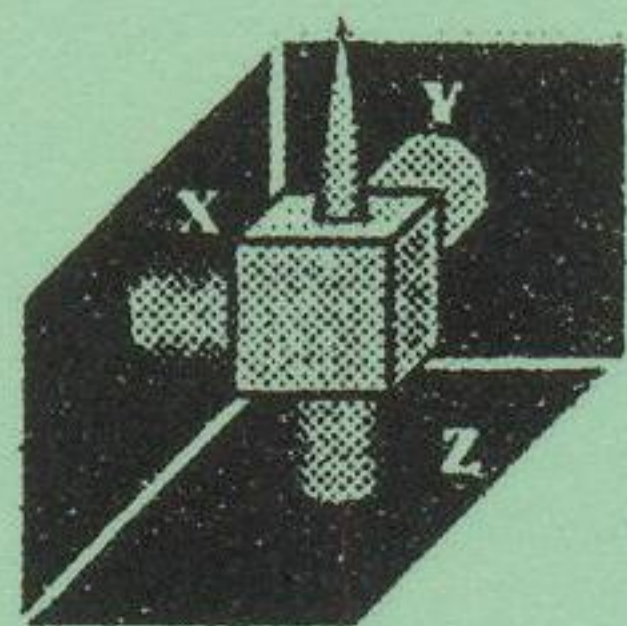


В.А. ВАСИН, Е.Н. ИВАШОВ, С.В. СТЕПАНЧИКОВ

# НАНОТЕХНОЛОГИЧЕСКИЕ ПРОЦЕССЫ И ОБОРУДОВАНИЕ ЭЛЕКТРОННОЙ ТЕХНИКИ





УДК 537.331

И 25

Рецензенты: **Васичев Б.Н.**, доктор физико-математических наук, профессор Российской экономической академии им. Г.В. Плеханова

**Кузькин В.И.**, кандидат технических наук, доцент МАТИ – Российского государственного технологического университета им. К.Э. Циолковского

**В.А. Васин, Е.Н. Ивашов, С.В. Степанчиков**

**И 25** Нанотехнологические процессы и оборудование электронной техники. – М.: МИЭМ.2009-264с.

**ISBN 978-5-94506-245-0**

Рассмотрены общие и специальные вопросы, связанные с описанием, анализом и перспективами применения методов и технологий получения различных наноразмерных структур. Представлены функциональные устройства и системы, а также оборудование для реализации нанотехнологических процессов.

**ISBN 978-5-94506-245-0**

УДК 537.331

©В.А. Васин, Е.Н. Ивашов, С.В. Степанчиков, 2009

©Московский государственный институт  
электроники и математики, 2009



Е.Н. ИВАШОВ

С.В.СТЕПАНЧИКОВ

**НАНОТЕХНОЛОГИЧЕСКИЕ  
ПРОЦЕССЫ И ОБОРУДОВАНИЕ  
ЭЛЕКТРОННОЙ ТЕХНИКИ**

**ФЕДЕРАЛЬНОЕ АГЕНСТВО ПО ОБРАЗОВАНИЮ**

**Государственное образовательное учреждение высшего профессионального образования**

**Московский государственный институт электроники и математики**

**(Технический университет)**

**Е.Н. ИВАШОВ**

**С.В.СТЕПАНЧИКОВ**

**НАНОТЕХНОЛОГИЧЕСКИЕ  
ПРОЦЕССЫ И ОБОРУДОВАНИЕ  
ЭЛЕКТРОННОЙ ТЕХНИКИ**

**Утверждено Редакционно-издательским советом института  
в качестве учебного пособия**

Москва 2009



УДК 537.331

ББК

И

Рецензенты: **Васичев Б.Н.**, доктор физико-математических наук, профессор Российской экономической академии им. Г.В. Плеханова

**Кузькин В.И.**, кандидат технических наук, доцент "МАТИ" – Российского государственного технологического университета им. К.Э. Циалковского

**Ивашов Е.Н., Степанчиков С.В.**

И Нанотехнологические процессы и оборудование электронной техники: Учебное пособие. – М.: Московский государственный институт электроники и математики (технический университет), 2009.- 264с.

В учебном пособии рассмотрены общие и специальные вопросы, связанные с описанием, анализом и перспективами применения методов и технологий получения различных наноразмерных структур. Представлены функциональные устройства и системы, а также оборудование для реализации нанотехнологических процессов.

Утверждено Редакционно-издательским советом МИЭМ (ТУ) в качестве учебного пособия.

**ISBN**

УДК 537.331

ББК

©Е.Н.Ивашов, С.В.Степанчиков, 2009

©Московский государственный институт электроники и математики (технический университет), 2009



## Предисловие

Основополагающие технологии, оборудование и материалы всегда играли главенствующую роль в истории развития человечества. По мнению специалистов, наступивший век будет веком нанотехнологий, которые и будут определять его облик.

Под термином "нанотехнология" понимают создание и исследование материалов, устройств и систем, структура которых реализуется в нанометровом диапазоне пространственных размеров.

В отличие от традиционных технологий для нанотехнологии характерен "индивидуальный" подход, при котором внешнее управление достигает отдельных молекул и атомов, что позволяет создавать из них как новые классы устройств с характерными нанометровыми размерами, так и "бездефектные" материалы с принципиально новыми физическим, химическими и биологическими свойствами.

Одной из составных частей нанотехнологии является создание многокомпонентных изделий электронной техники с новыми потребительскими качествами и назначением, таких как сверхскоростные микропроцессоры, сверхъемкая память, специальные нанороботы и др.

На сегодняшний день в нанотехнологии реализуется большое количество научно-технических программ во многих развитых странах мира. Нанотехнологии используются в оборонной промышленности и космонавтике, электронной и вычислительной технике, биотехнологии и медицине, энергетике и охране окружающей среды.

Перспектива использования нанотехнологий ставит новые задачи и в системе образования, в первую очередь – высшего профессионального.

Настоящее учебное пособие ставит целью ознакомить студентов, аспирантов и специалистов с основными нанотехнологическими процессами и оборудованием электронной техники, их возможностями и ролью в недалеком будущем.

Введение знакомит читателей с историческими и идеологическими аспектами нанотехнологии.

В первой главе рассмотрены особенности применения пьезоэлектрических сканеров в зондовой микроскопии, представлена модель сканера точной подстройки на основе пьезоэлектрического элемента.

Вторая глава посвящена вопросам построения магнитострикционных систем перемещения в нанотехнологии, рассмотрены эффекты магнитострикции и гигантской магнитострикции, магнито-упругие эффекты, информационная модель магнитострикционного эффекта.



В третьей и четвертой главе учебного пособия отражены вопросы формирования объектов в лазерно-зондовой и туннельно-зондовой нанотехнологии, в том числе особенности поглощения светового излучения материалом, механическое воздействие лазерного излучения на атомы, модель атомного нанозонда, магнитные свойства наноструктур, пути повышения плотности записи в устройствах магнитной памяти, методы получения нанокристаллических твердых тел для их последующей магнитоориентирующей обработки.

Пятая глава рассказывает об особенностях формирования искусственных нейронных сетей в туннельно-зондовой нанотехнологии. В главе рассмотрены структура искусственной нейронной системы и ее вычислительные особенности, вопросы моделирования элементов искусственных нейронных сетей на полупроводниковых нанообъектах - квантовых точках и ямах.

В шестой главе работы рассмотрены квантовые наноразмерные структуры для систем кодирования и криптографии, основные понятия защищенных информационных технологических систем, их определения и свойства, схемы управления квантовыми наноструктурами

Седьмая глава посвящена вопросам формирования рельефа подложки в туннельно-зондовой нанотехнологии. Рассмотрены физические основы формирования нанорельефа подложек.

В восьмой главе рассмотрены приповерхностные свойства нанообъектов, гетероструктур и нанообластей с различным химическим и фазовым составом, методы исследования данных свойств и эффектов, возникающих при взаимодействии с исследуемыми образцами.

Девятая глава знакомит читателей с особенностями управления качеством распознавания образов в кластерных системах обработки информации.

В каждой из вышеперечисленных глав учебного пособия представлены функциональные устройства и системы для реализации нанотехнологических процессов.

В десятой главе представлено оборудование молекулярно-лучевой эпитаксии для выращивания метастабильных наноструктур.



## **Введение. Исторические и идеологически аспекты нанотехнологии**

В общем виде развитие науки и техники представляет собой эволюционный и революционный процесс. Причем при переходе на более высокий уровень развития не исключена возможность возврата. Развитие науки - это смена парадигм, периодические скачкообразные изменения в стиле мышления, методологии и методике научного познания.

Рассмотрим это на примере развития нанотехнологии. Один *нанометр* (от греческого "нано" - карлик) равен одной миллиардной части метра. На этом расстоянии можно вплотную расположить примерно 10 атомов. Пожалуй, первым ученым, использовавшим эту единицу измерения, был Альберт Эйнштейн, который в 1905 г. теоретически доказал, что размер молекулы сахара равен одному нанометру. Но только через 26 лет немецкие физики Эрнст Руска, получивший Нобелевскую премию в 1986 г., и Макс Кнолл создали электронный микроскоп, обеспечивающий 15-кратное увеличение (меньше, чем существовавшие тогда оптические микроскопы), он и стал прообразом нового поколения подобных устройств, позволивших заглянуть в наномир.

1932 год. Голландский профессор Фриц Цернике, Нобелевский лауреат 1953 г., изобрел фазово-контрастный микроскоп - вариант оптического микроскопа, улучшавший качество показа деталей изображения, и исследовал с его помощью живые клетки (ранее для этого приходилось применять красители, убивавшие живые ткани). Интересно, что Цернике предлагал свое изобретение фирме "Цейс", но менеджеры не осознали его перспективности, хотя сегодня такие микроскопы активно применяются в медицине.

1939 год. Компания Siemens, в которой работал Руска, выпустила первый коммерческий электронный микроскоп с разрешающей способностью 10 нм. Днем рождения нанотехнологий считается 29 декабря 1959 г. Профессор Калифорнийского технологического института Ричард Фейнман (Нобелевский лауреат 1965 г.) в своей лекции "Как много места там, внизу" ("There's plenty of room at the bottom"), прочитанной перед Американским физическим обществом, отметил возможность использования атомов в качестве строительных частиц.

1966 год. Американский физик Рассел Янг, работавший в Национальном бюро стандартов, придумал пьезодвигатель, применяемый сегодня в сканирующих туннельных микроскопах и для позиционирования наноинструментов с точностью до 0,01 ангстрем (1 нанометр = 10 ангстрем).

1968 год. Исполнительный вице-президент компании Bell Альфред Чо и сотрудник ее отделения по исследованиям полупроводников Джон Артур обосновали теоретическую возможность использования нанотехнологий в решении задач обработки поверхностей и достижения атомной точности при создании электронных приборов.



1971 год. Рассел Янг выдвинул идею прибора Tорографнер, послужившего прообразом зондового микроскопа. Столь длительные сроки разработки подобных устройств объясняются тем, что наблюдение за атомарными структурами приводит к изменению их состояния, поэтому требовались качественно новые подходы, не разрушающие исследуемое вещество. Правда, вскоре работы над Tорографнер были прекращены, и признание к Янгу пришло только в 1979 г., после чего он получил множество наград.

1974 год. Японский физик Норио Танигучи, работавший в Токийском университете, предложил термин "нанотехнология" (процесс разделения, сборки и изменения материалов путем воздействия на них одним атомом или одной молекулой), быстро завоевавший популярность в научных кругах.

1982 год. В Цюрихском исследовательском центре «ЮМ» физики Герд Бинниг и Генрих Рорер (Нобелевские лауреаты 1986 г. вместе с Эрнстом Руской) создали сканирующий туннельный микроскоп (СТМ), позволяющий строить трехмерную картину расположения атомов на поверхностях проводящих материалов. СТМ действовал по принципу, схожему с заложенным в Tорографнер, но швейцарцы создали его независимо от Янга, добившись значительно большей разрешающей способности и распознав отдельные атомы в кальциево-иридиево-оловянных кристаллах. Главной проблемой в исследовании были фоновые помехи - острие микроскопа, позиционировавшееся с точностью до долей атома, сбивалось от малейших шумов и вибраций на улице.

1985 год. Трое американских химиков: профессор Райсского университета Ричард Смэлли, а также Роберт Карл и Хэрольд Крото (Нобелевские лауреаты 1996 г.) открыли фуллерены - молекулы, состоящие из 60 атомов углерода, расположенных в форме сферы. Эти ученые также впервые сумели измерить объект размером 1 нм.

1986 год. Герд Бинниг разработал сканирующий атомно-силовой зондовый микроскоп, позволивший наконец визуализировать атомы любых материалов (не только проводящих), а также манипулировать ими.

1986 год. Американский ученый Эрик Дрекслер, работавший в лаборатории искусственного интеллекта Массачусетского технологического института, написал книгу "Машины созидания" ("Engines of Creation"), в которой выдвинул концепцию универсальных молекулярных роботов, работающих по заданной программе и собирающих что угодно (в том числе и себе подобных) из подручных молекул. Эта идея была, видимо, навеяна Дрекслеру его основной деятельностью - в задачах искусственного интеллекта идея самовоспроизводящихся устройств встречается постоянно. Ученый уже тогда довольно точно предсказал немало грядущих достижений нанотехнологии, и, начиная с 1989 г. его прогнозы сбываются, причем нередко со значительным опережением сроков.

1987 - 1988 гг. В НИИ "Дельта" под руководством П. Н. Лускиновича заработала первая российская нанотехнологическая установка, осуществляв-



шая направленный уход частиц с острия зонда микроскопа под влиянием нагрева.

1989 год. Ученые Дональд Эйглер и Эрхард Швецер из Калифорнийского научного центра IBM сумели выложить 35 атомами ксенона на кристалле никеля название своей компании. Для первого в мире целевого переноса отдельных атомов в новое место они использовали СТМ производства IBM. Правда, такая надпись просуществовала недолго - атомы быстро разбежались с поверхности. Но сам факт наличия постороннего атома в молекулярной структуре некоторого вещества открывал потенциальную возможность создания молекулярных автоматов, тракующих наличие или отсутствие такого атома в некоторой позиции как логическое состояние.

1991 год. Японский профессор Сумио Лиджима, работавший в компании NEC, использовал фуллерены для создания углеродных трубок (или нанотрубок) диаметром 0,8 нм. На их основе в наше время выпускаются материалы в сто раз прочнее стали. Оставалось научиться делать такие трубки как можно более длинными - их размеры оказались напрямую связаны с прочностью изготавливаемых веществ. Кроме того, открылась возможность собирать из нанотрубок различные наномеханизмы с зацепами и шестеренками.

Компьютерщик Уоррен Робинет и химик Стэн Уильяме, сотрудники университета Северной Каролины, изготовили наноманипулятор - робот размером с человека, состыкованный с атомным микроскопом и управляемый через интерфейс виртуальной реальности. Оператор, манипулируя отдельными атомами, с его помощью мог физически ощущать многократно усиленную отдачу от модифицируемого вещества, что значительно ускоряло работу. Пытаться делать прикладные наноустройства без такого комплекса до того времени было немыслимо.

1991 год. В США заработала первая нанотехнологическая программа Национального научного фонда. Аналогичной деятельностью озаботилось и правительство Японии. А вот в Европе серьезная поддержка таких исследований на государственном уровне началась только с 1997 г.

1997 год. Эрик Дрекслер объявил, что к 2020 г. станет возможной промышленная сборка наноустройств из отдельных атомов. До сего времени почти все его прогнозы сбывались с опережением.

1998 год. Сиз Деккер, голландский профессор Технического университета г. Делфтса, создал транзистор на основе нанотрубок, используя их в качестве молекул. Для этого ему пришлось первым в мире измерить электрическую проводимость такой молекулы. Появились технологии создания нанотрубок длиной 300 нм. В Японии запущена программа "Astroboy" по развитию наноэлектроники, способной работать в условиях космического холода и при жаре в тысячи градусов.

1999 год. Американские ученые - профессор физики Марк Рид (Йельский университет) и профессор химии Джеймс Тур (райский университет) -



разработали единые принципы манипуляции как одной молекулой, так и их цепочкой.

2000 год. Немецкий физик Франц Гиссибл разглядел в кремнии субатомные частицы. Его коллега Роберт Магерле предложил технологию нанотомографии - создания трехмерной картины внутреннего строения вещества с разрешением 100 нм. Проект финансировала компания Volkswagen. Правительство США открыло Национальную нанотехнологическую инициативу (NNI). В бюджете США на это направление выделено 270 млн. долл., коммерческие компании вложили в него в 10 раз больше.

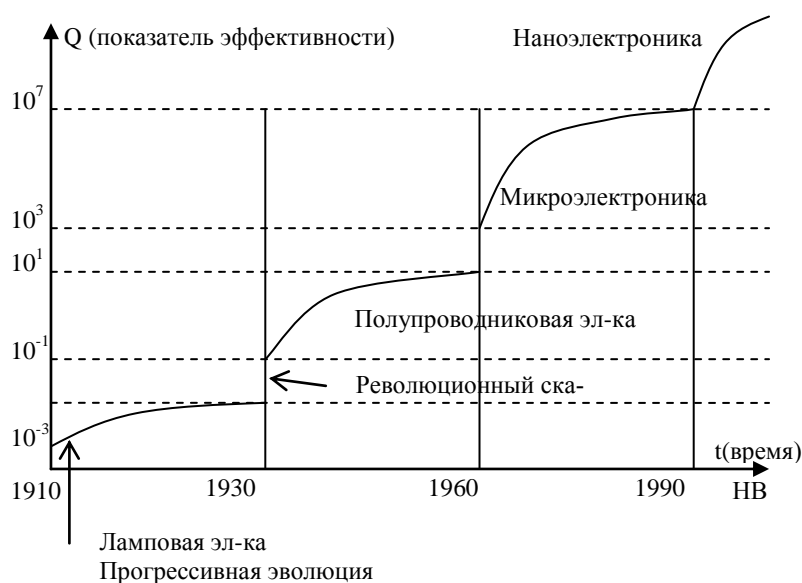
2001 год. Реальное финансирование NNI превысило запланированное (422 млн. долл.) на 42 млн.

2002 год. Сиз Деккер соединил углеродную трубку с ДНК, получив единый наномеханизм. Финансирование NNI составило 697 млн. долл. (на 97 млн. больше плана).

2003 год. Профессор Фенг Лью из университета Юты, используя разработки Франца Гиссибла, с помощью атомного микроскопа построил образы орбит электронов путем анализа их возмущения при движении вокруг ядра.

На NNI отпущено 770 млн. долл. В бюджете NNI 2004 г. заложена сумма 849 млн. долл.

Главным объединяющим началом сообщества ученых являются не нормы профессиональной этики, а единый стиль мышления, т.е. признание данным сообществом ученых определенных фундаментальных теорий и методов исследований. Эти положения в философии называются парадигмой. Каждая теория создается в рамках той или иной парадигмы. Теории, существующие в рамках различных парадигм, несопоставимы. Одна и та же теория не может входить в разные парадигмы без предварительного её серьёзного переосмысления.



**Рис В1.** Схематическая иллюстрация законов прогрессивной эволюции и скачкообразного развития нанoeлектроники



При смене парадигм невозможно осуществить преемственность теорий, т.е. перенести какие-то из старых парадигм в новые. Это относится, в первую очередь, и к нанотехнологии.

На первой фазе своего становления нанотехнология занимается в основном моделированием, с целью углубления фундаментальных представлений в различных областях знания, в первую очередь микроэлектронике, полупроводниковой и ламповой электронике (рис. В1).

В это связи целесообразно рассматривать законы развития электроники, которая составляет основополагающую часть всей нанотехнологии.

1. Закон прогрессивной эволюции электроники. В вакуумных устройствах с одинаковой функцией переход от поколения к поколению происходит при наличии необходимого научно - технического уровня и социально - экономической целесообразности. Прогрессивная эволюция продолжается до максимального значения показателя эффективности  $Q$ , например количества элементов в единице объема кристалла или изделия.

2. Закон скачкообразного развития наноэлектроники. Этот закон отражает революционные изменения в процессе развития (см. рис. В1). Переход к каждой очередной стадии происходит при исчерпывании природных возможностей человека в улучшении показателей эффективности выполнения функций данным устройствам.

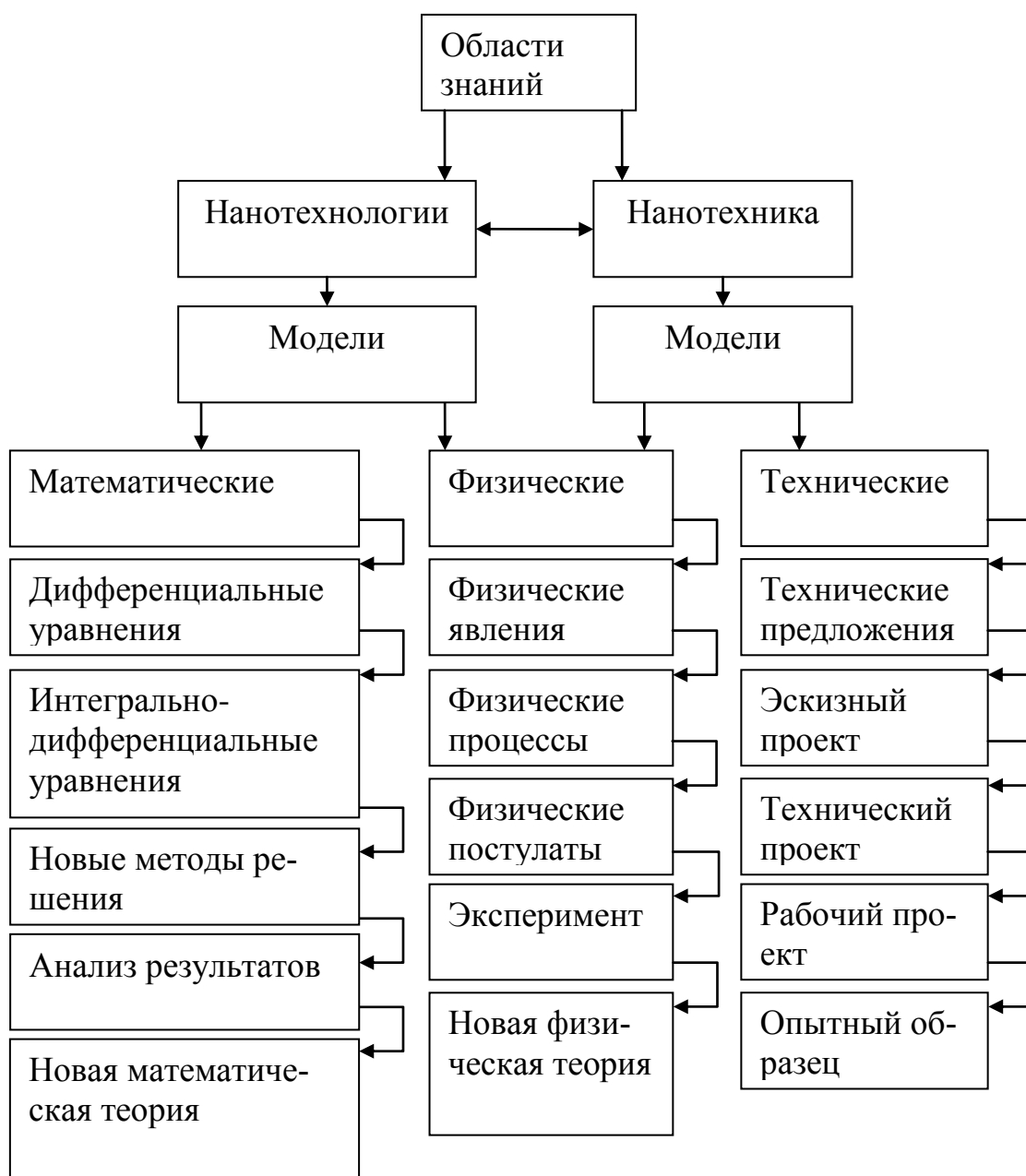
3. Закон соответствия между функцией и структурой. Главная суть закона заключается в том, что в правильно спроектированном электронном устройстве каждый элемент имеет вполне определённую функцию по обеспечению его работоспособности. Исключение элемента приводит к ухудшению какого либо показателя эффективности.

Первый и второй законы развития наноэлектроники тесно связаны с диалектическим законом перехода количественных изменений в качественные. Главная проблема в нанотехнологии – проблема верификации, заключающаяся в невозможности в настоящее время проверить некоторые теории опытным путём. Это связано с тем, что почти во всех экспериментах используется метод косвенных измерений. Соотношение принципов верификации и фальсификации, т. е. не подтверждение на истинность, а опровержение неистинности, это также проблема нанонауки и техники. Наряду с вышеуказанными проблемами в нанонауке и технике, как нигде более очевидными становятся размытые грани между различными категориями философии, такими как причина - повод - следствие, количество - условие - качество, единичное - особенное - всеобщее, сущность - наблюдение - явление, необходимость - действие - случайность, возможность - предрасположенность - действительность и др.

