunneling microscopy/spectroscopy (STM/STS) data available on the influence of nonequilibrium tunneling effects and electronic spectra reconstruction are reviewed and theoretically explained by self-consistently considering the nonequilibrium electron distribution and the change (due to the tunneling current) in the electron density of states near the tunneling junction. The paper discusses the basic ideas of self-consistent tunneling theory, which forms the basis for experimental research and which allows many effects observed in STM/STS experiments to be explained as well as new phenomena to be predicted.

**Keywords:** tunneling, scanning tunneling microscopy/spectroscopy, localized states, Keldysh diagram technique, nonequilibrium processes, interparticle interaction

PACS numbers: 05.60.Gg, 68.37.Ef, 73.40.Gk, **73.63. – b**

Bibliography — 39 references

*Received 12 December 2016*

*Uspekhi Fizicheskikh Nauk* **187** (11) 1147–1168 (2017)

*Physics–Uspekhi* **60** (11) (2017)

DOI: <https://doi.org/10.3367/UFNr.2017.01.038055>

DOI: <https://doi.org/10.3367/UFNe.2017.01.038055>