Анализ задач математического, физического и технического моделирования (рис. В2) в нанонауке и технике позволяет констатировать отсутствие в природе *физического нуля*, т. е. абсолютного ничего, пустоты. Так, например,

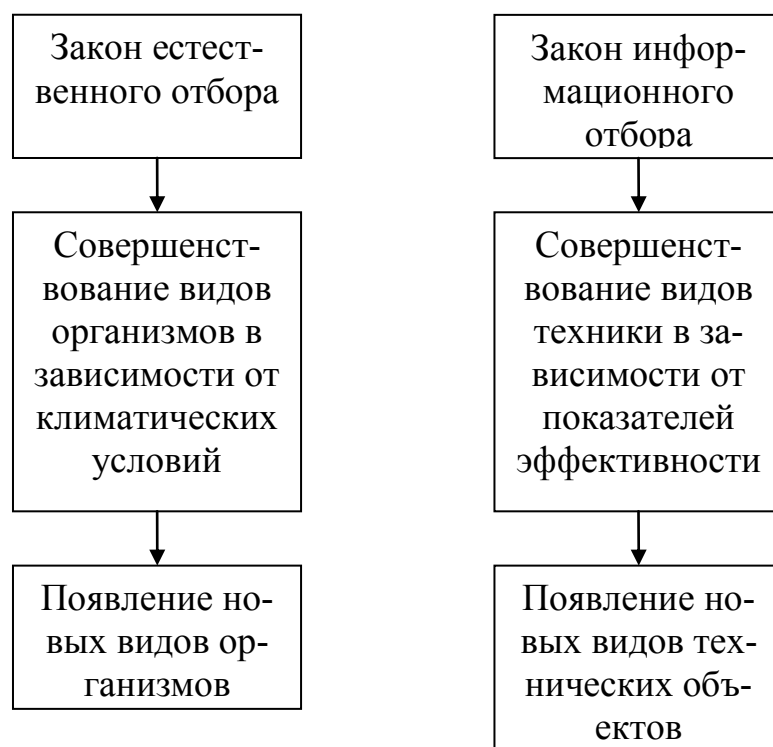


предельное остаточное давление в вакуумной камере - величина бесконечно малая, но не равная нулю, тоже самое можно сказать относительно силы тяжести или абсолютно отрицательной температуры (абсолютного нуля). Таким образом, в нанонауке и технике мы имеем дело с бесконечно малыми (не равными нулю) величинами. А, в общем случае, бесконечно малые величины - антиподы бесконечно больших. Например, при увеличении радиуса до бесконечно большой величины окружность превращается в прямую, а при уменьшении до бесконечно малой - в точку. Таким образом, прямая - синоним бесконечно большой величины, точка - бесконечно малой.



**Рис В2.** Модели нанонауки и нанотехники

Техноэволюция нанотехники осуществляется под действием закона информационного отбора Б. И. Кудрина. Действие этого закона проявляется в наследственных изменениях вида нанотехники точнее - популяций изделий, занимающих определенную экологическую нишу. В отличие от закона естественного отбора Ч. Дарвина в вакуумной технике имеет место более разумная целенаправленная изменчивость: появляются, как правило, только такие новые варианты конструкторских решений, которые по основным показателям (критериям эффективности) обеспечивают повышение конкурентоспособности (рис. В3), а подавляющее большинство изменений связано с улучшением наиболее актуальных показателей, которые в данный момент требуется улучшить, например, скорость откачки вакуумного насоса и величину предельного вакуума.



*Рис В3. Отличительные особенности естественного отбора Ч. Дарвина и информационного отбора Б.И. Кудрина*

Техноценоз - сообщество всех изделий и оборудования конкретного участка, цеха или предприятия для определенного момента или отрезка времени. Существующие НИИ или КБ нанотехники в основном сосредоточены на изучении и проектировании отдельных изделий, а изучением, проектированием техноценозов никто серьезно не занимается.

В каждом конкретном случае существует свой оптимальный состав оборудования в техноценозе, который находится между двумя крайними предельными случаями, когда все изделия в техноценозе различны и все изделия одинаковы.



Задача выбора оптимального состава оборудования техноценоза - очень сложная задача нелинейного программирования.

Попробуем разобраться, какую помощь могут оказать людям нанороботы и какую угрозу для человечества они представляют.

Перспективы просто фантастические, иначе не скажешь. Например, за счёт внедрения в организм молекулярных роботов, предотвращающих старение клеток, а также перестраивающих и "облагораживающих" ткани организма можно будет достигнуть бессмертия человека, не говоря об оживлении и излечении безнадежно больных и людей, которые были заморожены методами крионики.

Наноробот, введённый в организм человека, сможет самостоятельно передвигаться по кровеносной системе и очищать его от микробов или зарождающихся раковых клеток, а саму кровеносную систему - от отложений холестерина. Он сможет изучить, а затем и исправить характеристики тканей и клеток.

В промышленности произойдёт замена традиционных методов производства сборкой молекулярными роботами предметов потребления непосредственно из атомов и молекул, вплоть до персональных синтезаторов и копирующих устройств, позволяющих изготовить любой предмет.

Замена произойдёт и в сельском хозяйстве: комплексы из молекулярных роботов придут на смену "естественным машинам" для производства пищи (растений и животных) их искусственными аналогами. Они будут воспроизводить те же химические процессы, что происходят в живом организме, однако более коротким и эффективным путем.

Биологи смогут "внедряться" в живой организм на уровне атомов, в результате чего, станут возможными и "восстановление" вымерших видов, и создание новых типов живых существ, в том числе биороботов.

В кибернетике произойдёт переход к объёмным микросхемам, а размеры активных элементов уменьшаться до размеров молекул. Рабочие частоты компьютеров достигнут терагерцовых величин. Получат распространение схемные решения на нейроноподобных элементах. Появится долговременная быстродействующая память на белковых молекулах, ёмкость которой будет измеряться терабайтами. Станет возможным "переселение" человеческого интеллекта в компьютер.

За счёт внедрения логических наноэлементов во все атрибуты окружающей среды она станет "разумной" и исключительно комфортной для человека. На всё это, по разным оценкам, понадобится около 100 лет.

Однако новые открытия могут иметь и негативные последствия. Представим себе, что в устройстве, предназначенном для разборки промышленных отходов до атомов, произойдет сбой, и оно начнёт уничтожать полезные вещества биосферы, обеспечивающие жизнь людей. При этом самым неприятным может оказаться то, что это будут нанороботы, способные к самовоспроизводству (саморепликации, размножению).

Можно представить себе и нанороботов, запрограммированных на изготовление уже существующего оружия. Овладев секретом создания подобного робота или каким-то образом достав его, воспроизвести универсального "малыша" в большом количестве сможет небольшая группа людей или даже террорист-одиночка.

Отметим также принципиальную возможность создания, выборочно разрушительных устройств, например воздействующих на определённые этнические группы или заданные географические районы.

Нанороботы, вышедшие из-под контроля, могут стать оружием массового поражения.

Так или иначе, но главный шаг на пути создания нанороботов группа ньюйоркских учёных, по их признанию, уже сделала. Судя по тому, что на создание первой ДНК-машины ушло около 10 лет, первый наноробот появится максимум лет через 5-7.

Устройства микроэлектромеханических систем (MEMS) действуют как и устройства макроразмеров и даже выглядят также - с моторами, передачами и рычагами, изготовленными из стекла, керамики или металла. Но наноразмерные структуры, в частности NEMS, будут строиться и действовать совершенно по-другому: они формируются и функционируют на основе других физических законов. На молекулярном уровне перестают действовать законы механики, используемые для расчетов узлов обычных машин. Законы сопротивления материалов и гидравлики уже не применимы - вместо этого вступают в действие законы квантовой механики, которые приводят к совершенно неожиданным, с точки зрения классической механики, последствиям.

Сегодня практическая нанотехнология ориентирована на решение следующих задач:

- создание твердых тел и поверхностей с требуемой молекулярной структурой;
- создание новых химических веществ посредством конструирования молекул (с участием и без участия химических реакций);
- разработка устройств различного функционального назначения (компоненты наноэлектроники, нанооптики, наноэнергетики, нанороботы и нанокomпьютеры, нанолечения, наноинструменты и т.д.);
- создание наноразмерных самоорганизующихся и самореплицирующихся структур.

Инструментальный базис нанотехнологий, позволяющий ученым и исследователям не только визуализировать атомные структуры, но и манипулировать отдельными атомами и строить новые молекулы, основан на использовании так называемого эффекта туннелирования электронов. Его применение на вершинах зондов специальных конструкций позволило достичь высокой пространственной разрешающей способности управления атомно-молекулярными реакциями в отличие от известных групповых технологий



осаждения материалов, методов оптической литографии, эпитаксии, а также электронной литографии, где высокая энергия фокусируемых электронов приводит к значительному разрушению используемых материалов.

За 20 с небольшим лет с момента появления техники сканирующей зондовой микроскопии и изобретения сканирующего туннельного, а затем и атомно-силового микроскопов, в разных странах были получены впечатляющие результаты по наблюдению наноразмерных частиц и структур на их основе и поставлена задача создания технологических машин, позволяющих осуществить атомно-молекулярную сборку вещества и конструирование отдельных узлов и устройств различного функционального назначения.

Внедрение наносхемотехники и нанороботов позволит создать микроскопические компьютеры небывалой производительности. Более того, они станут саморемонтирующимися и самовоспроизводящимися. Это означает, что в зависимости от потребности вычислительной системы она будет увеличиваться и уменьшаться сама. Применение десятиатомных транзисторов позволит подойти вплотную к имитации мыслительных процессов человека и уже к середине XXI столетия создать настоящий искусственный интеллект - саморазвивающуюся мыслительную среду. Станет возможным также и внедрение человеческого сознания в компьютерные программы.

Впервые идея о новом направлении была высказана лауреатом Нобелевской премии Р.Фейнманом в 1959 г. Позже, в 80-х годах, появились приборы, способные оперировать с отдельным атомом, например, взять его и переставить на другое место. Созданы отдельные элементы нанороботов: опытный механизм шарнирного типа на основе нескольких цепочек ДНК, способный сгибаться и разгибаться по химическому сигналу, первые образцы нанотранзисторов или электронных переключателей, состоящие из небольшого числа атомов. Специалисты складывают из атомов названия и символы своих фирм. В нанотехнологию ежегодно инвестируются сотни миллионов долларов, разработками заняты многие десятки фирм.

Нанороботы - гипотетические механизмы размером десятки и сотни нанометров (миллионные доли миллиметра), разработка которых начата не так давно. Как и роботы обычных размеров, нанороботы будут иметь самые различные конструкции и назначения: смогут двигаться, производить механические и другие операции, управляться извне или встроенными компьютерами. Они смогут собирать механизмы, создавать новые вещества; для таких устройств используют названия «ассемблер» (сборщик) или «репликатор». Возможна настройка их на переработку или уничтожение каких-либо веществ. Венцом этого направления могут стать нанороботы, самостоятельно собирающие свои копии, т.е. практически способные к размножению.

Нанороботов условно разделяют на два вида: способных конструировать что-либо, например, самовоспроизводиться (ассемблеры), или деконструировать, разбирать (дизассемблеры). Молекулярные ассемблеры - основной инструмент человека для манипуляций в наномире. Любой вирус в оп-

ределённом смысле также является ассемблером. Нанороботов нередко так и называют - "искусственные вирусы".

Микроскопические роботы, способные манипулировать объектами размером в несколько нанометров ( $10^{-9}$  метра), могли бы оказаться весьма полезны во многих отраслях народного хозяйства. Причём, вполне мирного свойства - у военных, тоже проявляющих живейший интерес к теме, задачи всё же помасштабнее, не какие-то там атомы с молекулами с места на место переставлять.

В то же время, в настоящий момент у роботов отсутствуют навыки обращения с "предметами" меньше чем несколько микрон ( $10^{-6}$  метра). Правда, не очень понятно, уместно ли для таких микроскопических устройств название "робот". Ведь они не только интеллектом обделены и к обучению неспособны, но даже батарейку в их крохотное тельце не вставишь.

Конечно, за деятельностью этих механизмов невозможно наблюдать невооружённым глазом, требуется сканирующий электронный микроскоп. Идея изобретения состоит в том, чтобы использовать микроскоп не только для наблюдения, но и для обратной связи - отдачи роботу производственных указаний. При этом электронный луч будет буквально толкать робота под руку.

Для реализации такого взаимодействия использованы свойства сплавов с эффектом памяти формы (Shape Memory Alloys - SMA), пластически деформированные изделия из которых способны при нагревании восстанавливать свои первоначальные очертания. Собственно, SMA-сплавам на титано-никелевой основе уже давно прочат переворот в нанотехнологиях, однако идея использования луча микроскопа для нагрева манипулятора запатентована только сейчас.

Как показывает практика, манипуляция объектами размером меньше микрона требует создания манипуляторов микронного размера, причём сила воздействия такого привода должна быть неувовимо мала. Существующие типы приводов (электромагнитный, пьезоэлектрический) не удовлетворяют этим параметрам.

SMA-устройства раньше не делались меньше, чем в несколько сот микрон. Следовательно, было необходимо ответить на два вопроса. Во-первых, каковы минимальные размеры, при которых сплавы сохраняют свои свойства? И, во-вторых, насколько малый объект можно выборочно нагреть, чтобы привести устройство в действие?

Преыдущие исследования показали, что плёнка из SMA на титано-никелевой основе с добавлением кремния и оксида кремния толщиной в 100 нанометров (всего около 200 атомных слоев) всё еще способна предсказуемо менять форму при нагревании.

Что же касается электронного сканирующего микроскопа, то его лучом можно нагревать области микронного диаметра. Для нагревания до необходимой температуры образца размером, например, 4 x 10 x 100 мкм необхо-



димо  $1,3 \cdot 10^{-5}$  Дж, то есть возможной мощности луча  $2 \cdot 10^{-3}$  Вт достаточно, чтобы сделать это за 6 мс.

Путём деформации достаточно толстой перфорированной плёнки из SMA и последующего нагрева лучом микроскопа удалось продемонстрировать прототип манипулятора с диаметром рабочего элемента 2 мкм и длиной 20 мкм. Проект манипулятора уже достаточно подробно описан. У позиционирующего устройства "руки" будет шесть степеней свободы. Каждая будет управляться своим "храповиком", приводимым в действие давлением инертного газа, цилиндрами будут служить углеродные нанотрубки. Всё достаточно просто, однако пока такая "рука" пока не создана.

## Литература

1.Ивашов Е.Н., Некрасов М.И., Степанчиков С.В. Расчет надежности деталей, работающих под нагрузкой. Метод. указания по выполнению РГР. – М.: МИЭМ, 1998. – 27с.

2.Расчет и конструирование механических систем оборудования аэрокосмической и электронной техники: Учебное пособие для вузов / Н.В. Василенко, Е.Н. Ивашов, С.В. Степанчиков и др. – Томск: МГП «РАСКО» при издательстве «Радио и связь», 1999. – 304с.

3.Степин В.С., Горохов В. Г., Розоб М. А. Философия науки и техники. М: Контакт Альфа, 1995. - 384с.

4.Митчем К. Что такое философия техники. / Перевод с английского под редакцией В.Г. Горохова. М: Аспект Пресс, 1995. - 149с.

5.Ананян М.А., Косяков А.Н., Киселев М.В., Котенков А.Г., Лускинович П.Н., Обьедков О.В., Шавыкин А.Е., Филипов В.В. Нанотехнологический комплекс НК-100-1В: Сборник докладов 5-й Всероссийской конференции "Нейрокомпьютеры и их применение", 1999.

6.Ананян М.А. Наноробототехника: аспекты технологии, 2003.

7.Интернет: <http://www.membrana.ru>.

8.Кудрин Б.И. Научно-технический прогресс и формирование техноценозов // ЭКО, 1980, №8, С. 15-28.

9.Кудрин Б.И. Отбор: энергетический, естественный, информационный, документальный. Общность и специфика // Электрификация металлургических предприятий Сибири. Вып. 5. Томск: ТГУ, 1981, С. 111-187.

10.Половинкин А.И. Основы инженерного творчества. - М.: Машиностроение, 1988. - 368 с.

11.Васин В.А., Ивашов Е.Н., Степанчиков С.В. Исторические и идеологические аспекты нанотехнологии // Справочник. Инженерный журнал. 2008. №7, С. 55-59.

## Содержание

Предисловие.....	3
Введение. Исторические и идеологические аспекты нанотехнологии.....	5
Глава 1. Пьезосканеры в сканирующей зондовой микроскопии.....	17
Введение.....	17
1.1. Пьезоэлектрический сканер зондового микроскопа.....	18
1.1.1 Искажение и нестабильность изображения из-за неидеальности пьезосканеров.....	20
1.1.2 Конструкция и устройство пьезосканеров.....	21
1.2. Методика диагностики сканеров.....	23
1.2.1. Основные калибровочные характеристики.....	25
1.2.2. Методика измерения на стенде и диагностика пьезосканеров.....	27
1.2.3. Процедура корректировки Z – приводов.....	31
1.2.4. Диагностика XY – сканеров.....	32
1.3. Результаты корректировки пьезосканеров зондовых микроскопов.....	34
1.4. Материалы пьезосканеров.....	36
1.5 Модель сканера точной подстройки на основе пьезоэлектрического элемента.....	40
Заключение.....	44
Литература.....	44
Глава 2. Магнитострикционные системы перемещения в нанотехнологии.....	46
Введение.....	46
2.1. Магнитострикция. Основные сведения.....	46
2.1.1 Обменная магнитострикция.....	47
2.1.2 Спонтанная магнитострикция и инвар-эффект.....	49
2.1.3 Магнитострикция (магнитодипольная и одноионная).....	50
2.1.4 Гигантская анизотропная магнитострикция.....	52
2.1.5 Материалы с гигантской магнитострикцией.....	54
2.1.6 Влияние спонтанной магнитострикции на тепловое расширение ферромагнетиков.....	55
2.1.7 Влияние всестороннего сжатия на температуру магнитного упорядочения.....	59
2.2. Информационное моделирование магнитострикционного эффекта.....	61
2.2.1. Описание магнитострикционного преобразователя.....	61
2.2.2. Информационный метод проектирования новых технических решений.....	62
2.2.3. Основные понятия информационного моделирования.....	63
2.2.4. Информационная модель магнитострикционного эффекта.....	63



2.3. Расчет магнитного поля постоянного магнита для магнитоотрицательных систем перемещений.....	67
2.4. Устройства прецизионного перемещения на основе эффекта магнитоотрицательности.....	70
2.4.1. Устройство перемещения для нанотехнологии.....	70
2.4.2. Привод для нанотехнологии.....	71
2.4.3. Устройство наноперемещения зонда.....	72
Заключение.....	73
Литература.....	73

Глава 3. Формирование объектов в лазерно-зондовой .....	74
Введение.....	74
3.1. Механическое воздействие лазерного излучения на атомы.....	74
3.2. Особенности поглощения светового излучения материалом.....	78
3.2.1. Описание и назначение сканирующего ближнепольного оптического микроскопа (СБОМ).....	82
3.2.2. Атомный нанозонд с единичным фотоном.....	84
3.2.3. Устройство для получения нанометровых отверстий на подложке.....	87
Заключение.....	89
Литература.....	89

Глава 4. Формирование магнитоориентированных объектов методом туннельно-зондовой нанотехнологии.....	91
Введение.....	91
4.1. Туннельно-зондовая нанотехнология.....	92
4.2. Магнитные свойства наноструктур.....	96
4.2.1. Нанокристаллические материалы.....	96
4.2.2. Однодоменные ферромагнитные частицы.....	98
4.2.3. Суперпарамагнетизм.....	101
4.3. Магнитоориентированные объекты.....	102
4.3.1. Применение магнитоориентированных объектов в датчиках.....	103
4.3.2. Применение магнитоориентированных объектов в магнитных запоминающих устройствах.....	104
4.3.3. Трудности, связанные с увеличением плотности записи магнитных запоминающих устройств.....	104
4.3.4. Пути увеличения плотности магнитной записи информации.....	107
4.3.5. Методы получения магнитоориентированных объектов.....	112
4.4. Варианты технических устройств для формирования магнитоориентированных объектов.....	113
4.4.1. Устройство для ориентированной сборки наноструктур.....	113
4.4.2. Многозондовое устройство для ориентированной сборки магнитных наноструктур.....	115

4.4.3. Многозондовое устройство для формирования магнитоориентированных наноструктур.....	118
Заключение.....	119
Литература.....	120

Глава 5. Формирование искусственных нейронных сетей в туннельно-зондовой нанотехнологии.....	122
Введение.....	122
5.1. Особенности искусственных нейронных сетей.....	124
5.1.1. Модель технического нейрона.....	124
5.1.2. Обобщенная аппроксимационная теорема и вычислительные возможности искусственных нейронных сетей.....	126
5.1.3. Теоретический подход к возможности ускоренного обучения искусственных нейронных сетей за счет адаптивного упрощения обучающей выборки.....	127
5.2. Модели ускоренного обучения искусственных нейронных сетей.....	132
5.2.1. Модель нейросетевой структуры для оптимизации функционирования.....	134
5.2.2. Исследование искусственных нейронных сетей полупроводниковых нанообъектах-квантовых точках и ямах.....	135
5.3. Варианты технических устройств для получения элементов искусственных нейронных сетей.....	137
5.3.1. Искусственная нейронная сеть.....	138
5.3.2. Устройство для получения углеродных пленок.....	140
5.3.3. Устройство для получения нанодорожек.....	141
5.3.4. Устройство наноперемещений.....	142
Заключение.....	143
Литература.....	144

Глава 6. Квантовые наноразмерные структуры для систем кодирования и криптографии.....	145
Введение.....	145
6.1. Основные концепции построения «защищенных систем».....	147
6.1.1. Понятие «защищенная система» - определение и свойства.....	147
6.1.2. Стандарты безопасности «защищенных систем».....	148
6.1.3. Анализ существующих стандартов информационной безопасности.....	149
6.2. Моделирование «защищенных систем».....	151
6.2.1. Формальные модели безопасности.....	151
6.2.2. Дискреционная модель Хариссона-Руззо-Ульмана.....	151
6.2.3. Типизованная матрица доступа.....	156
6.2.4. Мандатная модель Белла-ЛаПадулы.....	160
6.2.5. Моделирование квантовых наноразмерных структур для	



систем кодирования и криптографии.....	166
6.3.Варианты технических устройств для получения элементов систем кодирования и криптографии.....	168
6.3.1.Устройство для получения нанодорожек.....	168
6.3.2.Устройство формирования изомерных квантовых точек.....	169
6.3.3.Устройство долговременной памяти.....	170
Заключение.....	171
Литература.....	171

Глава 7. Формирование рельефа подложки в туннельно-зондовой нанотехнологии.....	173
Введение.....	173
7.1.Основные направления в нанотехнологии.....	173
7.2.Физические основы процесса формирования нанорельефа подложки....	177
7.2.1.Основные нанотехнологические процессы.....	177
7.2.2.Типы нанотехнологических воздействий.....	183
7.2.3.Физические эффекты в туннельно-зондовой нанотехнологии.....	185
7.3.Нанотехнологические установки и устройства.....	188
7.3.1.Нанотехнологическая установка «Луч-2».....	188
7.3.2.Устройство определения нанорельефа подложки.....	192
7.3.3.Устройство определения отклонений острия зонда.....	193
7.3.4.Устройство перемещения зонда для нанотехнологии.....	194
7.3.5.Устройство кантилевера для нанотехнологии.....	195
7.3.6.Основные параметры зондов.....	196
Литература.....	198

Глава 8. Приповерхностные свойства нанобъектов, гетероструктур и нанобластей с различным химическим и фазовым составом.....	199
Введение.....	199
8.1.Основные принципы сканирующей зондовой микроскопии.....	200
8.2.Силовое взаимодействие зонда с поверхностью.....	202
8.2.1.Электростатическое взаимодействие.....	202
8.2.2.Сила Ван-дер-Ваальса.....	208
8.2.3.Капиллярные силы.....	210
8.2.4.Адгезионные силы.....	212
8.2.5.Упругие взаимодействия зонда с образцом.....	216
8.3.Эффекты, возникающие при исследовании приповерхностных свойства .....	221
8.3.1.Эффект прилипания-скольжения.....	221
8.3.2.Трибоэмиссия частиц, электромагнитных и акустических волн....	225
8.3.3.Трибохимические эффекты.....	226
8.4.Анализ химического состава поверхности.....	227

Заключение.....	228
Литература.....	229

Глава 9. Управление качеством распознавания образов в кластерных системах обработки информации.....	230
Введение.....	230
9.1. Теоретический анализ кластерных систем.....	231
9.2. Комментарии к генетическому алгоритму.....	236
9.3. Кластеры повышенной производительности.....	238
9.4. Коммуникационные библиотеки.....	241
9.5. Стандарты MPI.....	242
9.6. Оценка достоверности.....	244
Заключение.....	245
Литература.....	245

Глава 10. Оборудование молекулярно-лучевой эпитаксии для выращивания метастабильных наноструктур.....	247
Введение.....	247
10.1. Основные сведения о молекулярно-лучевой эпитаксии.....	247
10.2. Сверхвысоковакуумные установки и комплексы для молекулярно-лучевой эпитаксии.....	249
Литература.....	256

# Глава 1. Пьезосканеры в сканирующей зондовой микроскопии

## Введение

Повышение качества и микроминиатюризация изделий во многих областях науки и техники поставили проблему сверхточного позиционирования функциональных механизмов при выполнении установочных перемещений, в том числе по нескольким координатам одновременно. Сверхмалая с точки зрения механики длительность процессов в автоматизированном технологическом и исследовательском оборудовании (от 100 мс до 0,1 мс и менее) предъявляет жесткие требования к динамическим параметрам модулей линейных перемещений. При этом во многих случаях эти модули должны работать в вакуумной среде, которая создает проблемы даже для обычных диапазонов позиционирования.

Например, при изготовлении новейших поколений современных сверхбольших интегральных схем (СБИС) элементы электронной схемы формируются на полупроводниковом кристалле (подложке) посредством ряда технологических операций, из которых наиболее ответственной является операция совмещения изображений на шаблоне и подложке. При этом точность совмещения шаблона и подложки по шести степеням подвижности (например, для систем рентгенолитографии) не должна превышать 20 нм. Время точной установки подложки в микролитографическом оборудовании не должно превышать 100 мс. Модули установочных перемещений для перспективного микролитографического оборудования должны обеспечивать высочайшие требования по быстродействию и точности позиционирования одновременно.

Зеркала сверхбольших адаптивных составных телескопов (АСТ) состоят из многих сегментов, так как выполнить их как целые конструкции невозможно. При этом каждый сегмент необходимо выставлять и выверять относительно других с погрешностью позиционирования не более 50 нм по трем координатам одновременно и поддерживать в этом состоянии с учетом всех возмущающих воздействий. Постоянная времени позиционирующего модуля (время обработки управляющего сигнала) не должна превышать 100 мс. Такой телескоп (диаметр 25 м) строится впервые в нашей стране и в мире и без создания модулей, обеспечивающих подобную стабильность положения и быстродействия, функционировать не сможет.

Сканирующие туннельные и атомно-силовые микроскопы (СТМ и АСМ) работают как при атмосфере, так и в сверхвысоком вакууме (при давлении менее  $5 \cdot 10^{-8}$  Па). Существующие функциональные модули для сканирования – пьезотрубки и пьезостолбцы могут обеспечить точностные и динамические требования (разрешение по осям X, Y, Z порядка 0,1 нм, постоянная времени около 1 мс), но имеют диапазон перемещения не более 400 мкм. Поэтому создаваемые модули для перспективных широкопольных



зондовых микроскопов должны обеспечивать достаточный диапазон сканирования и сверхжесткие требования по позиционированию и быстродействию.

Рассмотренные примеры, число которых можно продолжать, позволили сформировать ряд принципиально новых требований к линейным модулям, а именно: сочетание достаточной величины установочных перемещений с высочайшей точностью позиционирования и быстродействием.

### **1.1. Пьезоэлектрический сканер зондового микроскопа**

Основным элементом сканирующей зондовой микроскопии (СЗМ), который и обеспечивает возможность работы прибора в режимах атомных разрешений, является сканер (пьезодвигатель). Сканер изготавливают из поликристаллических пьезоэлектрических материалов, которые, с одной стороны, обеспечивают высокую жесткость конструкции, а с другой – возможность перемещения на весьма малые расстояния, вплоть до единиц пикометров.

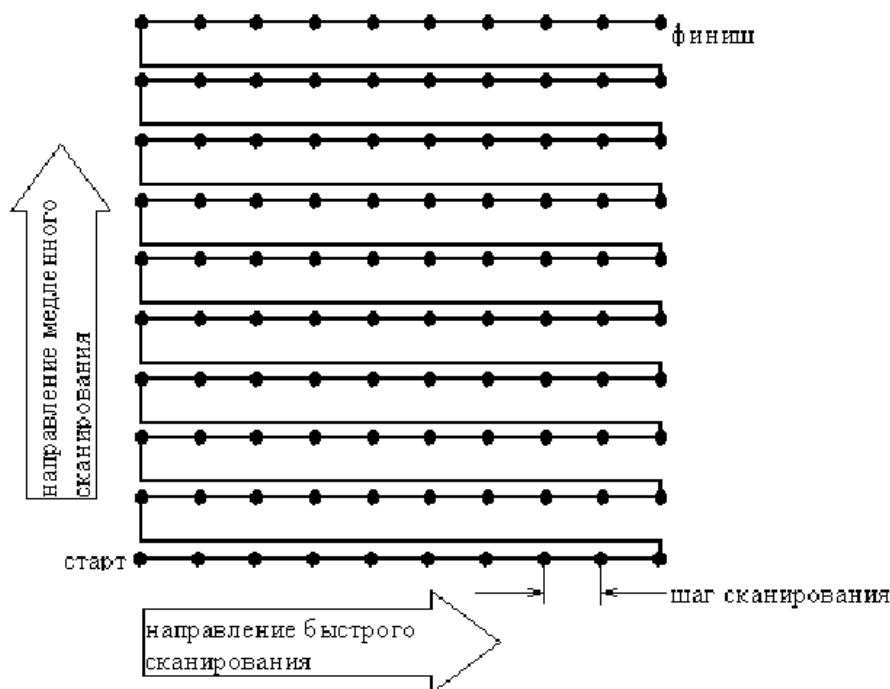
Кроме того, современное состояние СЗМ таково, что возникла необходимость перехода от чисто качественных измерений поверхностных структур к количественным, метрологически выверенным измерениям, что требует углубленного изучения поведения пьезокерамических материалов, используемых для изготовления сканеров, разработки методов прецизионных измерений и метрологической аттестации и коррекции пьезосканеров. Причем, для процедур литографии необходимы сканеры с ортогональностью в плоскости на уровне лучше десятых долей процента. Если неортогональность в плоскости еще допустимо исправлять методами программной коррекции, то неортогональность в плоскости оси Z, которая в некорректированной керамике может достигать десятков градусов, недопустима, так как приводит к тому, что низкая боковая частота сканера начинает быть существенной для работы петли обратной связи, что существенно снижает качество работы прибора.

Именно эти обстоятельства привели к необходимости углубленного изучения свойств сканера и разработки технологии их коррекции на физическом уровне.

Практически во всех сканирующих зондовых микроскопах пьезоэлектрический сканер используется как очень тонкое позиционирующее устройство, для того чтобы перемещать зонд относительно образца или образец относительно зонда. Сканер обеспечивает два независимых движения: сканирование вдоль поверхности образца (в плоскости XY) и перемещение в направлении, перпендикулярном к поверхности (по оси Z).

Сканер зондового микроскопа перемещает зонд относительно образца по типу растровой картины, как показано на рис. 1.1.

Сканер движется вдоль первой линии скана и обратно. Затем он смещается на шаг в перпендикулярном направлении на следующую линию сканирования, движется вдоль нее и обратно, затем смещается на третью линию и таким образом дальше. Путь отличается от традиционной растровой картины тем, что чередующиеся линии данных измеряемых сигналов не берутся в обратных направлениях. Измеряемые данные при сканировании собираются только в одном направлении, обычно называемом направлением быстрого сканирования, чтобы минимизировать ошибки регистрации, которые возникают из-за гистерезиса сканера. Перпендикулярное направление, в котором сканер перемещается от линии к линии, называется направлением медленного сканирования.



**Рис. 1.1.** Движение сканера СЗМ в плоскости (X, Y). Точками показаны места сбора информации

Пока сканер движется вдоль линии сканирования, данные изображения оцифровываются через одинаково расположенные интервалы. Данными является высота сканера по направлению Z для режима постоянной силы или режима постоянного тока. Для режима постоянной высоты данными являются отклонения кантилевера или туннельный ток.

Промежуток между точками данных называется шагом сканирования. Размер шага определяется полным размером сканирования и числом точек, данных на одну линию. В типичном сканирующем зондовом микроскопе размеры скана изменяются от 10 ангстрем до 100 микрон, и от 64 до 1000 точек данных на линию. Некоторые системы имеют 3000 точки данных на линию. Число линий устанавливается обычно равным числу точек на линии.

Таким образом, идеальная установка данных соответствует квадратной сетке измерений.

При работе сканеров имеют место такие нежелательные эффекты, как гистерезис, старение, крип, которые приводят к искажению получаемого изображения.

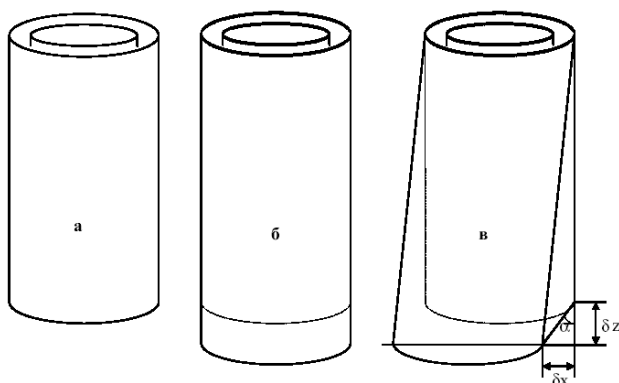
Пьезоэлектрические сканеры являются критическими элементами в сканирующих зондовых микроскопах, которые существенно влияют на качество получаемого изображения. Поэтому характеристики сканера должны удовлетворять определенным требованиям. Одним из самых важных требований, предъявляемых к сканеру, является требование ортогональности сканера.

### 1.1.1. Искажения и нестабильности изображения из-за неидеальности пьезопроводов

Сканер может состоять из одной или нескольких полых пьезотрубок. Возможны варианты, когда одна из них вставлена в другую (рис. 1.2.). В микроскопах линии СОЛБЕР в настоящее время такая конструкция пьезосканеров является основной для большинства моделей, поэтому остановимся на ней более подробно.

В двухтрубчатой коаксиальной конструкции внешняя трубка является Z приводом. При приложении управляющего напряжения к ее электродам, сканер обеспечивает движение зонда в вертикальном направлении (по оси Z). Внутренняя трубка отвечает за отклонение зонда по X и Y.

При приложении переменного напряжения к +X и -X электродам, происходит растяжение одного и сжатие противоположного сегмента трубки, заставляя сканер отклоняться назад и вперед в X направлении. Аналогично осуществляется перемещение по оси Y. Напряжение, приложенное к Z электроду, инициирует сжатие и растяжение сканера в зависимости от полярности приложенного напряжения относительно оси поляризации.



**Рис. 1.2.** Сканер и его работа: а – общий вид сканера до приложения напряжения; б – идеальная работа сканера; в – реальная работа сканера



Работа сканера, описанная выше, идеализирована, реально же возникает такой дефект, как неортогональность сканера (рис. 1.2,в). Это значит, что при приложении напряжения к Z электроду, сканер не только растягивается или сжимается по вертикали, но еще и отклоняется в сторону. Такой дефект может быть вызван несоосностью внутренней и внешней оси пьезотрубки, обусловленной неидеальностью технологического изготовления (при токарной обработке внешней трубки до толщины стенок 500 мкм проблематично ожидать точности, лучшей 50 мкм, а это составляет 10% от толщины), неоднородностями материала в самой пьезокерамике, неоднородностью поляризации, воздействием пайки и т.д.

В качестве величины, характеризующей неортогональность сканера, можно ввести такой параметр, как угол неортогональности сканера. Угол неортогональности сканера  $\alpha$  (см. рис. 1.2,в) можно определить следующим образом. Если считать, что наклон сканера происходит по координате X, то отклонение в направлении X можно обозначить как  $\delta x$ . Далее будем считать, что при приложении напряжения к Z электроду, сканер удлинился. Деформацию (удлинение) по оси Z обозначим как  $\delta z$ . Теперь рассмотрим треугольник, который можно видеть на рисунке 1.2. Это прямоугольный треугольник и легко определить  $\operatorname{tg}\alpha$ :

$$\operatorname{tg}\alpha = \frac{\delta x}{\delta z}, \text{ откуда: } \alpha = \operatorname{arctg} \frac{\delta x}{\delta z}.$$

### 1.1.2. Конструкция и устройство пьезосканеров

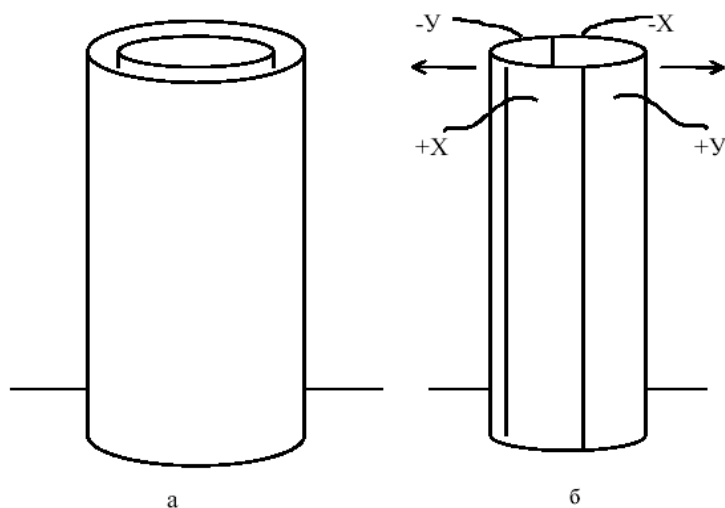
Непосредственно после изготовления пьезотрубки, дипольные моменты в ее керамике ориентированы беспорядочно. Для ориентации дипольных моментов используется процесс, называемый поляризацией. При поляризации сканеры достаточно нагреть не до температуры Кюри, а до 200°C – температуры, при которой существенно уменьшается коэрцитивная сила. Прикладывается постоянное напряжение к сканеру. После этого большинство из диполей становятся ориентированными. Затем сканер охлаждается, чтобы зафиксировать диполи в ориентированном состоянии. Поляризованный сканер может реагировать на напряжение расширением или сжатием.

Периодическое использование сканера помогает сохранить поляризацию сканера. Напряжение, приложенное для того, чтобы вызывать сканирующее движение, перестраивает неориентированные диполи, что уменьшает случайную ориентацию. Если сканер не переполаризован при регулярном использовании, то значительная часть диполей будет находиться в случайном состоянии (неполяризованном) в течение периода работы. Располяризация заметно ускоряется, если сканер подвергается влиянию температуры 200°C. Это означает, что если вы хотите добавить

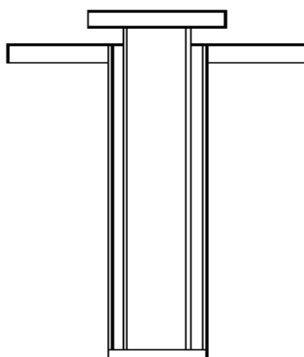
подогревательную стадию к вашему сканирующему зондовому микроскопу, вы должны термически изолировать ее от сканера.

Сканеры сканирующего зондового микроскопа, типа СОЛВЕР, имеют конструкцию, которая схематически представлена на рис. 1.3. Сканер состоит из двух полых пьезотрубок разного диаметра, вставленных одна в другую.

Верхний конец большой трубки через фланец жестко закреплен на головке микроскопа (рис. 1.4.). К нижнему, свободному концу через фланец крепится нижний конец трубки меньшего диаметра, которая вставляется внутрь большой трубки. К противоположному, свободному концу внутренней трубки крепится через фланец держатель зонда.



*Рис. 1.3. Схематическое изображение сканера: а – внешний вид сканера, б- внутренняя трубка*



*Рис. 1.4. Конструкция пьезосканера (вид в разрезе)*

Внутренняя трубка является приводом по осям X и Y. На внутреннюю и внешнюю поверхности трубки нанесены четыре управляющих электрода в виде четырех вертикальных четвертей: +X, +Y, -X и -Y (см. рис. 1.3).

Соответственно приложение управляющего напряжения к двум парам  $X$  электродов отклоняет свободный конец сканера в  $X$  направлении, а приложение напряжения к двум парам  $Y$  электродов отклоняет сканер в  $Y$  направлении.

Трубка большого диаметра является приводом по  $Z$  – координате. Ее внешняя и внутренняя поверхности покрыты управляющими электродами. Однако в отличие от  $X$ ,  $Y$  трубки внутренний и внешний электроды  $Z$  – трубки являются сплошными. При приложении управляющего напряжения к электродам, в зависимости от полярности,  $Z$  – трубка будет удлиняться или укорачиваться (по оси  $Z$ ).

При работе микроскопа напряжение, прикладываемое к  $Z$  –приводу ( $Z$  - электроду) сканера в каждой сканируемой точке поверхности, представляет собой измеряемую величину высоты рельефа поверхности в данной точке и используется для формирования массива данных изображения топографии. В некоторых случаях используются внешние датчики непосредственно для измерения высоты сканера. В вертикальном направлении сканеры сканирующих зондовых микроскопов могут различать измерения высоты от суб-ангстремного диапазона до 10 микрон.

Максимальный размер скана в плоскости  $XU$  для конкретного пьезоэлектрического сканера зависит от длины трубок сканера, диаметра трубок, толщины стенок трубок и пьезомодулей конкретной пьезоэлектрической керамики, из которой они сделаны. Обычно сканирующие зондовые микроскопы используют сканеры, которые имеют поперечный скан от 10 ангстрем до 100 микрон.

В работе сканеров наблюдаются дефекты такие, как гистерезис, старение, крип, различные отклонения, искажение изображения.

## **1.2. Методика диагностики сканеров**

Разработана методика диагностики сканеров, состоящая из двух основных частей.

- 1) Измеряются основные характеристики движения сканера в плоскости  $XU$  в радиальном и нормальном (по оси  $Z$ ) направлении при подаче на  $Z$  – привод управляющего напряжения фиксированной амплитуды.
- 2) На основе полученных зависимостей определяются основные параметры сканера, в частности параметры, характеризующие его неортогональность.

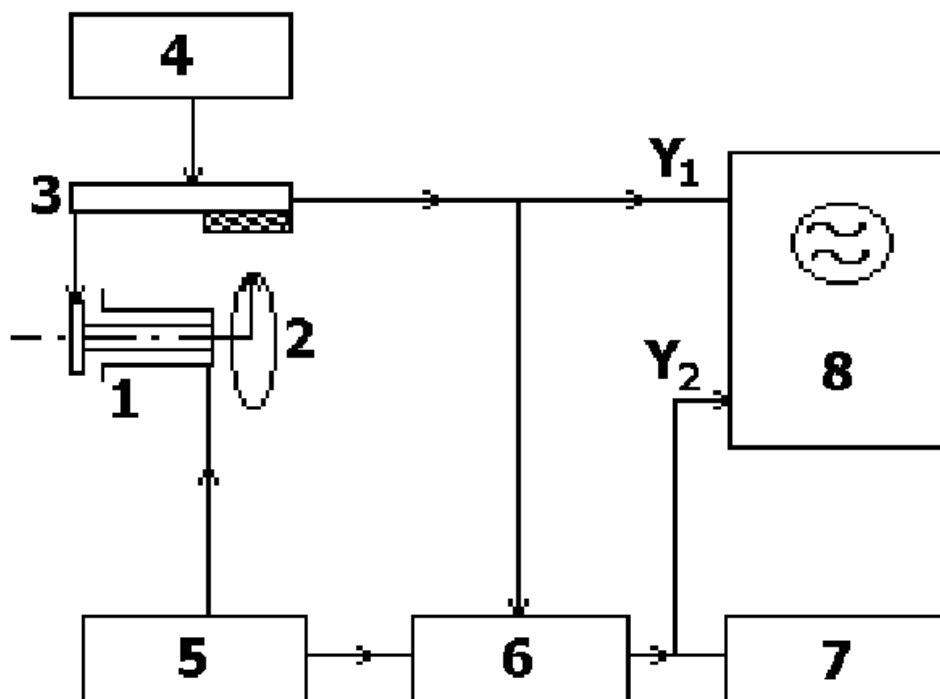
Любой реальный сканер обладает некоторой неортогональностью, и при приложении управляющего напряжения к  $Z$  – приводу сканера не только сжимается и растягивается в  $Z$  – направлении, но и колеблется в плоскости  $XU$ .

Для получения количественных характеристик параметров неортогональности сканеров разработан стенд диагностики. Стенд позволяет измерять основные характеристики сканеров, на основе которых получаются



численные значения параметра неортогональности по Z, а также параметры, характеризующие ортогональность в плоскости XY.

Необходимо установить (определить) дефект неортогональности сканера и устранить его. Определение дефекта есть первоочередная задача. Поэтому был создан стенд диагностики сканеров. Блок-схема стенда диагностики представлена на рис. 1.5.



**Рис. 1.5.** Блок-схема стенда диагностики сканеров:

- 1 – пьезосканер.
- 2 – шкала угловой позиции сканера (точность 1 угловой градус).
- 3 – датчик с зондом, чувствуюющим перемещения около 1 °А.
- 4 – блок питания датчика.
- 5 – генератор синусоидальных колебаний 20Гц, 10В<sub>эф</sub>.
- 6 – синхродетектор PAR124А, время интегрирования 3 сек.
- 7 – пятизначный цифровой вольтметр постоянного напряжения HEWLETT PACARD 3450В.
- 8 – двухлучевой осциллограф С1-83, предназначенный для установки по вертикали позиции датчика

Принцип работы стенда диагностики следующий. Основными элементами стенда диагностики X- Y и Z – приводов являются: датчик с зондом (или измерительная головка); двухлучевой осциллограф; синхродетектор (фазочувствительный усилитель PAR124А, включающий в

себя генератор синусоидальных колебаний и цифровой вольтметр постоянного напряжения).

Переменное напряжение определенной, фиксированной амплитуды (20Гц, 10В) подается с генератора на Z – привод (Z – пьезотрубку) тестируемого сканера. Это вызывает колебание трубки в направлении Z, на частоте возбуждающего напряжения. Измерительный датчик регистрирует величину отклонений. Зонд измерительного датчика измеряет угловую зависимость амплитуды колебаний. Датчик преобразовывает механические отклонения в электрический сигнал, пропорционально величине отклонения. Синхродетектор усиливает этот сигнал и преобразует его в постоянный, который подается на цифровой вольтметр. С датчика или измерительной головки сигнал также поступает на осциллограф. На осциллографе можно зрительно наблюдать сигнал.

При необходимости на стенде при изготовлении соответствующей оснастки могут исследоваться пьезопроводы различной геометрией.

### **1.2.1. Основные калибровочные характеристики**

Для получения качественных характеристик сканеров, исследуемых на стенде, необходимо определить калибровочные характеристики этого стенда. Основной калибровочной характеристикой является коэффициент, связывающий величину линейного перемещения проволочного зонда измерительной головки с величиной изменения напряжения на выходе измерительной головки.

Ниже описаны некоторые наиболее важные экспериментальные процедуры определения основной калибровочной характеристики.

Измерительная головка (датчика) регистрирует величину геометрического отклонения проволочного зонда и мембраны от некоторой позиции, условно принимаемой за нулевую. С помощью головки величина геометрического отклонения преобразуется в электрический сигнал – напряжение, находящееся в интервале  $\pm 13$  вольт. Столик с держателем сканера можно перемещать в вертикальном направлении с помощью кремальеры микроскопа, имеющей грубую и тонкую (прецизионную!) регулировки вертикального перемещения.

Последовательность операций во время калибровки следующая.

- 1) Игла датчика устанавливается в непосредственной близости от предметного столика, на котором расположен держатель сканера.
- 2) Далее с помощью ручки грубой регулировки кремальеры предметный столик устанавливается до положения касания зондовой иглой. Это состояние регистрируется по резкому отклонению луча осциллографа от первоначальной позиции. Обычная рабочая чувствительность осциллографа выбирается такой, что высота одной клетки экрана осциллографа соответствует 5 вольтам.

3) Затем с помощью прецизионного винта кремальеры на экране осциллографа достигают такого состояния, что луч на экране соответствует напряжению примерно  $-10V$  (точнее  $-10.1V$ ). Более точно эту величину считывают на цифровом вольтметре, который соединен параллельно с первым входом осциллографа. Далее производят первое считывание показаний с прецизионного микрометрического винта кремальеры. Цена деления на шкале этого винта 2 микрометра на деление.

4) Затем с помощью этого же прецизионного винта следует переместить луч осциллографа в позицию, соответствующую напряжению примерно  $+10V$  (точнее  $10.3V$ ), и снять показания со шкалы микрометрического винта.

Т.е. легко установить коэффициент между линейным приращением зонда и перемещением на выходе измерительной головки.

Перемещение луча осциллографа равно

$$S_{\text{л}} = 10.1 + 10.3 = 20.4V.$$

При перемещении луча осциллографа на  $20.4V$  прецизионный винт повернули на 3 деления. Одно деление шкалы прецизионного винта составляет 2 микрометра, следовательно, перемещение винта составляет

$$S_{\text{в}} = 6 \cdot 10000 = 60000 \overset{\circ}{A}.$$

$3 \cdot 2 = 6 \text{ мкм}$ . Известно, что 1 мкм соответствует  $10000^\circ A$ . Следовательно, перемещение винта составляет

В измерительном стенде при величине сопротивления  $100 \text{ МОм}$  в обратной связи операционного усилителя в измерительной головке, калибровочный коэффициент  $K$  датчика с зондом оказался равен

Если измерения производятся на переменном токе с использованием синхродетектора, то необходимо ввести поправочные коэффициенты,

$$K = \frac{S_{\text{в}}}{S_{\text{л}}} = \frac{60000 \overset{\circ}{A}}{20.4V} = 2940 \frac{\overset{\circ}{A}}{V}.$$

учитывающие чувствительность предусилителя. В случае использования предусилителя в режиме  $100 \text{ мВ}$  на всю шкалу (что соответствует напряжению на выходе синхродетектора в  $10V_{\text{эф}}$ ), то калибровочный

$$K_{\text{р}} = K \cdot \frac{100 \text{ мВ}}{10V} = \frac{2940 \cdot 100}{1000 \cdot 10} = 29.4 \frac{\overset{\circ}{A}}{V}.$$

коэффициент принимает окончательное значение

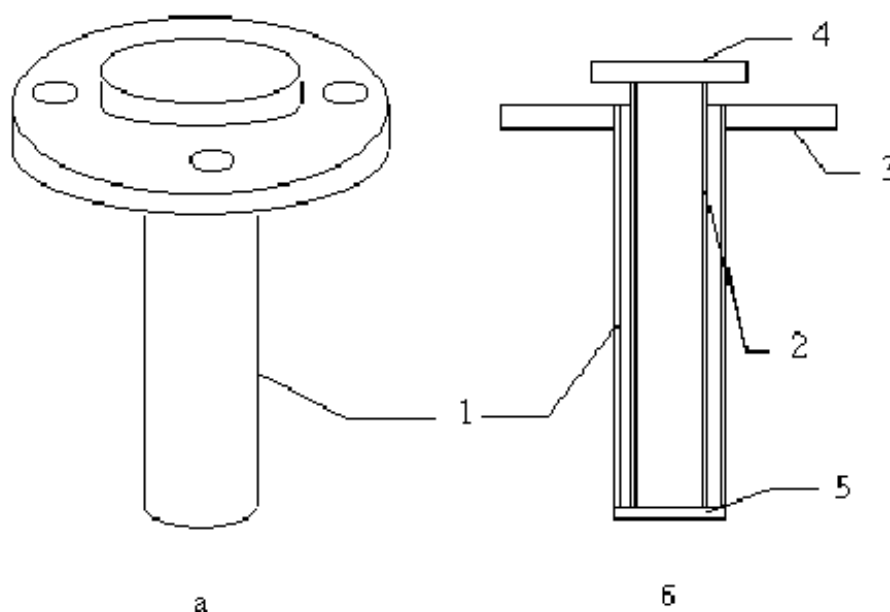
Если измерения проводятся при чувствительности предусилителя, равной  $200 \text{ мВ}$  (для торцевых измерений), то калибровочный коэффициент равен

$$k_{\text{т}} = K \cdot \frac{200 \text{ мВ}}{10V} = \frac{2940 \cdot 200}{1000 \cdot 10} = 58.8 \frac{\overset{\circ}{A}}{V}.$$

### 1.2.2. Методика измерения на стенде и диагностика пьезосканеров

- 1) К Z-приводу прикладывается управляющий сигнал определенной, фиксированной величины.
- 2) Измеряется угловая зависимость величины отклонения подвижного конца сканера в плоскости ХУ в радиальном направлении.
- 3) Измеряется угловая зависимость величины отклонения подвижного конца сканера в плоскости ХУ в направлении оси Z.

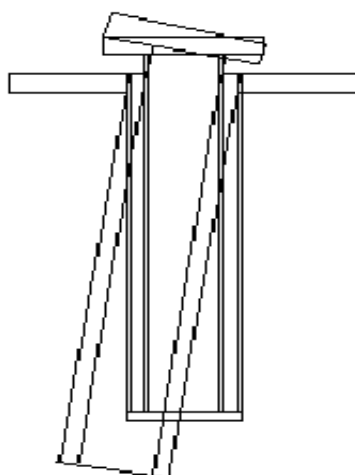
Для того чтобы описать методику измерений и диагностику пьезосканеров, необходимо подробнее рассмотреть конструкцию сканера. Конструкция сканера представлена на рис. 1.6.



**Рис. 1.6.** Конструкция сканера: а – внешний вид, б – вид в разрезе (1 – внешняя трубка сканера (Z-я трубка), 2- внутренняя трубка сканера (X-Y-я трубка), 3 – фланец для крепления сканера в микроскопе, 4 – фланец для крепления держателя образца, 5 – фланец для соединения внутренней и внешней пьезотрубок)

Внешняя трубка сканера, отвечающая за растяжение и сжатие сканера по оси Z, приклеивается к фланцу большего диаметра (3). Фланец большего диаметра жестко крепится (с помощью отверстий) непосредственно в корпусе зондового микроскопа. Таким образом, верхний конец большой трубки (трубки Z-привода) является жестко закрепленным.





*Рис. 1.7. Работа сканера и его неортогональность*

Противоположный конец большой трубки (трубки Z-привода) является свободным и к нему через фланец крепится трубка меньшего диаметра (трубка ХУ-привода). Трубка ХУ – привода вставляется внутрь большой трубки. К противоположному концу трубки ХУ – привода крепится фланец (4), к которому присоединяется держатель зонда. С противоположного конца сканера находится фланец (5), соединяющий внутреннюю и внешнюю трубки.

При приложении переменного напряжения к Z – приводу, сканер не только сжимается и растягивается в направлении Z, но и перемещается в плоскости ХУ (рис. 1.7). Обычно угол неортогональности в среднем составляет  $12^\circ$ .

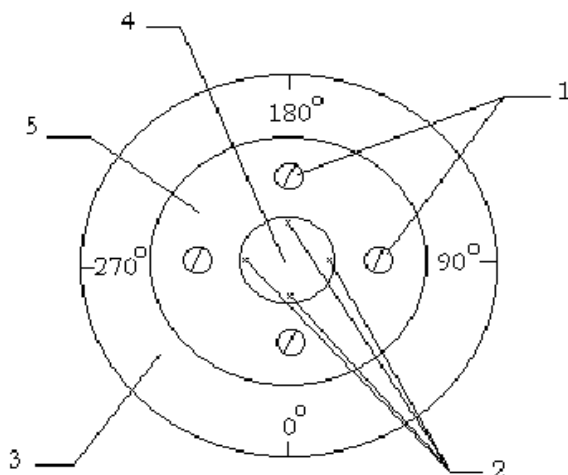
С помощью стенда диагностики можно выявить и проанализировать количественно этот дефект.

С помощью стенда можно провести торцевые и радиальные измерения. Перед началом любых измерений оборудование для измерений должно быть включено и «прогрето» в течение 30 минут. При таком режиме величина скорости дрейфов сигнала по постоянному напряжению уменьшается в несколько раз, что в конечном счете ведет к увеличению точности измерений.

Для торцевых измерений сканер устанавливается в держатель с градусной шкалой (рис. 1.8).

После того как установили сканер, необходимо проверить чувствительность предусилителя синхродетектора, необходимую для торцевых измерений. Обычно она составляет  $200\text{mV}$ . От сканера отходят 6 проводов, которые припаяны к электродам. На каждом из них имеются бирки, обозначающие, к какому электроду припаяны данные проводки. С

помощью присоединения маркированных по цвету (красные и черные) контактных зажимов типа «крокодил», которые связаны с генератором



**Рис. 1.8.** Установка сканера для торцевых измерений (вид сверху): 1 – винты для крепления сканера, 2 – точки, в которые устанавливается зонд для проведения измерений, 3- держатель с градусной шкалой, 4 – фланец для крепления держателя образца, 5 – фланец для крепления сканера в микроскопе

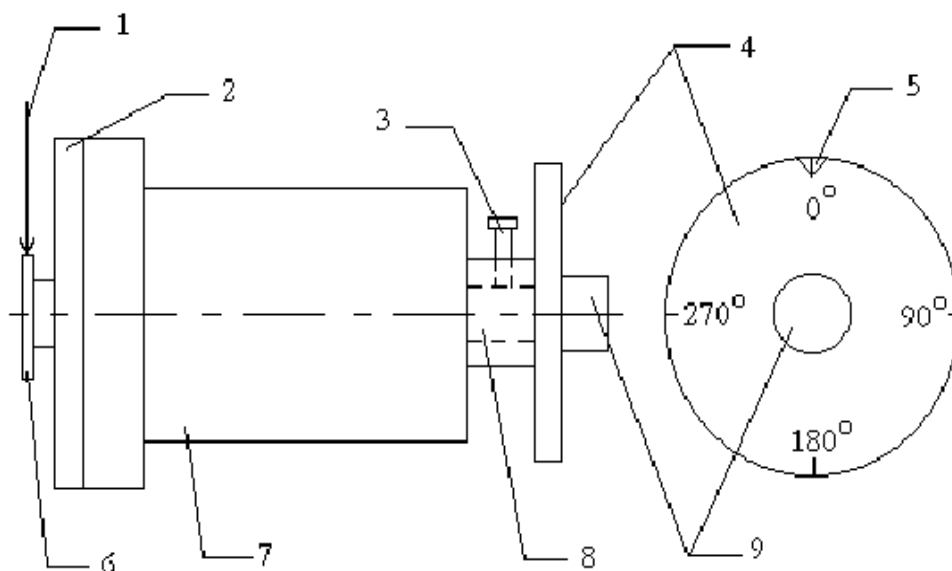
синусоидальных колебаний 20Гц,  $10V_{эф}$ , осуществляют вынужденные колебания Z – привода пьезосканера. Провод с клеммой Z соединяется с красным концом провода, идущим на выход генератора, а провод с клеммой 0 - с черным концом. После этого можно приступить непосредственно к измерениям. Для этого необходимо установить датчик с зондом таким образом, чтобы зондовая игла находилась над фланцем для крепления держателя образца. Причем, зондовая игла при соприкосновении с фланцем должна находиться примерно в одном миллиметре от края фланца.

Далее зондирующая игла опускается на фланец в положении  $0^\circ$  координатного столика (см. рис. 1.8). Затем с помощью ручки тонкой регулировки кремальеры и осциллографа добиваются такой позиции, чтобы показания датчика соответствовали  $0V$ . Затем снимается напряжение на синхродетекторе.

После этого поворачиваем держатель с градусной шкалой, а вместе с ним и сканер, на  $90^\circ$  относительно иглы зондирования. Затем игла зонда вновь опускается на фланец 4 в положении  $90^\circ$  (см. рис. 1.8) и проводится второе измерение. Аналогично проводятся измерения в оставшихся положениях ( $180^\circ$ ,  $270^\circ$  и  $360^\circ$ ).

Для проведения радиальных измерений сканер устанавливается в держатель, предназначенный для радиальных измерений (рис. 1.9). Сам держатель со сканером закрепляется на стержне 8, что связывает сканер с градусной шкалой. Так же, как и в торцевых измерениях, необходимо провод

сканера с клеммой Z соединить с красным концом провода, идущим на выход генератора, а провод с клеммой 0 - с черным концом. Далее необходимо



**Рис. 1.9.** Установка сканера для радиальных измерений: 1 – зондирующая игла; 2 – фланец для крепления сканера в микроскопе; 3 – винт, закрепляющий держатель со сканером на стержне 8; 4 – градусная шкала; 5 – стрелка, фиксирующая угловое перемещение сканера; 6 – фланец для крепления держателя образца; 7 – держатель, в котором установлен сканер; 8 – стержень для закрепления держателя 7; 9 – ручка для вращения угловой шкалы, а вместе с ней и сканера

установить датчик с зондом таким образом, чтобы зондирующая игла (см. рис. 1.9) при соприкосновении с фланцем для крепления держателя образца находилась точно по центру торца фланца 6. После этого устанавливается чувствительность предусилителя синхродетектора, необходимая при радиальных измерениях. Обычно она составляет 100мV. Затем можно переходить к непосредственным измерениям. Угловая шкала, а вместе с ней и сканер, фиксируется в таком положении, что стрелка угловой шкалы 5 фиксирует 0°. Зондирующая игла фиксируется также в положении 0°. После этого зондирующая игла опускается на боковую кромку фланца 6 в точку, соответствующую 0°. Далее сначала с помощью грубой регулировки, а затем с помощью ручки прецизионной настройки кремальеры микроскопа и осциллографа устанавливают столик в позиции, близкой к условному нулю. И далее снимается напряжение на синхродетекторе. После этого градусную шкалу 4, а вместе с ней и сканер поворачивают на 90° против часовой стрелки. В этом положении стрелка, фиксирующая угловое положение сканера 5, должна указывать на 90°. Далее зондирующая игла датчика опускается на торец фланца 6 в точку, соответствующую 90°, и проводятся

все операции, описанные при первом измерении. Аналогично проводятся измерения в оставшихся положениях ( $180^\circ$ ,  $270^\circ$  и  $360^\circ$ ).

После проведения торцевых и радиальных измерений полученные данные анализируются (проводятся необходимые расчеты).

### 1.2.3. Процедура корректировки Z приводов

Разработанная и изложенная выше методика диагностики сканера позволяет получить основные параметры, характеризующие неортогональность сканера по Z – координате. Этими параметрами являются угол неортогональности и угловое расположение максимальных и минимальных колебаний в плоскости ХУ. В результате исследования более 70 сканеров оказалось, что угол неортогональности лежит в интервале от  $3^\circ$  до  $55^\circ$ . Средний угол составляет  $12^\circ$ . Естественно, возникает вопрос, нельзя ли улучшить параметры уже изготовленного сканера при помощи некоторой процедуры корректировки.

Оказалось, что можно. В чем причина неортогональности сканера по Z – координате? Нежелательные отклонения Z – привода в плоскости ХУ вызваны следующими основными причинами: геометрическими дефектами трубки; неоднородностью толщины стенок трубок; локальными неоднородностями пьезоэлектрических свойств керамики; дефектами электродов трубки; возможными нарушениями неоднородности локальной поляризации, вызванными технологическими процессами изготовления (например; пайка проводящих электродов и т.д.).

Разработанная методика корректировки основана на следующем. Было сделано предположение, что неортогональность трубки можно уменьшить путем изменения топологии электродов, т.е. удалением определенной части площади электрода и в определенных местах электродов.

Данное предположение оказалось правильным. Действительно, используя процедуру локального удаления в нужных местах и нужной площади части управляющего электрода, удается существенно улучшить параметры неортогональности уже изготовленного сканера.

Разработанная процедура корректировки сканера производится примерно таким образом.

Используя основные параметры сканера, полученные после диагностики сканера, определяем угловое расположение минимальных колебаний в плоскости ХУ. Находим это угловое положение на самом сканере. В этом положении вблизи фланца для соединения внутренней и внешней пьезотрубок удаляем часть электрода. Удаление производится с помощью абразивного инструмента, например, алмазного надфиля или тонкой наждачной бумаги. Сначала (первая коррекция) пятно, т.е. площадь удаления электрода, имеет размер, примерно,  $3\text{мм} \times 3\text{мм}$ . После удаления такой части электрода проводят повторные радиальные и торцевые



измерения. Полученные данные после этих измерений анализируются и производится расчет. Данные, полученные после расчета анализируются, т.е. смотрим, уменьшился ли угол  $\alpha$ . Если угол неортогональности уменьшился незначительно, то производится вторая коррекция, т.е. пятно увеличивается в размере. Но это в том случае, если угловое расположение минимальных колебаний осталось прежним. Если же угловое расположение минимальных колебаний изменилось, то производится удаление электродов в этом новом положении. Далее, после вторичной коррекции, производят вновь радиальные и торцевые измерения. Затем вновь, как и ранее, данные, полученные после этих измерений, анализируются и проводится расчет. Если угол неортогональности уменьшился незначительно, то производится третья коррекция и т.д. В среднем надо провести пять коррекций. После коррекции получаем угол неортогональности, равный  $0,5^\circ$ .

#### 1.2.4. Диагностика ХУ - сканеров

Была также разработана методика диагностики ХУ – приводов сканеров. Методика основана на следующем. К У – приводу (к управляющему электроду У - привода) прикладывается тестирующее переменное напряжение фиксированной амплитуды (1В) и измеряют амплитуды колебаний сканера в направлении оси У и оси Х. Соответственно, если бы сканер был идеальным, то колебания в направлении оси Х должны были отсутствовать. В реальном сканере они обычно имеют место.

Неортогональность сканера по Х и У осям можно характеризовать «коэффициентом просачивания», который определяется как отношение амплитуды колебаний по Х и У при приложении управляющего напряжения к У – приводу. Амплитуда колебаний по оси У также является важным параметром, поскольку ее отношение к величине амплитуды приложенного напряжения является коэффициентом, который определяет размер скана.

Процедура диагностики ХУ – приводов сканера производится следующим образом.

Для диагностики Х – У приводов, сканер устанавливается в держатель, предназначенный для радиальных измерений. Необходимо провод сканера с клеммой –Х соединить с «красным» концом провода, идущим с выхода генератора, а провод с клеммой +Х - с «черным» концом. Далее необходимо установить датчик с зондом таким образом, чтобы зондирующая игла, при соприкосновении с боковой поверхностью фланца для крепления держателя образца, проходила точно по радиусу торца фланца. После этого устанавливается чувствительность предусилителя, необходимая при диагностике ХУ – приводов. Она обычно составляет 200мV. Зондирующая игла опускается на фланец, в точку, соответствующую  $0^\circ$ . Далее с помощью ручек грубой и тонкой подстройки кремальеры устанавливают взаимное расположение датчика и сканера в состояние, удобное для измерений. О

достижении такого положения судят по положению У, луча осциллографа: оно должно соответствовать напряжению 0 вольт. После этого снимается показание напряжения на синхродетекторе. Эта величина соответствует амплитуде колебаний по Х координате  $U_x$ .

Чтобы получить эту амплитуду в абсолютных единицах, ее необходимо умножить на калибровочный коэффициент:

$$k = k_T \cdot 10 = 58.8 \cdot 10 = 588 \frac{A}{V}.$$

Коэффициент  $k_T$  умножается на 10 потому, что сигнал на усилитель подается через дополнительный аттенюатор (1:10), который устанавливается перед предусилителем.

Тогда амплитуда колебаний по Х координате будет определяться по формуле:

$$A_x = k_T \cdot U_x.$$

Чтобы определить амплитуду колебаний по У координате, необходимо провод сканера с контактной клеммой  $-У$  соединить с красным концом провода, идущим на выход генератора, а провод с контактной клеммой  $+У$  с черным концом. После этого устанавливается чувствительность предусилителя, необходимая при диагностике ХУ – приводов. Обычно она составляет 200мV. Зондирующая игла опускается на боковую поверхность торца фланца, в точку, соответствующую  $90^\circ$ . И при достижении позиции, пригодной для измерения, снимают значение напряжения на синхродетекторе. Эта величина соответствует амплитуде колебаний по У координате  $U_y$ .

Амплитуда колебаний по У координате в относительных единицах определяется так же, как амплитуда колебаний по Х координате:

$$A_y = k_T \cdot U_y.$$

Необходимо также определить нежелательное «просачивание» сигнала из У – входа пьезопривода в Х – координату. Для этого необходимо провод сканера с клеммой  $-У$  соединить с красным концом провода, идущим с выхода генератора, а провод с клеммой  $+У$  - с черным концом. Зондирующая игла опускается на боковую поверхность торца фланца, в точку, соответствующую  $0^\circ$ . После достижения позиции, пригодной для измерений, производится считывание напряжения на синхродетекторе  $U$ .

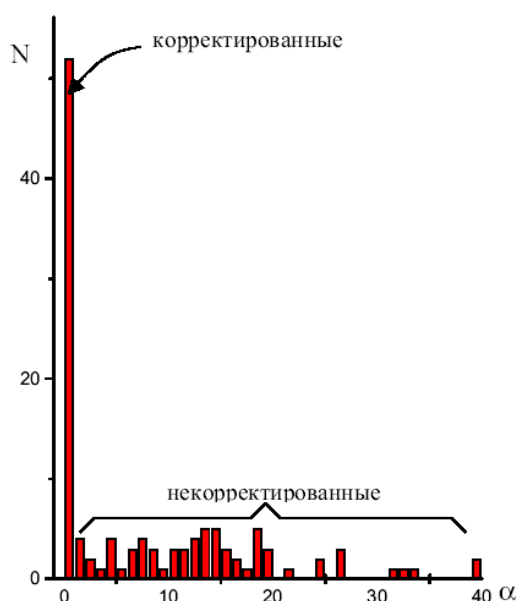
После этих манипуляций определяем «просачивание» сигнала из У-входа в Х – координату по формуле:

$$\Pi = \frac{U}{U_x} \cdot 100\%.$$

«Просачивание» сигнала из  $Y$  – входа в  $X$  – координату должно составлять не более 5 – 6%. Если «просачивание» много больше, то возникает дефект неортогональности в плоскости  $XU$ . Это значит, что задавая перемещение по  $X$ , будет происходить перемещение не только по  $X$ , но и по  $Y$ . Грубо говоря, будет наблюдаться искажение в снимаемом изображении.

### 1.3. Результаты корректировки пьезосканеров зондовых микроскопов

Было исследовано более 70 сканеров, используемых в сканирующих зондовых микроскопах типа «СОЛВЕР». Для исследованных сканеров на основе разработанной методики диагностики были получены основные параметры сканеров, до и после процедуры корректировки.

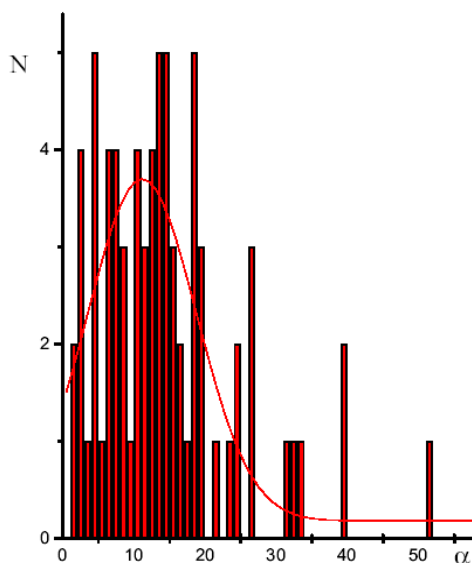


*Рис. 1.10. Гистограмма: распределение угла неортогональности для корректированных сканеров и некорректированных сканеров ( $\alpha$  – угол неортогональности сканера,  $N$  – количество сканеров)*

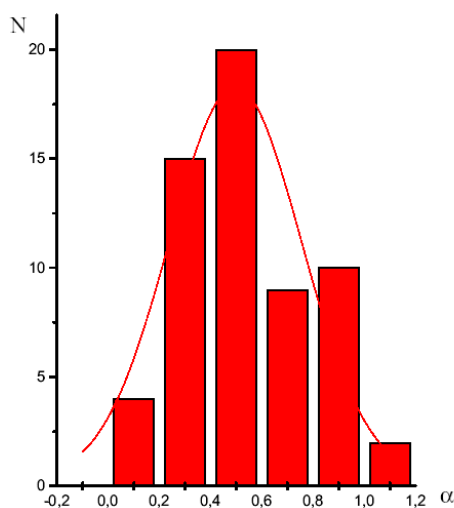
На рис. 1.10÷1.12. приведены гистограммы распределения основного параметра – угла неортогональности для сканеров до и после процедуры корректировки. На рис. 1.10. приведено распределение параметров сканера до процедуры корректировки и после процедуры корректировки.

На рис. 1.11 и 1.12 детально показаны гистограммы распределения угла неортогональности некорректированных и корректированных сканеров. Для некорректированных сканеров величина угла неортогональности по  $Z$  – координате лежит в интервале от  $3^\circ$  до  $50^\circ$ , средняя величина угла неортогональности сканера  $\alpha$  составляет  $12^\circ$ . На рис. 1.12 приведено распределение параметров сканеров после корректировки. Процедура

корректировки сканеров позволяет улучшить угол неортогональности. После корректировки средняя величина угла неортогональности от  $0^\circ$  до  $1^\circ$ .



*Рис. 1.11. Гистограмма: распределение угла неортогональности до корректировки сканеров ( $\alpha$  – угол неортогональности сканера,  $N$  – количество сканеров)*



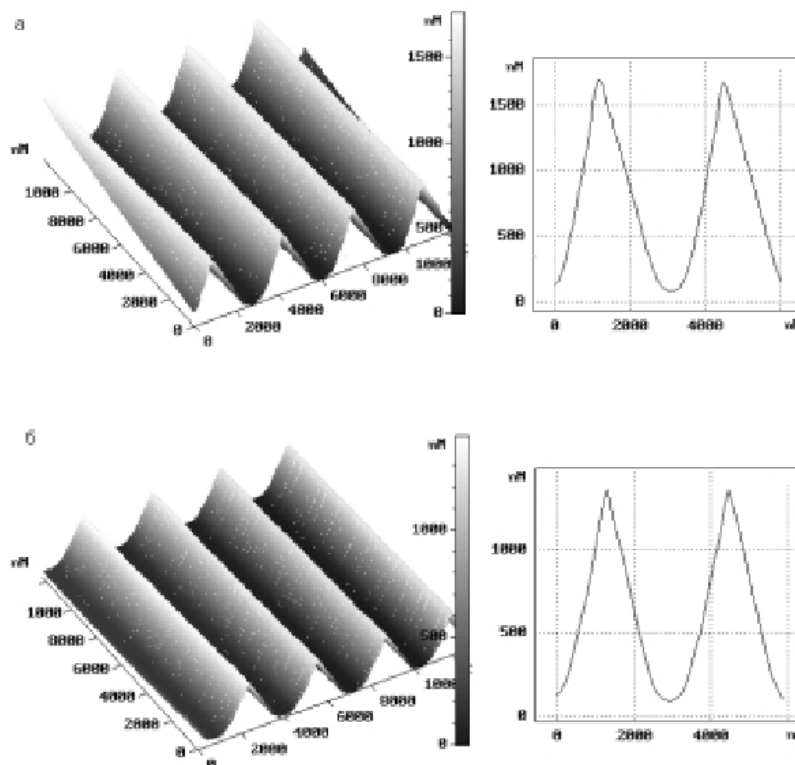
*Рис. 1.12. Гистограмма: распределение угла неортогональности после корректировки сканеров ( $\alpha$  – угол неортогональности сканера,  $N$  – количество сканеров)*

Таким образом, в результате процедуры корректировки удается уменьшить угол неортогональности сканера в среднем в 30 раз.

В качестве иллюстрации работы пьезосканера приведено изображение треугольной решетки, полученное на приборе СОЛВЕР – Р47 до



корректировки сканера и после корректировки сканера (рис. 1.13.). Если посмотреть на сечение решетки, представленное на рисунке слева, то можно увидеть, что до корректировки пьезосканера наблюдается небольшой наклон, который связан с неортогональностью сканера. Следует также отметить, что угол неортогональности сканера до корректировки составлял  $10^\circ$ . Соответственно если угол неортогональности больше, то и наклон, и несимметрия картинки относительно оси ординат будет больше. После коррекции угол неортогональности сканера составил  $0,9^\circ$ .



*Рис. 1.13. Изображение треугольной решетки: а – до корректировки сканера, б – после корректировки сканера*

#### 1.4. Материалы пьезосканеров

Пьезоэлектрические материалы условно можно разбить на две группы.

а) Пьезоэлектрические монокристаллы.

Природные пьезоэлектрические материалы имеют достаточно высокую стоимость. В связи с этим потребности бурно развивающейся электроники в настоящее время удовлетворяются синтетическими пьезоэлектрическими монокристаллами, которые выращиваются в специальных установках. Пьезоэлектрические свойства таких кристаллов с достаточно высокой повторяемостью можно задавать путем композиции входящих в него компонентов.

Выращенные кристаллы определенным образом режутся на пластины, некоторые (сегнетоэлектрики) поляризуются, и из них путем шлифования и нанесения электродов изготавливаются пьезоэлектрические элементы.

б) Пьезоэлектрическая керамика (пьезокерамика).

По физическим свойствам это поликристаллический сегнетоэлектрик, представляющий собой химическое соединение или твердый раствор (порошок) зерен (кристаллитов).

По химическому составу это сложный оксид, включающий ионы двухвалентного свинца или бария, а также ионы четырехвалентного титана или циркония. Путем изменения основного соотношения исходных материалов и введения добавок синтезируют разные составы пьезокерамики, обладающие определенными электрофизическими и пьезоэлектрическими характеристиками. Наибольшее распространение получила группа пьезокерамических материалов типа ЦТС (цирконата-титаната свинца). Вместе с тем используется керамика на основе титаната бария (ТБ) и титаната свинца (ТС). В последние годы разрабатываются новые пьезокерамические материалы со свойствами, позволяющими в некоторых случаях использовать их вместо более дорогостоящих пьезоэлектрических кристаллов. В частности, разработана и производится группа материалов на основе ниобата свинца, которая уже нашла практическое применение благодаря возможности ее использования в диапазоне частот до 30 и более МГц. Значительные исследования проводятся по созданию пьезокерамических композитных материалов, а также многослойной керамики. Зарубежные производители в зависимости от пьезоэлектрических свойств делят ее на сегнетожесткую и сегнетомягкую. В отечественной практике существует дополнительное деление на керамику средней сегнетожесткости, а также выделяются высокостабильные, высокотемпературные и т. п. материалы.

Качество пьезокерамики характеризуется следующими, принятыми за рубежом, основными параметрами:

$K_{33}^T$  ( $\epsilon_{33}^T/\epsilon_0$ ) — относительная диэлектрическая проницаемость;

$\text{tg } \delta$  — тангенс угла диэлектрических потерь при частоте 1 кГц в слабых полях;

$T_c$  ( $T_k$ ) — температура точки Кюри;

$K_p$ ,  $K_{33}$ ,  $K_{31}$ ,  $K_{15}$  — коэффициенты электромеханической связи;

$d_{33}$ ,  $d_{31}$ ,  $d_{15}$  — пьезоэлектрические модули;

$g_{33}$ ,  $g_{31}$ ,  $g_{15}$  — электрические коэффициенты по напряжению;

$N_L$ ,  $N_T$ ,  $N_R$  — частотные постоянные;

$S_{11}^E$ ,  $S_{33}^E$  — параметр эластичности;

$\rho$  — плотность;

$Q_m$  — механическая добротность.

Из природных пьезоэлектрических материалов предпочтение отдают кварцу, который при удовлетворительных пьезоэлектрических свойствах нечувствителен к влажности, химическим и температурным воздействиям, имеет высокую механическую прочность, линейность, большой динамический диапазон.

В качестве пьезоматериалов в нанотехнике обычно используют пьезокерамику состава ЦТС (цирконат-титанат свинца). Эти составы отличаются наивысшей пьезоактивностью, повышенной устойчивостью к воздействию электрических нагрузок, стабильностью свойств во времени.

Таблица 1

Основные параметры и характеристики выпускаемых пьезокерамических материалов.

Марка материала	Относительная диэлектрическая проницаемость $\epsilon_{33}^T / \epsilon_0$	Пьезомодуль, пКл/Н		Максимальная рабочая температура, °С, Траб.	Максимальное рабочее давление кг/см <sup>2</sup> , дсж
		$d_{33}$	$d_{31}$		
Группа 1					
ЦТС-19М	1900-2100	450-500	150-210	250	200
ЦТС -19	1500-1850	200	100	290	300
ЦТС-83Г	1600-1700	320-340	150-170	300	300
Группа 2					
ПКП-50	800-1000	150-180	50-60	350	300
ТС-4	500	80-100	30-40	400	1000
Группа 3					
ТВ-3	160	14-16	-	600	1500
ТНВ-1	115	8-12	-	700	1500
НТВ-1	130	24-26	-	550	1500
ТНВ-2	115	12-14	-	900	1500

На основе указанных выше пьезокерамических материалов освоено производство пьезокерамических элементов традиционных геометрических форм в виде дисков, пластин, шайб, цилиндров.

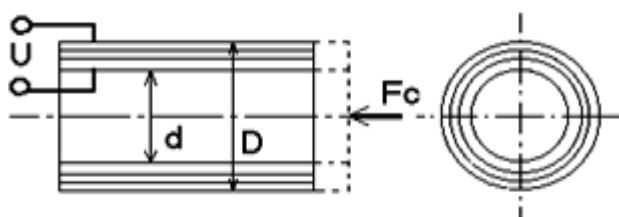


Рис. 1.14. Пьезодвигатель – трубка

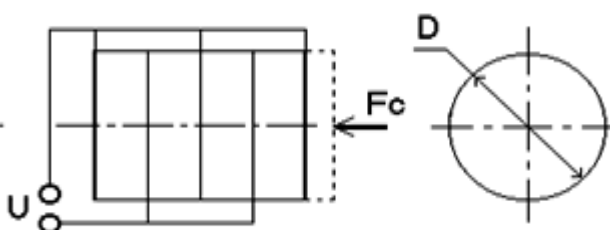


Рис. 1.15. Пьезодвигатель - столбец

При работе пьезопреобразователей используются поперечный и продольный пьезоэффекты. Рассмотрим соответственно две схемы пьезопреобразователей: трубка и столбец.

Перемещение конца пьезодвигателя-трубки (рис. 1.14) равно:

$$\xi = 2d_{31} \ell U / (D-d) - 4S_{11} \ell F_c / (\pi(D^2 - d^2)),$$

пьезодвигателя столбца (рис. 1.15):

$$\xi = nd_{33} U - 4n S_{33} h F_c / (\pi D^2),$$

где  $d_{31}$ ,  $d_{33}$  – пьезомодули, соответственно поперечного и продольного пьезоэффекта, характеризующие величину деформации образца при воздействии единичного электрического поля, м/В;

$\ell$ ,  $h$ ,  $d$ ,  $D$  – геометрические параметры трубки и столбца, м;

$n$  – количество дисков в пьезостолбце;

$U$  – напряжение питания пьезодвигателя, В;

$S_{11}$ ,  $S_{33}$  – величины, обратные модулю упругости пьезоматериала трубки и столбца;

$F_c$  – внешняя силовая нагрузка, Н.

Величина продольного пьезомодуля  $d_{33}$  для большинства пьезоматериалов более чем в два раза превышает поперечный пьезомодуль  $d_{31}$ .

Максимальное усилие  $F_{\text{сmax}}$

для пьезотрубки равно:

$$F_{\text{сmax}} = \pi d_{31} (D+d) U / (2S_{11}),$$

для пьезостолба:

$$F_{\text{сmax}} = \pi d_{33} D^2 U / (4S_{33} h).$$

Одной из особенностей любого сегнетоэлектрического материала является то, что он изменяет свои свойства в точке Кюри  $T_c$ . При температуре  $T > T_c$  кристалл не проявляет себя как сегнетоэлектрик, однако до

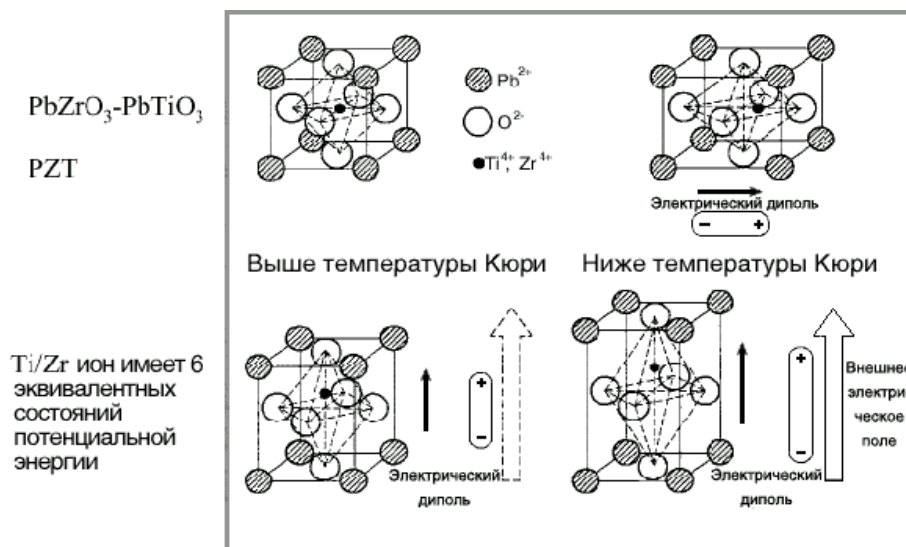


Рис 1.16. Элементарные диэлектрические решётки

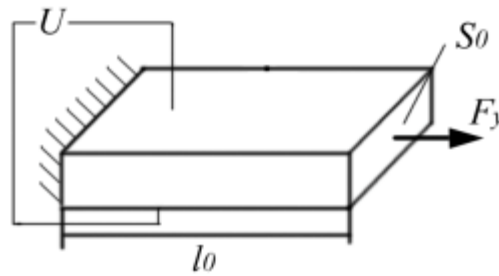
тех пор, пока  $T < T_c$ , он сегнетоэлектрик. Большинство кристаллов могут находиться в различных кристаллических фазах, которые устойчивы при различных температурах и диапазонах давлений (рис. 1.16). Переход между фазами сопровождается изменением термодинамических характеристик (упругости, оптических и тепловых свойств, объема, энтропии и т.д.). Во время перехода атомы перемещаются таким образом, что кристалл сменяет один кристаллический класс на другой. Вообще переход происходит при различных температурах при нагревании и охлаждении (температурный гистерезис). Фазовый переход первого порядка отличается сильными и резкими изменениями в кристаллической структуре. В течение перехода второго порядка изменения менее сильны, и переход непрерывен. Фазовые переходы второго порядка не обладают температурным гистерезисом. Фазовые переходы часто сопровождаются наличием новых физических явлений (сегнетоэлектричество, ферромагнетизм, сверхпроводимость). Для  $\text{BaTiO}_2$  точка Кюри равняется  $120^\circ\text{C}$ . Выше этой температуры  $\text{BaTiO}_2$  принадлежит кубическому кристаллическому классу и теряет таким образом свои сегнетоэлектрические и пьезоэлектрические свойства, ниже точки Кюри кристалл тетрагональный, от 0 и до  $-70^\circ\text{C}$  дальнейшие фазовые переходы происходят от орторомбического к тригональному кристаллическому классу. Смена кристаллической группы симметрии, связанной с фазовым переходом, становится причиной появления новых коэффициентов в тензоре материала. Материалы, с пригодным для использования коэффициентом продольной деформации, охватывают минералы, монокристаллические вещества и полимеры. Обычно пьезоэлектрический эффект наиболее ярко выражен в монокристаллических веществах. Среди материалов, годных к использованию в микросистемах, пьезоэлектрический коэффициент обычно лежит в диапазоне  $1-500 \times 10^{-12}$  м/В.

### **1.5. Модель сканера точной подстройки на основе пьезоэлектрического элемента**

Рассмотрим пьезоэлектрический элемент (рис. 1.17.) поперечного сечения  $S_0$ , закрепленный одним концом на неподвижном основании, первоначальный размер  $l_0$  которого изменяется под действием электрического поля  $E$  в направлении действия поля на значение  $\Delta$ , и перемещающий исполнительный орган массой  $m_0$ , сочлененный с его свободным концом. Кроме исполнительного органа, перемещаются также элементы объема активной части пьезосканера (ПС), масса которого  $m_k$  чаще всего меньше, чем масса исполнительного органа. При этом инерционностями перемещающихся элементарных объемов ПС можно пренебречь либо, имея в виду, что ПС закреплен одним концом и эта его часть не принимает участия в движении, уточнить значение суммарной



приведенной подвижной массы  $m_{\Sigma}$  включив в нее кроме массы ИО  $m_0$  часть массы ПС  $m_k$ .



**Рис. 1.17.** Пьезоэлектрический элемент

Основное динамическое усилие в рассматриваемом случае действует на ПС также в направлении приложенного поля, поэтому для описания пьезоэффектов можно воспользоваться уравнением

$$P = -d_n \sigma + \epsilon_0 \chi E, \quad (1)$$

где  $P$  – поляризованность пьезоэлемента,  $\sigma$  – механическое напряжение в пьезоэлементе,  $\epsilon_0 = 8,85 \cdot 10^{-12}$  Ф/м – электрическая постоянная,  $\chi$  – диэлектрическая восприимчивость пьезоматериала,  $E$  – напряженность поля.

Из уравнения обратного пьезоэффекта получим

$$\Delta Y S_0 / l_0 = d_n Y S_0 E - S_0 \sigma, \quad (2)$$

где  $\Delta$  – деформация пьезоэлемента,  $Y$  – модуль упругости,  $S_0$  – поперечное сечение пьезоэлемента,  $l_0$  – первоначальная длина пьезоэлемента.

Выражение в левой части (2) описывает усилие упругой деформации  $F_y$ , первое слагаемое в правой части (2) — усилие в ПС, вызванное приложенным электрическим полем,  $F_{\Xi}$ . Механическое напряжение, входящее во второе слагаемое правой части (3.2), определяется статическим усилием  $F_c$ , приложенным к ИО и действующим на ПС со стороны ИО, а также комплексом усилий, возникающих при движении ИО: динамическим усилием  $F_{дин}$ , пропорциональным ускорению ИО, и демпфирующим усилием  $F_d$ , пропорциональным скорости смещения ИО и вызванным процессами диссипации энергии в ПС. Описанная детализация позволяет записать (2) в виде следующего условия равновесия сил, приложенных к подвижной границе активной части ПС:

$$F_y = F_{\Xi} + F_c + F_d + F_{дин}, \quad (3)$$

где  $F_y = K_y \Delta$  – усилие упругой деформации ПС, Н;  $F_{\Xi} = d_n Y S_0 E$  – усилие в ПС, вызванное приложенным электрическим полем, Н;  $F_c$  – статическое усилие, приложенное к ИО, Н;  $F_d = -K_d d\Delta/dt$  – демпфирующее усилие в ПС, Н;  $F_{дин} = -m_{\Sigma} d^2 \Delta / dt^2$  – динамическое усилие в ПС, Н;  $K_y = Y S_0 / l_0$  – коэффициент упругости ПС, Н/м;  $K_d$  — коэффициент внутреннего демпфирования ПС, кг/с;  $m_{\Sigma}$  — эквивалентная приведенная суммарная масса подвижных частей, кг.

Воспользуемся также выражением для тока смещения в диэлектрике  $I_{см}$ . Учитывая, что для пьезоэлектрических материалов  $\chi \gg 1$ , запишем

$$I_{см} = S_0 dp/dt, \quad (4)$$

Введем  $\sigma$  из (2) в уравнение прямого пьезоэффекта (1) и продифференцируем полученную зависимость по времени, полагая модуль Юнга и диэлектрическую восприимчивость не зависящими от напряженности электрического поля и от механической напряженности в ПС. Подставив результат дифференцирования в (4), получим

$$I_{см} = C_0 dU/dt + K_n V, \quad (5)$$

где  $C_0 = \epsilon_0 \lambda S_0 (1 + K_{эм}^2) / l_0$  - емкость ПС, Ф;  $K_{п} = K_y d_{п}$  — коэффициент прямого пьезоэффекта, К/м;  $U = El_0$  — напряжение, приложенное к электродам ПС, В;  $V = d\Delta/dt$  - скорость исполнительного органа, м/с;  $K_{эм}^2 = d_{п}^2 Y / (\epsilon_0 \chi)$  — коэффициент электромеханической связи.

Установим связь между усилием  $F_э$  и электрическим напряжением, приложенным к электродам ПС:

$$F_э = K_0 U, \quad (6)$$

где  $K_0 = K_y d_{п}$  - коэффициент обратного пьезоэффекта, Н/В.

Отметим, что коэффициенты обратного и прямого пьезоэффектов равны между собой.

Примем, что заряд ПС осуществляется от источника ЭДС  $E_{п}$  через эквивалентное внутреннее сопротивление последнего  $R_{вт}$ . Для такой цепи справедливо уравнение:

$$(E_{п} - U) / R_{вт} = I_{см}. \quad (7)$$

После преобразований выражений (5) - (7) получим

$$dF_э/dt = ((K_0 e_n - F_э) / C_0 R_{э2}) - (K_0 K_n / C_0), \quad (8)$$

Уравнение (8) в сочетании с двумя нижеприведенными (9) и (10), которые могут быть выведены из выражений для составляющих усилия  $F_B$ :

$$dV/dt = (F_э + F_c - K_y \Delta - K_d) / m_{\Sigma} \quad (9)$$

$$d\Delta/dt = V \quad (10)$$

образуют полную систему дифференциальных уравнений, описывающих ПС с ИО при сделанных выше допущениях. На основании уравнений (8)-(10) составлена структурная схема. Сворачивание структуры приводит к следующей передаточной функции ПС, управляемого от источника ЭДС:

$$W_{\Delta, e_n}^э(p) = \Delta(p) / e_{п}(p) = \frac{K_0 / K_y}{1 + \left( R_{вт} C_0 + \frac{K_d}{K_y} + R_{вт} K_{п} \frac{K_0}{K_y} \right) p + \left( \frac{m_{\Sigma}}{K_y} + R_{вт} C_0 \frac{K_d}{K_y} \right) p^2 + R_{вт} C_0 \frac{m_{\Sigma}}{K_y} p^3}, \quad (11)$$

Выражение (11) показывает, что в ПС кроме демпфирования, обусловленного внутренними диссипативными процессами, имеет место также демпфирование, связанное с прямым пьезоэффектом.

Полученное выше описание ПС (8) – (11) пригодно не только для простейшего случая однородного плоского элемента. В области частот от нуля до первого механического резонанса оно удовлетворительно описывает процессы в ПС всех типовых конструкций, обеспечивающих повышенные перемещения, – наборных столбцах, кольцеобразных, биморфных и мембранных ПС. Наилучшее совпадение динамических характеристик модели и объекта получается при использовании для моделирования экспериментально определенных параметров  $m_{\Sigma}$ ,  $K_y$ ,  $K_0$ ,  $K_{\Pi}$ ,  $C_0$ ,  $R_{BT}$ .

Явление прямого пьезоэффекта затрудняет формирование управляющего поля, особенно при реализации быстропротекающих динамических процессов. Усилители на мощных транзисторах имеют выходное сопротивление порядка десятков Ом. Это позволяет при анализе динамики ПС малой емкости (биморфные элементы) пренебрегать малой электростатической постоянной времени  $T_{\Sigma} = R_{BT}/C_0$ . При управлении ПС большой емкости для компенсации явлений, связанных с прямым пьезоэффектом, необходимо применять специальные технические меры, такие, как организация специальных быстродействующих контуров регулирования управляющего поля, включение в схему усилителя дополнительных диодов, образующих цепи замыкания наведенных токов. Пренебрежение в (11) внутренней обратной связью по скорости ( $K_{\Pi} \approx 0$ ), влиянием внутреннего демпфирования ( $K_D \approx 0$ ) и инерционностью преобразователя ( $T_{\Sigma} \approx 0$ ), приводит к передаточной функции консервативного звена:

$$W_{\Delta \neq 0}^{\Sigma}(p) = \frac{\Delta(p)}{e_{\Pi}(p)} \approx \frac{K_0 / K_y}{1 + \frac{m_{\Sigma}}{K_y} p^2} \quad (12)$$

По частоте  $\omega_M = (K_y/m_{\Sigma})^{1/2}$  и амплитудам первых колебаний реально наблюдаемая затухающая переходная функция, соответствующая передаточной функции (11), мало отличается от незатухающей переходной функции, отвечающей выражению (12). Консервативное звено (12) может рассматриваться как модель самого грубого первого приближения для многих ПС.

## Заключение

В нанотехнологии перемещение и позиционирование связано с малыми размерами. В современном технологическом и исследовательском оборудовании используются различные типы устройств прецизионных перемещений. Но мы выделили устройства на пьезопреобразователях, отличительной чертой которых является их безынерционность.

Для пьезопреобразователей характерны получение микроперемещений с погрешностью не более  $\pm 0,1$  нм, высокая чувствительность и плавность перемещений, высокое быстродействие – постоянная времени пьезосканера не более 0,2 мс, возможность реверсивных перемещений. При работе пьезосканер не нагревается и тем самым не влияет на точность перемещения. Однако диапазон перемещения пьезопреобразователей не превышает 400 мкм, а использование редуцирующих механизмов приводит к уменьшению жесткости и снижению передаваемого усилия.

## Литература

1. Михайлов В.П. Автоматизация управления процессами сверхточного позиционирования линейных модулей реологического типа. – Дисс. на соиск. уч. ст. д.т.н. – М.: МГТУ им Н.Э. Баумана, 2003 – 291 с., ил.
2. Расчет и конструкция деталей аппаратуры САУ. Учебник для техникумов/ В.П. Савостьянов, Г.А. Филатова, В.В. Филатов – М.: Машиностроение, 1982. – 328 с., ил.
3. Быков В.А. Приборы и методы сканирующей зондовой микроскопии для исследования и модификации поверхностей. – Дисс. на соиск. уч. ст. д.т.н., - М.: ГНЦ НИИ физических проблем имени Ф.В. Лукина, 2000 – 393 с., ил.
4. Исаакян В.А. Микроскопия вчера, сегодня, завтра. – Материал для конференции, Москва.
5. Научно - образовательный центр физики твердотельных наноструктур ННГУ. Исследование топографии поверхности твердых тел методом атомно-силовой микроскопии в неконтактном режиме. Описание лабораторной работы. – Н. Новгород: НГУ им. Н.И. Лобачевского, 2003.
6. Неволин В.К. Физические основы туннельно-зондовой нанотехнологии. Учебное пособие. – М.: МИЭТ. 2000.
7. Никишин В. И., Лускинович П.Н. Нанотехнология и наноэлектроника // Электронная промышленность. – 1991.
8. Сушхара К., Мори И., Тобзе Т., Ито Т., Табста М., Синодзаки Т. Пьезоэлектрический столик с двумя поступательными и одной

вращательной степенями свободы для субмикронных литографических систем. // ПНИ. – 1989.

9.Бухараев А.А., Овчинников Д.В., Бухараева А.А. Диагностика поверхности с помощью сканирующей силовой микроскопии. // Заводская лаборатория. – 1997.

10.Снайдер К., Лазанн А. Растровый туннельный микроскоп на основе концентрических трубчатых пьезоэлементов. // ПНИ. – 1988.

11.Володин А.П., Панич А.Е. Применение пьезокерамических материалов ПТЭ в низкотемпературных сканирующих туннельных микроскопах. / ПТЭ. – 1989.

12.Володин А.Н. Новое в сканирующей зондовой микроскопии. // ПТЭ, №6. – 1998.

13.Быков В.А. Микромеханика для сканирующей зондовой микроскопии и нанотехнологии. Микросистемная техника. – 2000.

14.Смажевская Е.Г., Фельдман Н.Б. Пьезоэлектрическая керамика. М.: Советское радио, 200 стр. 1971.

15.Привод для нанотехнологии. Патент РФ на ПМ №40538. Опубликовано 10.09.04, Б.И. №25. Ивашов Е.Н., Павлов А.Ю., Пискарев Д.А., Степанов М.В.

16.Стереометрическое устройство для наноперемещений. Патент РФ на ПМ №36924. Опубликовано 27.03.2004 Б.И. №9. Ивашов Е.Н., Павлов А.Ю.

17.Устройство перемещения для нанотехнологии. Патент РФ на ПМ №4052. Опубликовано 10.09.04, Б.И. №25. Ивашов Е.Н., Павлов А.Ю., Пискарев Д.А., Степанов М.В.

18.Пьезопривод. Патент РФ на изобретение №2086417. Опубликовано 10.08.97, Б.И. №22. Григорьев И.Ю., Ивашов Е.Н., Степанчиков С.В. и др.

19.Вакуумный пьезопривод. Свидетельство РФ об официальной регистрации программы для ЭВМ №2001611099. Зарегистрировано в Реестре 29.08.04. Земсков О.В., Ивашов Е.Н., Степанчиков С.В.

## Глава 2. Магнитострикционные системы перемещения в нанотехнологии

### Введение

В настоящее время ни одно исследование в области физики поверхности и тонкопленочных технологий не обходится без применения методов сканирующей зондовой микроскопии. Развитие сканирующей зондовой микроскопии (СЗМ) послужило также основой для развития новых методов в нанотехнологии – технологии создания структур с нанометровыми размерами.

Одной из важных технических проблем в сканирующей зондовой микроскопии является необходимость прецизионного перемещения зонда и образца.

Для решения этой проблемы применяются различные типы устройств перемещения, осуществляющих перемещение объектов с высокой точностью.

Широкое применение получили различные механические редукторы, в которых грубому перемещению исходного движителя соответствует тонкое перемещение смещаемого объекта. Механические редукторы в основном применяются для грубой подводки зонда к подложке.

Для точного позиционирования в СЗМ применяют сканирующие элементы. В настоящее время наибольшее распространение получили сканирующие элементы на основе пьезоэлектриков.

Перспективным является применение магнитострикционных материалов для создания прецизионных систем перемещения, так как это позволяет повысить точность позиционирования в десятки раз по сравнению с пьезоэлектрическими приводами.

### 2.1. Магнитострикция. Основные сведения

Магнитострикция (от лат. натяжение, сжатие) — изменение формы и размеров тела при его намагничивании. Это явление свойственно как сильно магнитным (ферромагнитным), так и парамагнитным и диамагнитным веществам. Магнитострикция — результат проявления взаимодействий в магнитных телах. Магнитострикция неизменно привлекает внимание не только физиков, но также и инженеров с точки зрения конструирования новых приборов и технических устройств.

Магнитострикция оценивается безразмерной величиной — относительным изменением размеров магнетика  $\lambda = dl/l$ . Здесь  $dl$  — удлинение (или укорочение) при включении магнитного поля  $H$ , а  $l$  — длина образца. В экспериментах обычно измеряется  $\lambda_{11}$  — продольная магнитострикция, когда напряженность поля  $H$  совпадает с направлением измерения,  $\lambda_{\perp}$  —



поперечная магнитострикция, когда указанные направления взаимно перпендикулярны. Величины  $\lambda_{11}$  и  $\lambda_{\perp}$  малы (даже для ферромагнетиков), и для их измерения применяются специальные методы и установки.

Различают два вида магнитострикции: изотропную (обменную) и анизотропную (магнитодипольную и одноионную). Ниже в упрощенной и качественной форме объясняются механизмы их возникновения.

### 2.1.1. Обменная магнитострикция

Этот вид магнитострикции возникает в результате изменения обменного взаимодействия между магнитными моментами атомов  $M_{ат}$  в кристаллической решетке. Магнетизм атома обусловлен электронами (ядро атома дает очень малый вклад в магнетизм атома, и им обычно пренебрегают). Электроны атома участвуют в создании  $M_{ат}$  двояко. Во-первых, каждый электрон, вращаясь вокруг ядра, образует микроскопический замкнутый ток, величина его равна произведению микроскопического тока на площадь орбиты электрона. Этот магнитный момент называется орбитальным  $M_{орб}$  и изображается в виде вектора, направленного перпендикулярно площади орбиты. Во-вторых, каждый электрон обладает своеобразным "собственным" магнитным моментом (согласно выводам квантовой механики). Его называют спиновым  $M_{сп}$  (от англ. spin — вращение).

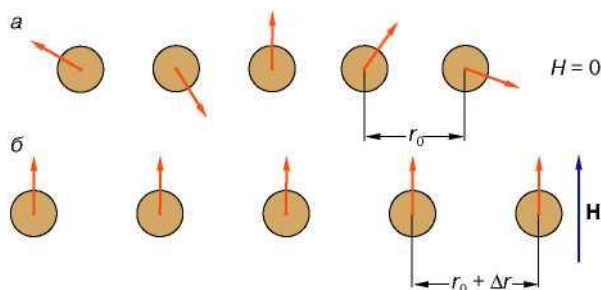
Векторное сложение  $M_{орб}$  и  $M_{сп}$  дает  $M_{ат}$ . Следует отметить, что внутри атома  $M_{орб}$  и  $M_{сп}$  связаны магнитными силами (спин-орбитальным взаимодействием).

В кристаллах ферромагнетиков, как было показано в 30-е годы русским теоретиком Я.И. Френкелем и немецким теоретиком В. Гейзенбергом, между электронами соседних магнитных атомов возникает особый вид взаимодействия, который они назвали обменным. Это электростатическое взаимодействие, однако оно не простое (кулоновское), а квантовое. В механизме обменного взаимодействия электронов важная роль отводится направлению спинов соседних атомов. Обменным его назвали потому, что в процессе данного взаимодействия электроны соседних магнитных атомов как бы обмениваются своими местами. Результатом обменного взаимодействия электронов является то, что моменты  $M_{сп}$  электронов устанавливаются параллельно друг другу, возникает спонтанная или самопроизвольная намагниченность (то есть без участия внешнего поля  $H$ ). Поскольку  $M_{сп}$  и  $M_{орб}$  взаимосвязаны, то можно говорить, что спонтанная намагниченность  $I_s$  создается упорядочиванием магнитных моментов  $M_{ат}$  (намагниченность  $I_s$  — это число однонаправленных  $M_{ат}$  в  $1 \text{ см}^3$  ферромагнетика).

Спонтанная намагниченность обращается в нуль при температуре  $T_C$ , называемой точкой Кюри. Величину обменного взаимодействия можно оценить по величине  $T_C$ . При этой температуре тепловое движение разрушает

упорядоченное расположение моментов  $M_{ат}$ , созданное обменным взаимодействием. Отсюда следует, что чем больше обменное взаимодействие в ферромагнетике, тем выше должна быть температура  $T_C$  для разрушения магнитного порядка.

Возникновение обменной магнитострикции можно наглядно пояснить следующим образом. Представим себе, что имеется цепочка магнитных атомов при температурах, близких к  $T_C$ . В области  $T_C$  большинство моментов  $M_{ат}$  находятся в разупорядоченном состоянии, так как действие теплового движения на атомы больше, чем обменное взаимодействие. Пусть  $r_0$  — равновесное расстояние между атомами, соответствующее этому состоянию ферромагнетика (рис. 2.1,а). Включим теперь поле  $H$ . Моменты  $M_{ат}$  повернутся по полю (рис. 2.1,б), но это приведет к изменению обменной энергии (поскольку, согласно теории, данная энергия зависит от направления спинов взаимодействующих электронов, принадлежащих соседним атомам).



**Рис. 2.1.** К объяснению явления обменной магнитострикции

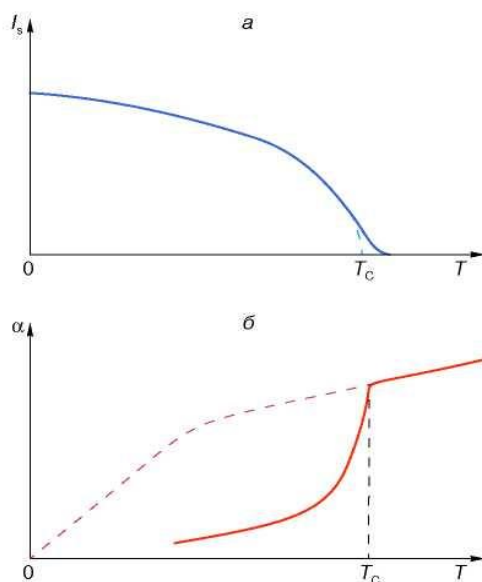
Состоянию ферромагнетика на рис. 2.1,б будет соответствовать другое равновесное расстояние между атомами:  $r_0 + \Delta r$ , где  $\Delta r$  есть не что иное, как обменная магнитострикция. В ферромагнетиках, обладающих кубической симметрией, величина  $\Delta r$  не зависит от направления в кристалле, следовательно, обменная магнитострикция будет изотропной. Это означает, что в кубическом кристалле величина  $\Delta r$  будет одной и той же во всех направлениях последнего. Эта магнитострикция будет проявляться в изменении объема кристалла  $\Delta V/V$ , при этом в большинстве ферромагнетиков она положительна, то есть при включении внешнего поля  $H$  объем образца увеличивается.

Подобного рода магнитострикция сопутствует процессу намагничивания ферромагнетика, при котором под влиянием  $H$  происходит ориентация моментов  $M_{ат}$ . Процесс напоминает намагничивание парамагнетиков, поэтому он получил название парапроцесса. Парапроцесс особенно интенсивен в области точки Кюри, и обменная магнитострикция здесь достигает наибольшей величины.

В ферромагнетиках, обладающих гексагональной структурой, например в редкоземельном металле гадолинии Gd, парапроцесс и обменная магнитострикция обладают анизотропией.

## 2.1.2. Спонтанная магнитострикция и инвар-эффект

Магнитострикция, обусловленная изменением обменного взаимодействия, проявляется не только при приложении магнитного поля  $H$ , но также при изменении температуры ферромагнетика (при отсутствии  $H$ ). Это тепловая магнитострикция (иногда называемая термострикцией) особенно велика в области точки Кюри. В самом деле, из вида температурной зависимости спонтанной намагниченности  $I_s$ , представленной на рис. 2.2,а, следует, что число разупорядоченных моментов  $M_{ат}$  особенно бурно возрастает при приближении к  $T_C$ . Это приводит к некоторому изменению обменной энергии, что, в свою очередь, вызывает обменную магнитострикцию  $(\Delta V/V)_T$ , однако в противоположность действию парапроцесса отрицательную (так как она сопутствует разупорядочиванию моментов  $M_{ат}$ ).



**Рис. 2.2.** К объяснению возникновения инвар-эффекта:  
*а* - температурная зависимость спонтанной намагниченности инварного сплава (схематически);  
*б* - температурная зависимость коэффициента теплового расширения инварного сплава (штриховая линия - температурный ход *а* при отсутствии влияния спонтанной магнитострикции)

У некоторых ферромагнетиков эффект спонтанной магнитострикции оказывает существенное влияние на тепловое расширение, так как приводит к частичной компенсации последнего. На рис. 2.2,б штриховой линией схематически показан температурный ход коэффициента теплового расширения  $\alpha = 1/l(\Delta l/\Delta T)$  ферромагнитного сплава 36%Ni + + 64%Fe при отсутствии компенсирующего действия обменной магнитострикции, сплошная кривая — зависимость  $\alpha(T)$ , экспериментально наблюдаемая.

Видно, что в определенном интервале температур  $\alpha$  может приобретать очень низкие значения.

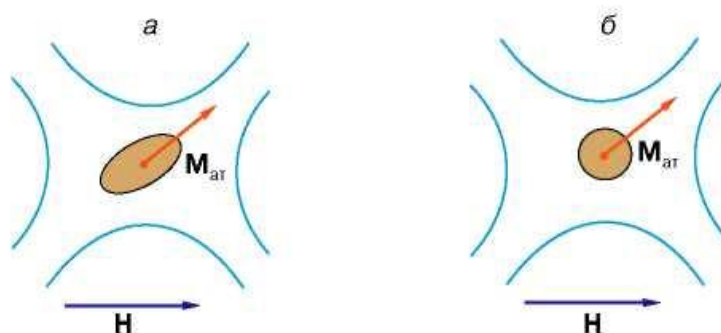
Указанный выше сплав носит название инвара (не изменяющего свои размеры при нагреве). В настоящее время существует большое число сплавов типа инвар; природа их малого коэффициента теплового расширения магнитная. Явление компенсации коэффициента теплового расширения спонтанной магнитострикцией получило название инвар-эффекта. В гадолинии инвар-эффект анизотропен, то есть различен по разным осям гексагонального кристалла.

### 2.1.3. Магнитострикция (магнитодипольная и одноионная)

Кроме рассмотренной выше обменной магнитострикции в ферромагнетиках при приложении поля  $H$  возникает анизотропная магнитострикция. Она сопутствует процессам намагничивания в полях более слабых, чем те, в которых проявляется парапроцесс. Анизотропия ее состоит в том, что  $\lambda$  по различным осям кристалла имеют разные величины и знаки. Характерная черта анизотропной магнитострикции состоит в том, что при ней меняется форма образца (при ничтожно малом изменении объема).

В теории рассматриваются два механизма анизотропной магнитострикции: 1) магнитодипольный и 2) одноионный. В первом из них рассчитывается магнитное взаимодействие магнитных моментов  $M_{ат}$ , расположенных в узлах кристаллической решетки, при этом магнитные моменты  $M_{ат}$  уподобляются магнитным диполям (то есть маленьким магнетикам с северным и южным полюсами).

Магнитодипольное взаимодействие в кристаллах кубической симметрии вдоль ребра и диагоналей куба будет различным, следовательно, равновесные расстояния между магнитными атомами в этих направлениях будут также различными, то есть магнитострикции будут разными по величине в этих направлениях. Однако данный механизм дает малый вклад в анизотропию магнитострикции ферромагнетиков.



**Рис. 2.3.** К объяснению одноионного механизма возникновения анизотропной магнитострикции: а - случай Tb и Dy, б - случай Gd

Как показали исследования, главным для анизотропной магнитострикции является одноионный механизм. Определяющую роль в нем играет наличие у магнитного атома или иона (то есть заряженного атома) орбитального магнитного момента  $M_{орб}$ . Согласно теории, в этом случае электронное орбитальное облако приобретает несферическую (анизотропную) конфигурацию (на рис. 2.3,а оно условно изображено в виде эллипсоида). Наглядно механизм возникновения анизотропной магнитострикции можно представить следующим образом. Пусть анизотропный магнитный ион находится в кристаллической решетке в окружении других ионов, создающих электростатическое поле (оно обычно называется кристаллическим).

На рис. 2.3 условно показаны тонкими линиями кристаллические поля, создаваемые окружающими ионами, отражающими локальную симметрию кристалла. При приложении поля  $H$  магнитный момент иона  $M_{ат} = M_{сп} + M_{орб}$  ориентируется в направлении  $H$  и одновременно с ним поворачивается анизотропное электронное облако иона, которое возмущает электростатическое поле окружающих ионов. В результате кристаллическая решетка испытывает анизотропные деформации в соответствии с симметрией кристалла. Эти деформации есть не что иное, как анизотропная магнитострикция.

Подобного вида анизотропная магнитострикция очень велика в редкоземельных металлах (не всех), так как их ионы обладают большими величинами орбитальных моментов  $M_{орб}$ .

Редкоземельный металл гадолиний не имеет орбитального момента ( $M_{орб} = 0$ ) и его атомный момент  $M_{ат}$  содержит только спиновую составляющую ( $M_{ат} = M_{сп}$ ). Вследствие этого электронное облако его иона имеет сферическую форму. Как следует из рис. 2.3,б, при приложении поля  $H$  поворот электронного облака не приводит к возмущению кристаллического поля окружающих ионов, следовательно, здесь не работает механизм одно-ионной магнитострикции. В Gd и его соединениях наблюдаемая анизотропная магнитострикция, по-видимому, обязана магнитодипольному механизму.

Ситуация с анизотропной магнитострикцией в металлах Fe, Ni, их сплавах и ферритах близка к случаю Gd. Намагничивание в них происходит в основном за счет спиновых моментов и в небольшой степени за счет орбитальных моментов. В этих магнетиках кристаллическое поле так сильно воздействует на  $M_{орб}$ , что они как бы закрепляются в решетке и теряют способность вращаться в направлении магнитного поля. Это явление принято называть замораживанием орбитального момента. Однако в некоторых из этих соединений замораживание  $M_{орб}$  происходит не полностью. Поэтому в данных веществах возникает анизотропная магнитострикция одноионной природы, но много меньшая по величине, чем в редкоземельных магнетиках.

## 2.1.4. Гигантская анизотропная магнитострикция

Было установлено, что анизотропная магнитострикция редкоземельных металлов тербия Tb, диспрозия Dy, их сплавов и ферритов-гранатов при низких температурах превышает анизотропную магнитострикцию железа, кобальта, никеля и их сплавов в десятки, сотни и даже тысячи раз (в монокристаллах). Несколько позднее такая гигантская магнитострикция была обнаружена в так называемых интерметаллических соединениях TbFe<sub>2</sub> и DyFe<sub>2</sub>. В них эта магнитострикция реализуется не только при низких температурах, но и выше комнатных температур.

В табл. 2.1 приведены результаты измерения продольной магнитострикции насыщения  $(\lambda_s)_{11}$  (то есть в полях  $H = H_s$ ) при температуре жидкого азота 78 К в поликристаллах Tb и Dy (для сравнения даны соответствующие значения  $(\lambda_s)_{11}$  для поликристаллов Fe, Co и Ni)

Таблица 2.1

Металл	$(\lambda_s)_{11} \cdot 10^6$
Tb	1230
Dy	1400
Fe	-15
Co	-85
Ni	-40

Огромных величин достигает  $\lambda_s$  в гексагональных кристаллах Tb и Dy (табл. 2.2). При этом она особенно велика вдоль гексагональной оси с и несколько меньше вдоль а. (На рис. 2.4,а показаны направления этих осей.) Видно, что эти значения в сотни раз больше, чем в кристалле Ni. (На рис. 2.4,б показаны направления осей кубического кристалла Ni символами [111], [100] и [110].)

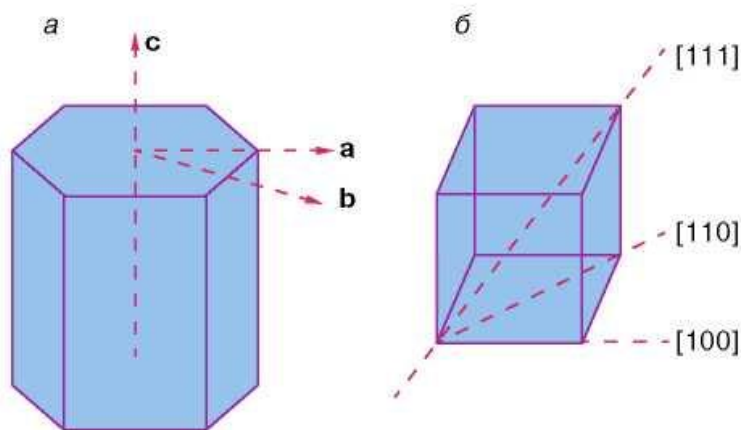
Магнитострикция в кристаллах Tb, Dy, Gd и Ni .

Таблица 2.2

Металл	T, К	$(\lambda_s)_{11} \cdot 10^6$	Кристаллическая ось
Tb	4,2	5 460	H    a
	4,2	22 000	H    c
Dy	4,2	8 500	H    a
	4,2	21 000	H    c
Gd	78	-80	H    c
	78	-16	H    a
Ni	78	-60	H    [111]

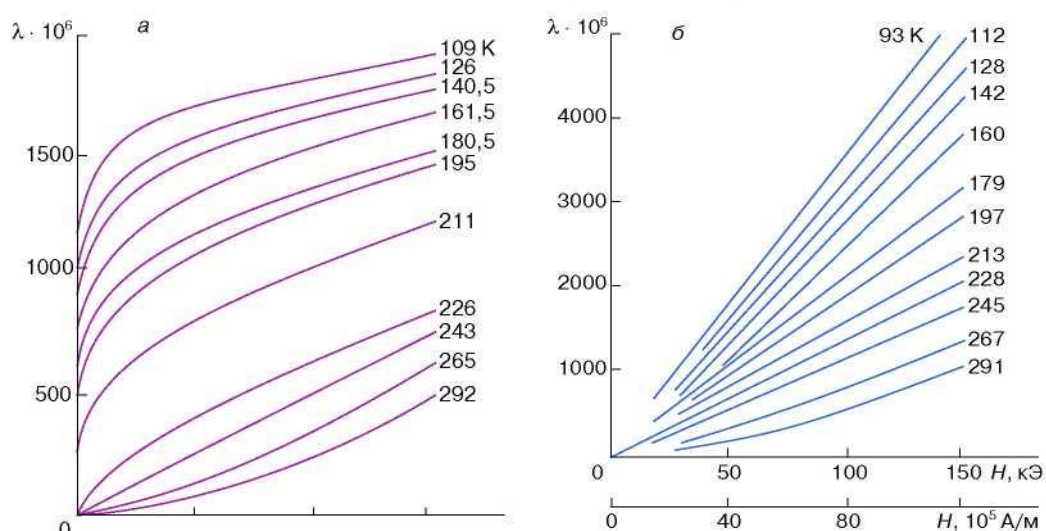
На рис. 2.5 представлены кривые магнитострикции  $\lambda(H)$  гексагонального монокристалла Tb вдоль его оси с и в базисной плоскости при разных температурах.





**Рис. 2.4.** Направления осей: а - в гексагональном кристалле и б - в кубическом кристалле

В поле, приложенном в базисной плоскости, магнитострикция сравнительно быстро стремится к насыщению (рис. 2.5,а), тогда как в направлении гексагональной оси с тенденция к насыщению не обнаруживается даже в очень сильном поле ( $H = 150$  кЭ).



**Рис. 2.5.** Зависимость магнитострикции в кристалле тербия от напряженности магнитного поля  $H$ : а - базисная плоскость ( $H \perp c$ ); б - гексагональная ось ( $H \parallel c$ )

Анализ приведенных результатов измерений показал, что за гигантскую магнитострикцию в Tb и Dy ответственны два механизма: одноионная магнитострикция и обменная магнитострикция (возникающая вдоль оси с, так как в этом направлении обменное взаимодействие сильно зависит от межатомного расстояния).

Было обнаружено, что ряд ферритов-гранатов  $R_3Fe_5O_{12}$  ( $R = Tb, Dy, Ho$ ) в области гелиевых и азотных температур обладает гигантской магнитострикцией (порядка  $10^{-3}$ ).

В табл.2.3 приведены результаты измерений ее в кристаллах этих веществ. Видно, что она носит также анизотропный характер в направлениях

[100] и [111] и возникает, как и в металлах Tb и Dy, за счет механизма одноионной анизотропии. Дело в том, что в ионах Tb, Dy (и других редкоземельных ионах) ответственная за магнетизм 4f-электронная оболочка находится в "глубине" иона; выше располагаются другие оболочки, которые экранируют ее от действия внутрискристаллического поля, в результате чего орбитальный момент  $M_{орб}$  не замораживается этим полем.

Таблица 2.3

Феррит-гранат	$\lambda_{[100]} \cdot 10^5$		$\lambda_{[111]} \cdot 10^5$	
	4,2 К	78 К	4,2 К	78 К
Tb <sub>3</sub> Fe <sub>5</sub> O <sub>12</sub>	1270	67	2460	560
Dy <sub>3</sub> Fe <sub>5</sub> O <sub>12</sub>	-1400	-254	-550	-145
Ho <sub>3</sub> Fe <sub>5</sub> O <sub>12</sub>	-665	-82	-632	-56

Необычайно высокая анизотропная магнитострикция наблюдалась в актинидных магнетиках (урановых, нептуниевых и др.). Так, в соединениях US при T= 4,2 К величина  $\lambda_{[111]} = 7000 \cdot 10^{-6}$ , а в NpFe<sub>2</sub>  $\lambda_{[111]} = -8000 \cdot 10^{-6}$ . В этих соединениях, так же как и в редкоземельных магнетиках, ответственны за возникновение гигантской анизотропной магнитострикции орбитальные магнитные моменты 5f-оболочек.

### 2.1.5. Материалы с гигантской магнитострикцией

Магнитострикция в кристалле создает магнитоупругую (добавочную) анизотропию, которая оказывает сильное влияние на ход кривой намагничивания. Металлы Tb, Dy, Ho, Er и ферриты-гранаты этих металлов (например, Tb<sub>3</sub>Fe<sub>5</sub>O<sub>12</sub>) при низких температурах имеют необычайно высокие эффекты магнитострикции, на два — три порядка больше, чем магнитострикция в металлах, сплавах и ферритах группы Fe (при 100 К для Tb  $\lambda = 5,3 \cdot 10^{-3}$ , для Dy  $\lambda = 8,0 \cdot 10^{-3}$ , для Ni при той же температуре  $\lambda = 4,0 \cdot 10^{-5}$ ). Установлено, что такие гигантские магнитострикции могут быть получены и при комнатных температурах путем использования ферромагнитных соединений: DyFe<sub>2</sub>, HoFe<sub>2</sub>, DyFe<sub>3</sub> и др. Точки Кюри этих соединений, поскольку в них входят атомы Fe, выше комнатной температуры. При этом здесь, как и в случае редкоземельных материалов высокого магнитного насыщения, необходимо принять меры к снижению "вредного" влияния огромной магнитной анизотропии, т.е. уменьшить величину поля насыщения  $H_s$ .

Редкоземельные магнитные материалы обладают уникальными свойствами, которые нельзя получить в материалах, синтезируемых на основе группы Fe. Использование огромных величин магнитной анизотропии и магнитострикции в редкоземельных веществах и нахождение

путей управления ими, бесспорно, будут иметь большое значение для получения магнитных материалов с заданными свойствами.

### 2.1.6. Влияние спонтанной магнитострикции на тепловое расширение ферромагнетиков

В ферромагнетиках ситуация становится иной при температурах ниже температуры магнитного упорядочения  $T_C$ . Здесь возникает корреляция в ориентации спинов на дальних расстояниях, вследствие чего появляется кооперативный эффект — деформация образца вследствие спонтанной магнитострикции, в которой суммируются локальные деформации вблизи атомов за счет ориентации их спинов. Благодаря этому кооперативному эффекту сумма энергий взаимодействий отдельных атомов — внутренняя энергия образца будет несколько по-другому зависеть от средних межатомных расстояний при  $T < T_C$ , чем при  $T > T_C$ . Этот эффект можно учесть рассматривая энергетические вклады в полную энергию твердого тела. В первую очередь необходимо учитывать энергию обменного взаимодействия, которая при охлаждении ниже  $T_C$  возрастает пропорционально квадрату спонтанной намагниченности. Согласно теории молекулярного поля эта зависимость может быть описана формулой:

$$E_0 = -AI_s^2, \quad (1)$$

где  $A$  — интеграл обменного взаимодействия,  $I_s$  — спонтанная намагниченность.

Так как  $A$  зависит от межатомных расстояний и атомного объема, то следует учесть эту зависимость. Наиболее просто сделать это для атомного объема. В этом случае магнитоупругая обменная энергия

$$E_{0y} = -\gamma I_s^2 \omega, \quad (2)$$

где  $\omega = \Delta V/V_0$  — относительное изменение объема тела, а  $\gamma = \Delta A/\Delta \omega$  — крутизна обменного интеграла, которая показывает, как резко  $A$  изменяется с атомным объемом.

При отклонении от положения равновесия атомов изменяется также и упругая энергия:

$$E_y = +\frac{k\omega^2}{2}, \quad (3)$$

где  $k$  — объемный упругий модуль, характеризующий жесткость кристаллической решетки. Легко показать, что минимальное значение полной энергии достигается при относительном изменении объема  $\omega_s$ :

$$\omega_s = -\frac{\gamma I_s^2}{k}. \quad (4)$$

Величина  $\omega_s$  называется спонтанной объемной магнитострикцией. Она налагается на обычное тепловое расширение  $\Delta V/V(T)$  при  $T < T_C$  (кривая 2 на

рис. 1.6). В результате кривая теплового расширения существенно трансформируется (кривая 3 на рис. 2.6). Для  $\mathcal{V} > 0$  получим  $\omega_s < 0$  и  $\omega_s > 0$  для  $\mathcal{V} < 0$ . Отклонения кривой теплового расширения возможны как в сторону увеличения коэффициента теплового расширения, так и его уменьшения.

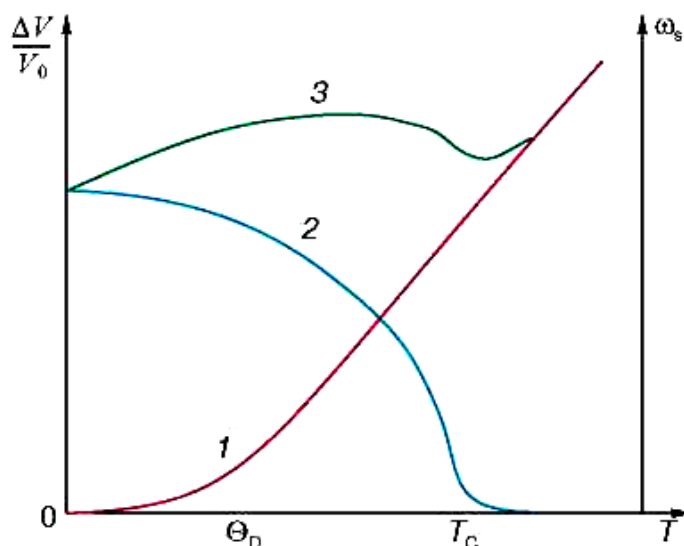
Коэффициент объемного теплового расширения может быть рассчитан по кривой 3 (рис. 2.6) по формуле:

$$\alpha_V = \frac{1}{V_0} \frac{V}{\Delta T}. \quad (5)$$

Из кривой 3 (рис. 2.6) и формулы (5) следует, что  $\alpha$  может принимать в ферромагнетике не только положительное значение, но может быть отрицательным, а также иметь нулевые значения.

Из соотношений (2) и (4) следует, что величина спонтанной магнитострикции содержит важную информацию о величине и знаке крутизны обменного интеграла  $\mathcal{V}$ . В анизотропных магнетиках изменение  $\Lambda$  при изменении межатомных расстояний вдоль различных кристаллографических направлений сильно отличается как по величине, так и по знаку. В этом случае аномалии теплового расширения при  $T < T_C$  вдоль различных кристаллографических осей неодинаковы, а спонтанная магнитострикция  $\lambda_c$  неодинакова вдоль различных кристаллографических направлений, то есть анизотропна. В ряде ферромагнетиков при  $T < T_C$  существует температурная область с нулевым тепловым расширением ( $\alpha = 0$ ). Такие температурные интервалы можно выделить на кривой 3 (см. рис. 2.6). Описанные выше аномалии теплового расширения наблюдаются не только в ферромагнетиках, но и в других магнитоупорядоченных веществах, как в кристаллических, так и в аморфных. Они получили название инварных эффектов по названию инварных Fe—Ni-сплавов, где впервые наблюдалось нулевое тепловое расширение при  $T < T_C$ .

Спонтанная магнитострикция и инварные эффекты наблюдаются не только в ферромагнетиках, но также и в других магнитоупорядоченных веществах (ферри-магнетиках, антиферромагнетиках, спиральных магнетиках и др.). В последние годы были обнаружены гигантские аномалии теплового расширения при  $T < T_C$  с относительным изменением линейных размеров за счет спонтанной магнитострикции  $\Delta l/l = 10^{-2} - 10^{-3}$  в урановых и редкоземельных магнетиках, сплавах Fe—Rh и некоторых аморфных магнетиках.



*Рис. 2.6. Тепловое расширение ферромагнетика. Температурные зависимости изменения относительного объема  $\Delta V/V_0$  парамагнетика (1), спонтанной объемной магнитострикции  $\omega_s$  ферромагнетика (2) и относительного изменения объема  $\Delta V/V_0$  ферромагнетика (3) выше и ниже температуры Кюри  $T_C$*

Спонтанная магнитострикция налагается на тепловое расширение, обусловленное тепловыми колебаниями атомов. Это проявляется в аномальном изменении размеров магнитоупорядоченного вещества (ферромагнетика, антиферромагнетика и других магнетиков) при нагреве и охлаждении по сравнению с неупорядоченными магнитными средами. Существует и другой родственный магнитоупругий эффект — изменение размеров магнитоупорядоченного вещества при действии магнитного поля. Этот эффект называется индуцированной полем магнитострикцией или просто магнитострикцией. Она является следствием спонтанной магнитострикции, однако на ее величину влияют и другие факторы. Рассмотрим их более подробно.

Спонтанная магнитострикция является результатом локальных деформаций, обусловленных магнитоупругим взаимодействием магнитных электронов с окружающими атомами. Его величина определяется зависимостью от межатомных расстояний обменных взаимодействий и взаимодействий электронных магнитных слоев ионов с локальными электрическими полями.

Рассмотрим более детально магнитострикцию в ферро- и антиферромагнетиках на основе редкоземельных элементов. В них магнетизм связан с магнитным моментом 4f-подоболочки иона редкой земли (РЗ). Если она имеет орбитальный момент  $L \neq 0$  и соответственно асимметричное распределение электронной плотности, то локальная магнитострикционная деформация (вокруг РЗ-иона) будет велика (из-за анизотропного магнитоупругого взаимодействия 4f-подоболочки с окружающими атомами).

Суммирование этих деформаций при  $T < T_C$  дает результирующую спонтанную магнитострикцию образца. Следовательно, чем больше ионов редких земель с  $L \neq 0$  в магнетике, тем она больше. В общем случае она анизотропна и неодинакова при изменении угла, который образует направление магнитострикционного растяжения с намагниченностью домена  $I_s$ . Домен — это область однородного намагничивания. Известно, что реальные ферро- и антиферромагнетики имеют многодоменную структуру. Следовательно, внутри каждого домена будет существовать анизотропная магнитострикционная деформация — спонтанная магнитострикция.

При включении поля происходит рост тех доменов, у которых векторы спонтанной намагниченности  $I_s$  ориентированы наиболее энергетически выгодно по отношению к магнитному полю  $H$ . Поскольку магнитная энергия  $E_M = -I_s H \cos \alpha$ , то это означает рост доменов с векторами  $I_s$ , образующими наименьший угол  $\alpha$  с полем  $H$ .

Величина спонтанной магнитострикции не изменяется, если направление вектора  $I_s$  изменить на обратное. В ферромагнетике, где существует одна ось легкого намагничивания (ОЛН), имеются только два типа доменов с векторами  $I_s$ , направленными вдоль ОЛН. В одном домене  $I_s \uparrow \uparrow$  ОЛН, в другом  $I_s \uparrow \downarrow$  ОЛН. Поэтому в поле  $H \uparrow \uparrow$  ОЛН процессы смещения не приводят к заметной индуцированной полем магнитострикции.

Иная ситуация возникает в ферромагнетике с несколькими ОЛН, ориентированными под определенными углами (например,  $90^\circ$  в кубической или  $60^\circ$  в гексагональной кристаллической решетке). В этом случае при включении поля  $H$  вдоль одной из ОЛН процессы смещения доменов с анизотропной магнитострикцией приводят к индуцированной магнитострикции, сравнимой по величине со спонтанной магнитострикцией.

Весьма существенным для величины магнитострикции является также высокая степень ориентации магнитных моментов ионов внутри доменов. Для выполнения этого условия необходимо, чтобы энергия взаимодействия магнитных моментов ионов  $\mu$  с эффективным обменным полем  $H_e$  превышала тепловую энергию:

$$\mu_0 H_e \gg k_B T, \quad (6)$$

где  $\mu_0$  — величина магнитного момента иона;  $k_B$  — постоянная Больцмана.

Итак, для наблюдения высоких значений магнитострикции в 4f-магнетиках необходимы:

высокая концентрация РЗ-ионов с  $L \neq 0$ ;

наличие многих ОЛН;

большие обменные поля, действующие на ионы РЗ.

Всем перечисленным условиям удовлетворяют в области температур  $T < 200$  К редкоземельные ферро- и антиферромагнитные металлы (тербий, диспрозий и др.). В них наблюдаются гигантские значения спонтанной и индуцированной магнитострикции ( $\Delta l/l \sim 10^{-2} - 10^{-3}$ ). Однако в области комнатных температур величина  $H_e$  в них недостаточно велика, чтобы

преодолеть дезориентирующее действие теплового движения (температуры Кюри и Нееля не превышают  $T = 300$  К).

Величина эффективного обменного поля, действующего на РЗ-ионы, достигает более высоких значений в интерметаллических соединениях, содержащих кроме РЗ ионы 3d-металлов (Fe, Co). В этом случае "подмагниченные" 3d-электроны (коллективизированные 3d-электроны) "переносят" обменное взаимодействие от 3d-ионов к 4f-ионам. Такое косвенное обменное взаимодействие между 3d- и 4f- ионами приводит к сильному эффективному полю, действующему на РЗ-ионы,  $H_e \sim 3 \cdot 10^6$  Э. Это поле достаточно, чтобы ориентировать магнитные моменты РЗ-ионов даже при комнатной температуре, в результате чего в некоторых из этих соединений с кубической кристаллической решеткой спонтанная магнитострикция  $\lambda_c \sim 10^{-2}-10^{-3}$  существует вплоть до комнатных температур. В магнитном поле в этих материалах также индуцируется гигантская магнитострикция  $\lambda \sim 10^{-3}$  при  $T \sim 300$  К. К таким соединениям относится ферромагнетик TbFe<sub>2</sub>. Для снижения величины внешнего магнитного поля, необходимого для получения максимального эффекта, разработаны композиции типа (Tb, Dy)Fe<sub>2</sub>, где за счет сплавления РЗ с разными знаками констант магнитной анизотропии гигантская магнитострикция достигается в сравнительно небольших магнитных полях ( $\sim 10^2$  Э).

### 2.1.7. Влияние всестороннего сжатия на температуру магнитного упорядочения

К числу магнитоупругих эффектов следует также отнести смещение температур магнитного упорядочения при действии давления. Согласно теории молекулярного поля, температура Кюри ферромагнетика с делокализованными атомными магнитными моментами 3d-ионов определяется соотношением:

$$T_c = \frac{2}{3} z S(S+1) \frac{A}{k_B}, \quad (7)$$

где  $A$  — интеграл обменного взаимодействия между локализованными спинами  $S$ ,  $z$  — число ближайших атомов. Из формулы (1.7) следует, что изменение  $T_c$  с давлением:

$$\frac{\Delta T_c}{\Delta P} = \frac{2}{3 k_B} z S(S+1) \frac{\Delta A}{\Delta P}. \quad (8)$$



Поскольку изменение относительного объема тела  $\Delta \omega = -\kappa/P$ , где  $\kappa$  — сжимаемость, то, определив смещение  $T_C$  с давлением, можно вычислить и кривизну обменных интегралов  $\Delta A/\Delta \omega$ .

Для температурного перехода парамагнетизм-ферромагнетизм при  $T_C$  в редкоземельных ферромагнитных металлах, где обменное взаимодействие между локализованными моментами 4f-ионов осуществляется через электроны проводимости, выполняется соотношение:

$$T_C = \frac{2}{3k_B} G A_{\text{косв}}. \quad (9)$$

Здесь  $G = (g_j - 1)^2 J(J + 1)$  - фактор де Жана,  $A_{\text{косв}}$  -интеграл косвенного обменного взаимодействия между ионами редких земель,  $g_j$  — фактор Ланде, определяется квантовыми числами, характеризующими основные состояния иона редкой земли,  $J$  — квантовое число полного механического момента этого иона. Изменение  $T_C$  с давлением в этом случае:

$$\frac{\Delta T_C}{\Delta P} = \frac{2}{3k_B} G \frac{\Delta A}{\Delta P}. \quad (10)$$

Давление приводит к изменению атомного объема и межатомных расстояний. Из формул (8) и (10) следует, что смещение температуры Кюри при всестороннем сжатии содержит информацию о зависимости обменных интегралов от давления и средних межатомных расстояний.

Проведенные исследования смещения температур  $T_C$  с давлением показали, что давление смещает  $T_C$  в кристаллических и аморфных 3d- и 4f-магнетиках со скоростью примерно 0,1—10 К на 1 кбар. В настоящее время во многих лабораториях существуют установки, которые позволяют создавать давление более 300 кбар. У большинства магнетиков  $T_C$  понижается с давлением, поэтому под действием столь больших давлений ферромагнетик превращается в парамагнетик.

Зависимость обменных интегралов в кристаллических магнетиках от межатомных расстояний в кристаллах с некубической решеткой, как правило, резко анизотропна, то есть значение  $\Delta A/\Delta a$  и  $\Delta A/\Delta c$  ( $a$  и  $c$  — параметры кристаллической решетки) могут различаться как по величине, так и по знаку. Это связано с тем, что при всестороннем сжатии изменение перекрытия электронной плотности магнитных электронов вдоль разных осей неодинаково. Заметный вклад в эффект смещения температур Кюри с давлением вносит изменение энергетического спектра электронов проводимости, а также изменение обменного взаимодействия между локализованными магнитными электронами и электронами проводимости. Для зонных магнетиков необходимо учитывать влияние атомного объема на

параметры зонной структуры (ширина зоны и плотность состояний на уровне Ферми).

Проведенные в последнее время исследования влияния всестороннего сжатия на температуры магнитных фазовых превращений редкоземельных магнетиков установили, что сжатие вызывает главным образом усиление антиферромагнитных взаимодействий как в кристаллическом, так и в аморфном состоянии. Это явление объясняется возрастанием при сжатии перекрытия электронных плотностей 5d- и 3d-электронов, которые являются "переносчиками" антиферромагнитных обменных взаимодействий. В кристаллических материалах действие высоких давлений в ряде случаев приводит к превращению ферромагнетиков в антиферромагнетики с угловым расположением магнитных моментов, а в аморфных материалах — к превращению ферромагнетиков в сперомагнетики и спиновые стекла.

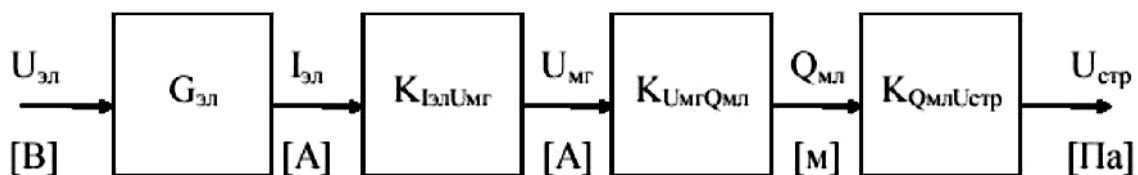
Таким образом, влияние магнитного упорядочения на упругие свойства проявляется в магнитоупругих эффектах, которые содержат важную информацию о магнитоупругих взаимодействиях в твердых телах, о зависимости от межатомных расстояний электронной структуры, обменных и магнитокристаллических взаимодействий. Многие магнитоупругие эффекты весьма перспективны для использования в технике, а некоторые (магнитострикция) уже нашли значительное применение.

## **2.2. Информационное моделирование магнитострикционного эффекта**

### **2.2.1. Описание магнитострикционного преобразователя**

Рассмотрим магнитострикционный эффект подробнее. В магнитострикционном преобразователе происходит композиционное преобразование энергии: из электромагнитной в механическую, а затем из механической в стрикционную (акустическую).

При этом происходит преобразование входного электромагнитного сигнала в механическое напряжение звукопровода (звено 1), вследствие которого по нему распространяется упругая (ультразвуковая) волна (звено 2). Схема преобразования прямого магнитострикционного эффекта представлена на рис. 2.7.



**Рис.2.7.** *Прямой магнестрикционный эффект:  $U_{эл}$  - электрическое напряжение, приложенное к обмотке катушки [В];  $I_{эл}$  - электрический ток, наведенный в обмотке катушки [А];  $U_{мг}$  - магнитодвижущая сила [А];  $G_{эл}$  - электрическая проводимость при постоянном потоке (при  $U_{мг} = const$ ), т.е. величина, обратная активному сопротивлению обмотки плюс сопротивление индуктивности рассеяния [1/Ом];  $K_{I_{эл}U_{мг}}$  - коэффициент электромагнитной связи [витки];  $K_{U_{мг}Q_{мл}}$  - коэффициент магнитомеханической связи [А/м];  $Q_{мл}$  - механический линейный заряд, т. е. деформация [м];  $K_{Q_{мл}U_{стр}}$  - коэффициент механострикционной связи [м/Па = м<sup>2</sup>\*с<sup>2</sup>/кг<sup>2</sup>];  $U_{стр}$  - стрикционное воздействие, т.е. механическое напряжение [Па]*

### 2.2.2. Информационный метод проектирования новых технических решений

При конструировании систем перемещения перед инженерами возникают задачи поиска не только новых физических принципов действия, но и задачи сравнения и улучшения характеристик существующих технических решений. Широкое внедрение САПР даже на начальных этапах разработки систем перемещения позволит расширить объем используемых специалистами знаний, сократить время и трудоемкость создания новых изделий за счет выбора наиболее эффективных решений при их моделировании на ЭВМ, и, как следствие, уменьшить стоимость. Поэтому актуальной становится задача создания автоматизированной системы синтеза новых технических решений систем перемещения для систем управления различного назначения, поиска и анализа прототипов и аналогов.

Одним из способов реализации такого подхода является использование автоматизированного банка данных физико-технических эффектов (ФТЭ) и изобретений. Увеличение объема информации в банке данных, связанное с появлением новых ФТЭ и изобретением новых промышленных образцов, значительно расширяет возможности эффективного синтеза технических решений и анализа существующих конструкций. Пополнение имеющегося банка данных хотя бы одним ФТЭ позволяет получить большее количество новых вариантов физического принципа действия (ФПД) технического устройства, и, следовательно, необходимая часть системы - это автоматизированный отбор наилучших из полученных вариантов.

Информационная модель цепей (ИМЦ) различной физической природы и аппарат параметрических структурных схем (ПСС) являются одним из подобных методов. Использование ИМЦ и ПСС требует большего времени для составления паспортов ФТЭ, но в то же время позволяет сделать процесс поискового конструирования технических устройств более содержательным, раскрывает качественные и количественные связи, акцентирует внимание разработчиков на наиболее ответственных узлах объекта. И, кроме того, позволяет перейти к структурной схеме устройства в целом и по ней составить уравнение, т.е. дать количественную оценку характеристик устройства.

### 2.2.3. Основные понятия информационного моделирования

Целью определенной физической природы называется идеализированная материальная среда, имеющая определенные геометрические размеры и характеризующаяся физическими константами, присущими только явлениям данной физической природы.

Величины цепи одной и той же физической природы изменяются в широких пределах и характеризуют внешнее воздействие на цепь данной физической природы и ее реакцию на него.

Параметры характеризуют относительную неизменность материальной среды, в которой протекают физические процессы.

Информационная модель оперирует обобщенными величинами и параметрами:  $N$  - мощность,  $P$  - импульс,  $Q$  - заряд реакции,  $U$  - сила воздействия,  $I$  - скорость реакции,  $R$  - сопротивление,  $G = 1/R$  - проводимость,  $C$  - емкость,  $W = 1/C$  - жесткость,  $L$  - индуктивность,  $D = 1/L$  - дедуктивность.

Синтезированные параметрические структурные схемы в ИМЦ ранжируются с помощью 10 параметров, важнейший из которых - чувствительность.

### 2.2.4. Информационная модель магнитострикционного эффекта

Примем в качестве величины воздействия величину индукции магнитного поля  $B$  с размерностью  $[Tл] = [\frac{кг}{с^2 \cdot А}]$ , так как именно данная величина является базовой при прямом магнитострикционном эффекте. Если принять индукцию в качестве величины воздействия, то в качестве параметра реакции должна выступать величина с размерностью  $\frac{м^2 \cdot А}{с}$ , которая есть отношение магнитного момента  $p_m$  к времени воздействия  $t$  или более точно  $\frac{dP_m}{dt}$ . Физический смысл данной величины есть скорость

изменения магнитного момента, возникающего из-за деформации образца в магнитном поле.

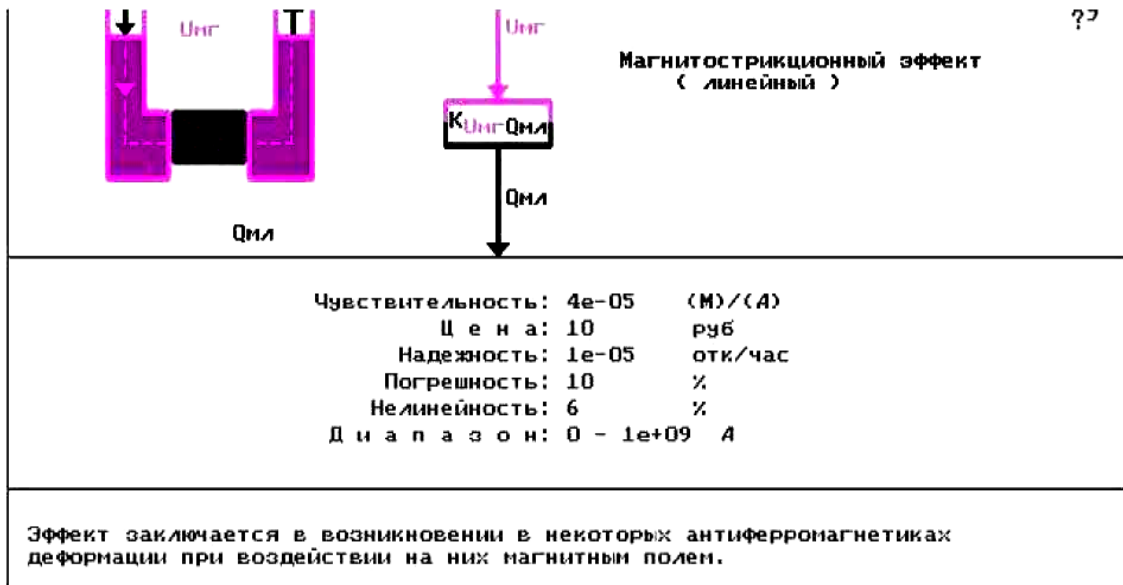


Рис. 2.8. Магнитострикционный эффект

Тогда на основе понятия полного термодинамического потенциала внешнего воздействия, а также после применения теории подобия и анализа размерностей могут быть выведены основные величины и параметры:

<b>Величины:</b>		
Воздействие	$U = B \text{ [Тл]}$	(1)
Реакция	$I = \frac{dp_m}{dt} = \left[ \frac{M^2 * A}{c} \right]$	(2)
Заряд	$Q = p_m \text{ [M}^2 * \text{A]}$	(3)
Импульс	$P = \int B * dt = \left[ \frac{кг^2}{A * c} \right] = \text{[Тл * с]} = \text{[кг/Кл]}$	(4)
<b>Параметры:</b>		
Емкость	$C = \frac{Q}{U} = \frac{p_m}{B} = \left[ \frac{M^2 * A^2 * c^2}{кг^2} = \frac{Bб * A}{M^3 * M} \right] = \rho_m * H$	(5)
Сопротивление	$R = \frac{\int B * dt}{p_m} = \left[ \frac{кг^2 * c}{A^2 * M^2} \right]$	(6)
Проводимость	$G = \frac{p_m}{\int B * dt} = \left[ \frac{A^2 * M^2}{кг^2 * c} \right]$	(7)
	$W = \frac{1}{C} = \frac{B}{p_m} = \left[ \frac{кг^2}{M^2 * A^2 * c^2} \right]$	(8)
Индуктивность	$L = P/I = \frac{(\int B * dt) * dt}{dp_m} = \left[ \frac{кг^2}{A^2 * M^2} \right]$	(9)

Где

$\rho_m$  - плотность магнитного заряда;  
 $H$  - напряженность магнитного поля;

$p_m$  - магнитный момент.

Физический смысл емкости  $C$  есть (с точностью до коэффициента)

механический момент  $\vec{M} = \vec{p} \times \vec{B}$  образца.

Далее, воспользуемся известной зависимостью:

$$\frac{\Delta V}{V} \cong \frac{H^2}{2} \cdot \left[ \beta \chi - \left( \frac{\partial \chi}{\partial p} \right)_T \right] \Rightarrow H \cong \sqrt{\frac{2 \cdot \Delta V}{V} \cdot \frac{1}{k}} = \sqrt{\frac{2\lambda}{k}}$$

где:

$\lambda_v$  - объемная магнитострикция;

$H$  - напряженность магнитного поля;

$$\beta = -\frac{1}{V} \cdot \left( T \frac{V}{T_p} \right)_{H,T}$$

- изотермическая сжимаемость материала;

$\chi = \frac{p_m}{HV}$  - изотермическая магнитная восприимчивость (причем обычно  $\chi$  не зависит от  $H$  и может считаться характеристическим свойством образца);

$k$  - коэффициент, равный

$$\left[ \beta \chi - \left( \frac{\partial \chi}{\partial p} \right)_T \right].$$

Из равенства:

$$B = \mu_0 \cdot \left( 1 + \frac{C}{T - T_c} \right) \cdot H,$$

где  $\mu_0$  - магнитная постоянная,  $T$  - температура материала,  $T_c$  - точка Кюри,  $C$  - постоянная Кюри, получаем:

$$B = \mu_0 \cdot \left( 1 + \frac{C}{T - T_c} \right) \cdot \sqrt{\frac{2\lambda_v}{k}},$$

что позволяет связать величину воздействия с параметрами материала. Здесь  $\mu_0$  - магнитная постоянная;  $T$  - температура материала;  $T_c$  - точка Кюри;  $C$  - постоянная Кюри;  $\lambda_v$  - объемная магнитострикция.

Таким образом, в рамках информационного метода возможно построить модель магнитострикционного эффекта, которая пригодна для автоматизации научно-технического творчества и учитывает свойства магнитострикционных материалов.

В магнитострикционных эффектах осуществляется преобразование входной величины  $U_{MG}$  магнитной цепи в величину  $Q_{ML}$  механической линейной цепи. Таким образом, схема прямого магнитострикционного эффекта есть  $U_{MG} - K_{U_{MG}Q_{ML}} - Q_{ML}$ , где  $K_{U_{MG}Q_{ML}}$  есть коэффициент преобразования или чувствительность магнитострикционного эффекта.

Согласно определению чувствительность - отношение изменения выходной величины  $B_{ВЫХ}$  к изменению входной величины  $B_{ВХ}$ :

$$K(B_{-j}) = \frac{dB_{-i}}{dB_{-j}}$$

В общем случае  $B_{\text{ВЫХ}i}$  и  $B_{\text{ВХ}j}$  - величины разной физической природы ( $i \neq j$ ).

Известно, что для характеристики интенсивности магнитострикционных колебаний в переменном магнитном поле при наличии постоянного подмагничивания вводится магнитострикционная постоянная  $a$ .

$$a = \frac{\Delta\sigma}{\Delta B} \quad \text{при} \quad \frac{\Delta l}{l} \neq 0 \quad [A/M]$$

Важным параметром магнитострикционных материалов является коэффициент магнитомеханической связи  $K_{\text{ММ}}$ :

$$K_{\text{ММ}} = \sqrt{\frac{4 \cdot \pi \cdot a^2 \cdot \mu}{E}} = \sqrt{\frac{\mu \cdot E \cdot \lambda_s}{B}}$$

где

$\lambda_s$  - магнитострикция при насыщении намагничивания;

$\mu$  - действующая магнитная проницаемость материала звукопровода при наличии постоянного подмагничивающего поля [ $H/A^2$ ];

$E$  - модуль упругости, [Па].

Коэффициент магнитомеханической связи :

$$K_{\text{УМГ}Q_{\text{мл}}} = k \cdot a = F/\Phi \quad (\text{при } x=0) = M/x \quad (\text{при } \Phi=0) \quad [A/M]$$

где

$x$  - деформация, м;

$a$  - магнитострикционный коэффициент (см. выше, А/м);

$k$  - безразмерный коэффициент, зависящий от устройства преобразователя;

$\Phi'$  - производная магнитного потока по времени В;

$\Phi$  - магнитный поток Вб;

$F$  - сила, Н.

Таким образом, коэффициент магнитомеханической связи  $K_{\text{ММ}}$  характеризует эффективность энергетических преобразований подводимой магнитной энергии в энергию упругих волн и наоборот может выступать в качестве параметра чувствительности при описании магнитострикционного эффекта с помощью информационного метода.

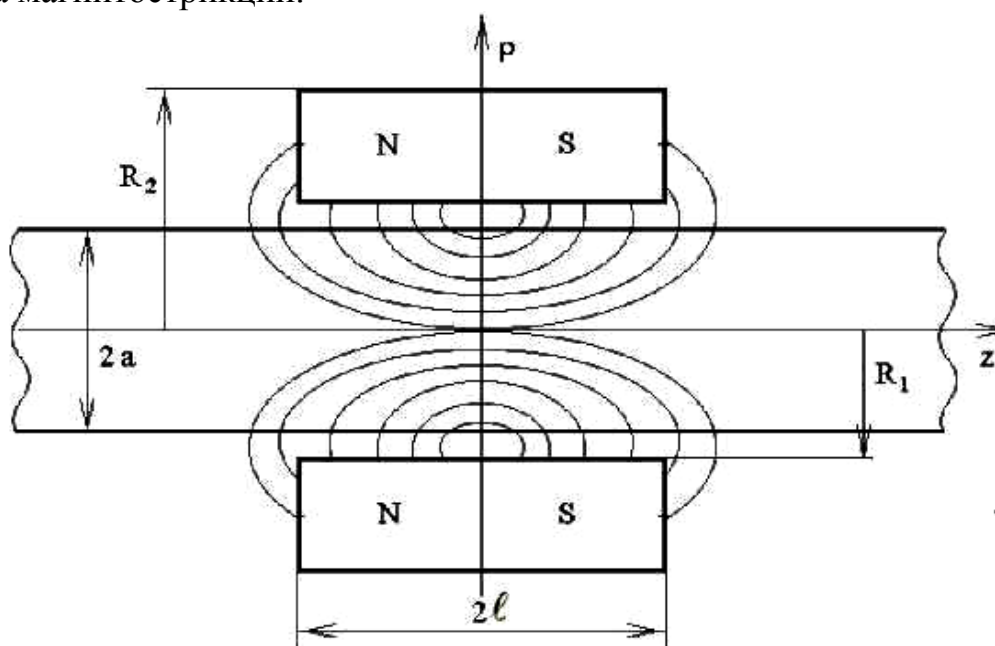
Поскольку предложенная модель магнитострикционного эффекта связывает параметры ИМЦ и параметры ФТЭ со свойствами магнитострикционных материалов, то она дает возможность учесть свойства материалов при проектировании датчиковой аппаратуры, что особенно важно при разработке прецизионной и высокочувствительной техники.



### 2.3. Расчет магнитного поля постоянного магнита для магнитострикционных систем перемещений

Напряженность магнитного поля находится в точках цилиндрического волновода намагниченного кольцевым постоянным магнитом с заданной постоянной индукцией внутри магнита. В работе трехмерный интеграл с дельта функцией сводится к двумерным интегралам, по которым вычисляется продольная составляющая вдоль волновода напряженности магнитного поля и радиальная составляющая. По результатам расчета приведены и проанализированы графики.

Исследованиями установлено, что конструкции преобразователей перемещений на основе магнитострикционного эффекта являются наиболее простыми и надежными в эксплуатации, и имеют наибольшее распространение в промышленности. Однако, на сегодня практически нерешенной остается задача оптимизации проектирования данного класса преобразователей из-за отсутствия полной аналитической модели, описывающей электромагнитомеханические процессы в МПП на основе эффекта магнитострикции.



*Рис. 2.9. Распределение магнитного поля в волноводе МПП в зоне действия кругового постоянного магнита*

В настоящей работе рассматривается один из главных этапов проектирования - расчет постоянного магнитного поля. Магнитострикционные устройства, основанные на применении эффекта Видемана, предполагают наложение на материал постоянного магнитного поля. Чаще всего его образуют с помощью постоянного магнита. Поле последнего зависит от конфигурации магнита. Мы будем рассматривать кольцевой магнит, находящийся на волноводе (рис. 2.9).

По результатам работы напряженность кольцевого постоянного магнита внутри звукопровода определяется функциями:

$$H_z(\rho, z) = \frac{1}{4\pi\mu_0} \int_{R_1}^{R_2} \int_0^{2\pi} \int_{-\ell}^{\ell} \frac{(z - z_1) \rho_M}{R^3} \rho_1 d\rho_1 d\theta_1 dz_1,$$

$$H_\rho(\rho, z) = \frac{1}{4\pi\mu_0} \int_{R_1}^{R_2} \int_0^{2\pi} \int_{-\ell}^{\ell} \frac{(\rho - \rho_1 \cos \theta_1) \rho_M}{R^3} \rho_1 d\rho_1 d\theta_1 dz_1, \quad (11)$$

Где  $R = \sqrt{(z - z_1)^2 + \rho^2 + \rho_1^2 - 2\rho\rho_1 \cos \theta_1}$ ; (12)  
 $\rho_1, \theta_1, z_1$  - координаты точки внутри магнита;  
 $\rho_M = B_0 [\delta(z_1 + 1) - \delta(z_1 - 1)]$ ,

где:  $\rho_M$  - плотность фиктивных магнитных зарядов;  
 $\delta(x)$  - дельта функция Дирака.

Вычисление тройных интегралов весьма затруднительно.

Используя свойства функции (12), тройные интегралы (11) можно привести к двойным интегралам:

$$H_z(\rho, z) = \frac{B_0}{4\pi\mu_0} \int_{R_1}^{R_2} \int_0^{2\pi} \frac{(z + \ell) \rho_1 d\rho_1 d\theta_1}{\left( (z + \ell)^2 + \rho^2 + \rho_1^2 + 2\rho\rho_1 \cos \theta_1 \right)^{\frac{3}{2}}} -$$

$$- \frac{B_0}{4\pi\mu_0} \int_{R_1}^{R_2} \int_0^{2\pi} \frac{(z - \ell) \rho_1 d\rho_1 d\theta_1}{\left( (z - \ell)^2 + \rho^2 + \rho_1^2 + 2\rho\rho_1 \cos \theta_1 \right)^{\frac{3}{2}}}$$

$$H_\rho(\rho, z) = \frac{B_0}{4\pi\mu_0} \int_{R_1}^{R_2} \int_0^{2\pi} \frac{(\rho - \rho_1 \cos \theta_1) \rho_1 d\rho_1 d\theta_1}{\left( (z + \ell)^2 + \rho^2 + \rho_1^2 + 2\rho\rho_1 \cos \theta_1 \right)^{\frac{3}{2}}} -$$

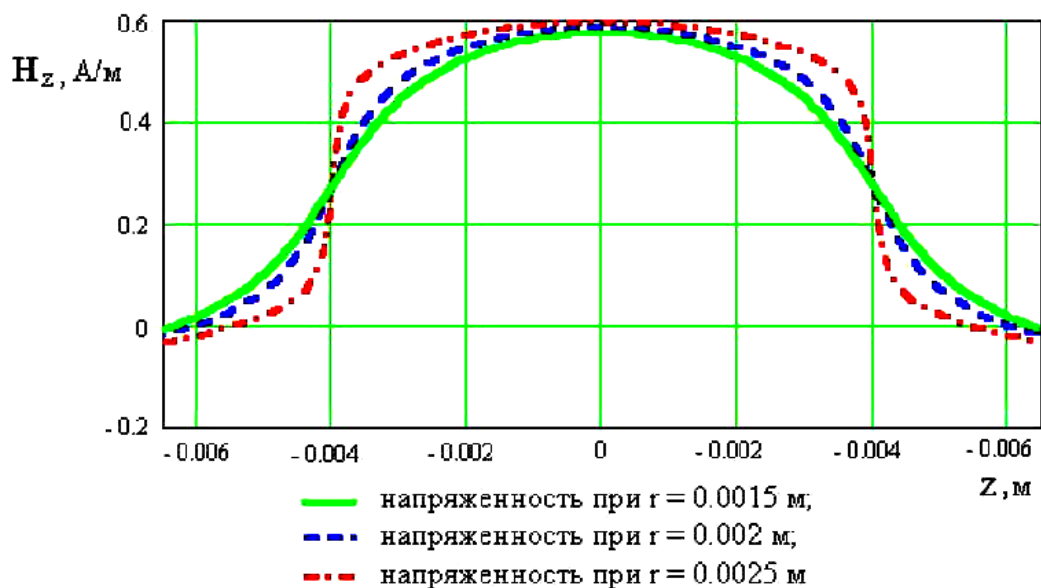
$$- \frac{B_0}{4\pi\mu_0} \int_{R_1}^{R_2} \int_0^{2\pi} \frac{(\rho - \rho_1 \cos \theta_1) \rho_1 d\rho_1 d\theta_1}{\left( (z - \ell)^2 + \rho^2 + \rho_1^2 + 2\rho\rho_1 \cos \theta_1 \right)^{\frac{3}{2}}}$$

В интегралах уже расставлены пределы интегрирования. Их значения можно получить с помощью двойных квадратур.

Произведен расчет в программе MatchCad и построены графики изменения напряженности магнитного поля вдоль волновода. Исходными данными являются следующие значения:

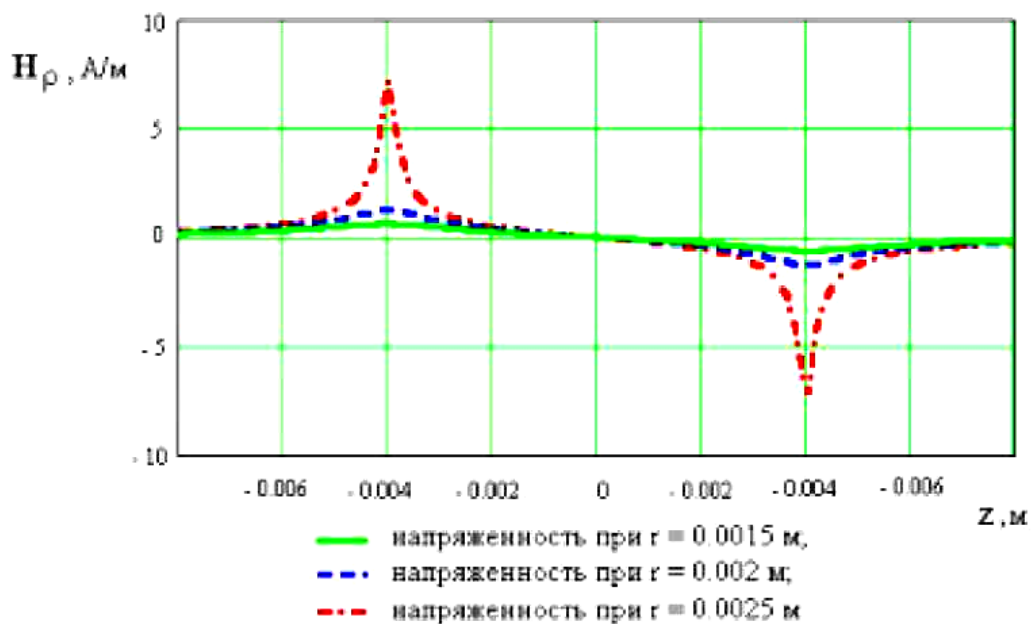
1) длина магнита  $l = 0,004$  м;

- 2) радиус волновода  $a = 0,0025$  м;
- 3) внутренний радиус магнита  $R_1 = 0,0026$  м;
- 4) внешний радиус магнита  $R_j = 0,03$  м;
- 5) магнитная индукция внутри магнита  $B_0 = 10^{-6}$  Тл.



**Рис. 2.10.** Продольная составляющая напряженности магнитного поля

На рис. 2.10, представлен график изменения напряженности магнитного поля в точках цилиндра с радиусами 0.0015 м, 0.002 м, 0.0025 м вдоль оси  $OZ$ . Мы видим, что график функции имеет колокообразную форму, причем при приближении к поверхности волновода возрастает. Так как длина магнита  $l = 0.004$ , то магнитное поле выходит за пределы магнита в одну и другую сторону, причем это влияние не распространяется больше, чем на 25% длины магнита.



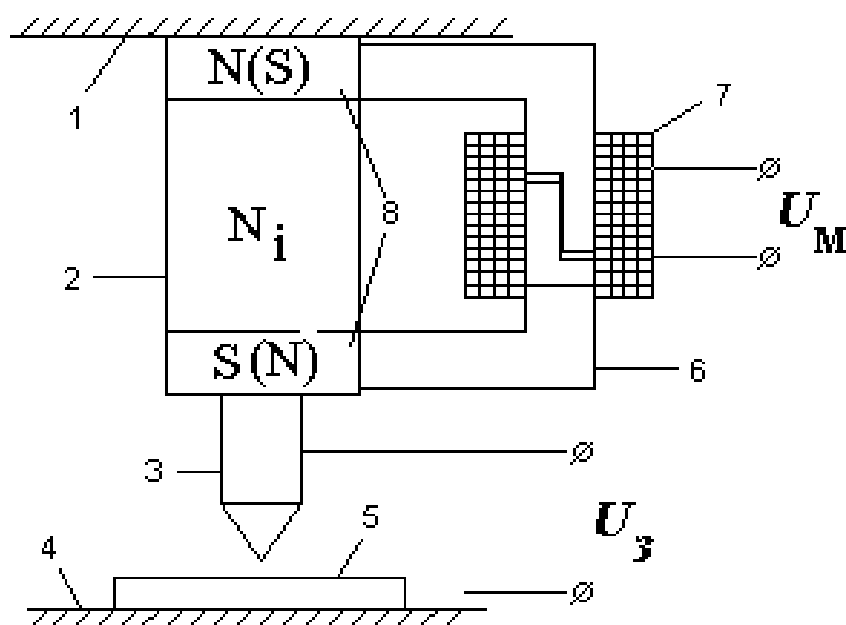
**Рис. 2.11.** Радиальная составляющая напряженности магнитного поля

На рис. 2.11 представлен график изменения радиальной составляющей для точек тех же цилиндров. Следует отметить, что по оси волновода  $\rho = 0$  поле отсутствует и увеличивается его напряженность по мере приближения к поверхности волновода. На поверхности волновода по торцам магнита наблюдается резкое возрастание напряженности магнитного поля и затем на расстоянии снова 25% длины магнита затухает до нуля. Внутри магнита напряженность медленно изменяясь переходит от северной стороны магнита к южной. Для магнитоэлектрики наиболее важный участок изображен на рис.2 от -0.004 до 0.004 м.

## 2.4. Устройства прецизионного перемещения на основе эффекта магнитоэлектрики

### 2.4.1. Устройство перемещения для нанотехнологии

В основу разработки положена задача обеспечить возможность использования магнитного поля для оперативного управления величиной перемещения зонда при выполнении различных нанотехнологических операций.



**Рис. 2.12.** Устройство перемещений для нанотехнологии

Устройство перемещения (рис. 2.12) содержит привод 2, связанный с неподвижным основанием 1 и зондом 3, подложку 5, установленную на подложкодержателе 4. Привод выполнен в виде цилиндра из магнитоэлектрического материала, торцы которого связаны с полюсами электромагнита 8, которые жестко связаны с его сердечником 6, выполненным составным, с возможностью осевого перемещения составных частей относительно друг друга внутри катушки индуктивности 7, один

полюс электромагнита жестко связан с неподвижным основанием, а другой с зондом.

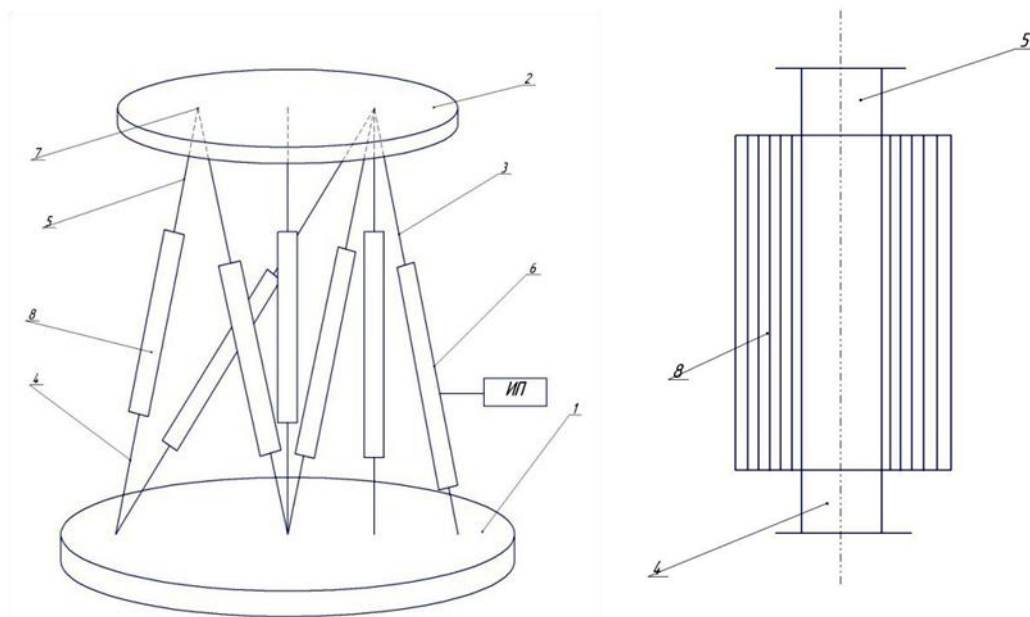
Устройство работает следующим образом.

При подаче напряжения на катушку индуктивности 7 происходит перемещение привода 2, а вместе с ним и зонда 3, и таким образом реализуется нанотехнологический процесс.

Применение устройства перемещения для нанотехнологии позволяет реализовать нанотехнологические процессы на подложке с использованием магнитного поля вместо электрического и повысить точность перемещения.

### 2.4.2. Привод для нанотехнологии

В основу разработки положена задача обеспечить возможность использования магнитного поля для оперативного управления величиной перемещения подложкодержателя при выполнении различных нанотехнологических операций.



*Рис. 2.13. Привод для нанотехнологии*

Привод для нанотехнологии, содержащий основание 1, выходное звено 2, шесть стержней 3, шарнирно связанных одним концом 4 с основанием, а другим 5 - с выходным звеном, приводы перемещения 6 выходного звена, кинематически связанные с соответствующими стержнями 3, отличающиеся тем, что шарнирные связи 7 выполнены упругими, стержни 3 – из магнитострикционного материала, причем стержни 3 установлены в катушку индуктивности 8, связанную с источником питания.

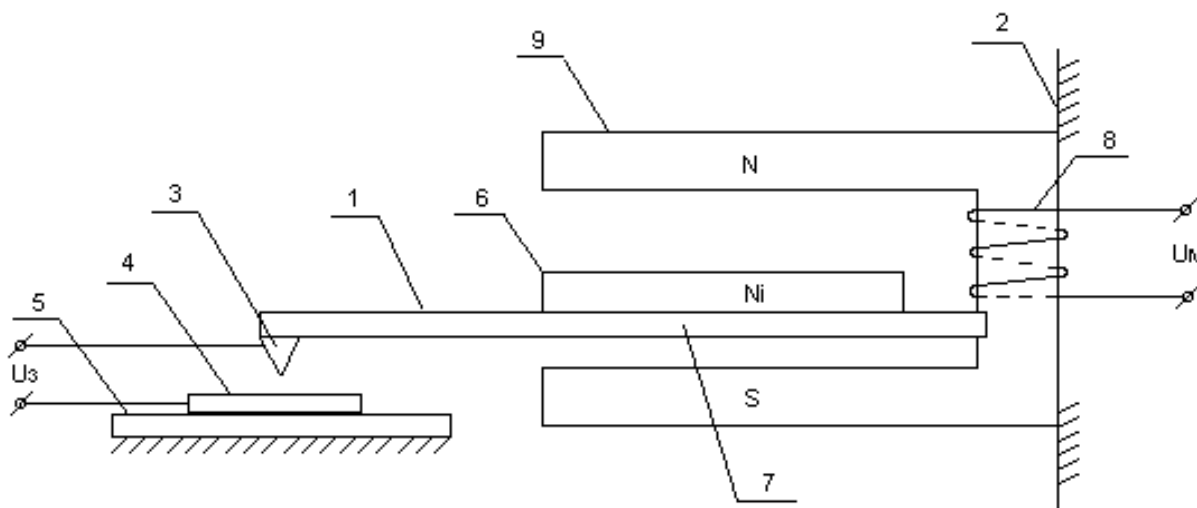
Устройство работает следующим образом.

При подаче напряжения на катушку индуктивности 8 происходит перемещение привода 6, а вместе с ним и выходного звена 2, и таким образом реализуется нанотехнологический процесс.

Применение устройства перемещения позволяет реализовать нанотехнологические процессы на подложке с использованием магнитного поля вместо электрического и повысить точность перемещения.

### 2.4.3. Устройство наноперемещения зонда

В основу разработки положена задача обеспечить возможность использования магнитного поля для оперативного управления величиной перемещения зонда при выполнении различных нанотехнологических операций.



*Рис. 2.14. Устройство наноперемещений зонда*

Устройство наноперемещений зонда, содержащее привод 1, связанный с неподвижным основанием 2 и зондом 3, подложку 4, установленную на подложкодержателе 5, отличающееся тем, что привод 1 выполнен в виде магнитоупругого биморфа – жестко связанных между собой двух пластин, одна из которых выполнена из магнитоупругого материала 6, а другая - из немагнитного, диэлектрического 7, на консольной части которой закреплен зонд 3, а другая часть немагнитной, диэлектрической пластины 7 жестко связана с сердечником электромагнита 8, между полюсами которого 9 расположен магнитоупругий биморф.

Устройство работает следующим образом.

При подаче напряжения на электромагнит 8 происходит перемещение пластины из магнитоупругого материала 6, а вместе с ним и привода 1, и таким образом реализуется нанотехнологический процесс.

На основании результатов, полученных при тестировании уже существующих магнитоупругих приводов, можно говорить о

повышении точности позиционирования в десятки раз по сравнению с пьезоэлектрическими приводами.

### Заключение

Для систем перемещения в нанотехнологии особый интерес представляют сплавы, в которых гигантская магнитострикция реализуется в области комнатных температур. Особо большие перспективы практического использования имеют так называемые интерметаллические соединения редкоземельных элементов с металлами группы железа.

Редкоземельный сплав Terfenol-D ( $Tb_{0.27}Dy_{0.73}Fe_{1.95}$ ) дал возможность приступить к практическому использованию эффекта "гигантской магнитострикции" для силового позиционирования, ведь в этом случае в магнитных полях  $\sim 0.5$  Т относительное удлинение достигало  $10^{-3}$ , а возникающее усилие достигало  $100$  Н/мм<sup>2</sup> сечения стриктора. Таким образом, стриктор длиной 1 см и таком же диаметре мог обеспечить силовое позиционирование в пределах 10 мкм с усилием до  $10^5$  Н. Так же этот сплав обладает пониженной магнитной анизотропией с сохранением высокой магнитострикции что позволяет управлять гигантской магнитострикцией с помощью малого поля.

### Литература

1. В.Л. Миронов. Основы сканирующей зондовой микроскопии. Нижний Новгород, РАН Институт физики микроструктур. 2004г.
2. Физика и химия редкоземельных элементов. Справочник / Под ред. Е.М. Савицкого. М.: Металлургия, 1982.
3. Тейлор К. Интерметаллические соединения редкоземельных металлов. М.: Мир, 1974.
4. Никитин С.А. Магнитные свойства редкоземельных металлов и их сплавов. М.: Изд-во МГУ, 1989. 248 с.
5. Белов К.П. Магнитострикционные явления и их технические приложения. М.: Наука, 1987. 159 с.
6. Андреев А.С., Никитин С.А., Спичкин Ю.И. Зависимость интегралов обменного взаимодействия от атомного объема в аморфных сплавах и кристаллических соединениях R—Fe // Вестн. МГУ. Сер. 3. 1993. Т. 34, № 1. С. 55-71.
7. Надеев А. И. Магнитострикционные интеллектуальные преобразователи параметров движения. Монография / Астрахан. гос. техн. ун-т. - Астрахань, АГТУ, 1999. - 155 с- деп. в ВИНТИ 22.07.99 № 2385-В99.
8. Электрический микродвигатель постоянного тока для работы в чистых помещениях. Патент РФ на изобретение №2003214. Опубликовано 15.11.93. Бюл. №41-42. Ивашов Е.Н., Оринчев С.М., Степанчиков С.В. и др.



## Глава 3. Формирование объектов в лазерно-зондовой микроскопии

### Введение

Лазерные системы позволяют генерировать световые пучки в широком диапазоне длинны волн и плотности потока излучения. Поглощение мощного светового излучения сопровождается процессами тепло- и массопереноса, развитие которых зависит от энергетических и пространственно-временных характеристик излучения. При действии сфокусированного мощного лазерного излучения на поверхность материала вещество нагревается, плавится, частично испаряется и ионизируется. В неоднородно нагретом веществе возникает течение жидкости, паров, плазмы и окружающего газа. Перемещение вещества, в свою очередь, оказывает существенное влияние на распространение лазерного излучения, изменяя фокусировку и условия поглощения и отражения излучения. Возникают сложные многофазные гидро- и газодинамические течения, согласованные с распространением лазерного излучения в сильно поглощающей и преломляющей оптически нелинейной среде. Соответствующий выбор длины волны излучения, интенсивности, времени воздействия, вида и давления окружающей атмосферы позволяет осуществлять различные технологические процессы, ряд которых невозможен без применения лазеров.

### 3.1. Механическое воздействие лазерного излучения на атомы

В исследованиях воздействия излучения на микрочастицы вещества прежде всего обращают внимание на электромагнитный характер воздействия, приводящий, в частности, и к передаче энергии излучения во внутренние степени свободы микрочастиц. Помимо этого существует, однако, и прямое механическое воздействие, которое обычно считается слабым, но в определенных условиях и оно способно проявиться довольно ярко.

С тех пор как была открыта квантовая природа излучения, стало ясно, что квант излучения (фотон) обладает энергией  $\hbar\omega$  и импульсом  $\hbar k$ , где  $\omega$  — частота (круговая) излучения;  $k$  — волновой вектор;  $\hbar$  — постоянная Планка. Величина волнового вектора связана с длиной волны излучения  $\lambda$  соотношением  $k = 2\pi/\lambda$ . При поглощении фотона частицей вещества последней передается импульс  $\hbar k$ . В случае оптической или более длинноволновой области спектра величина импульса фотона очень мала по сравнению с характерным значением импульса, которым обычно обладает частица вещества, хотя бы вследствие ее теплового движения. Например, если в качестве частицы вещества выступает атом массы  $M$  со скоростью теплового движения  $v_T = \sqrt{2k_B T/M}$  ( $k_B$  — постоянная Больцмана,  $T$  — температура), то отношение импульса фотона к импульсу такого атома есть

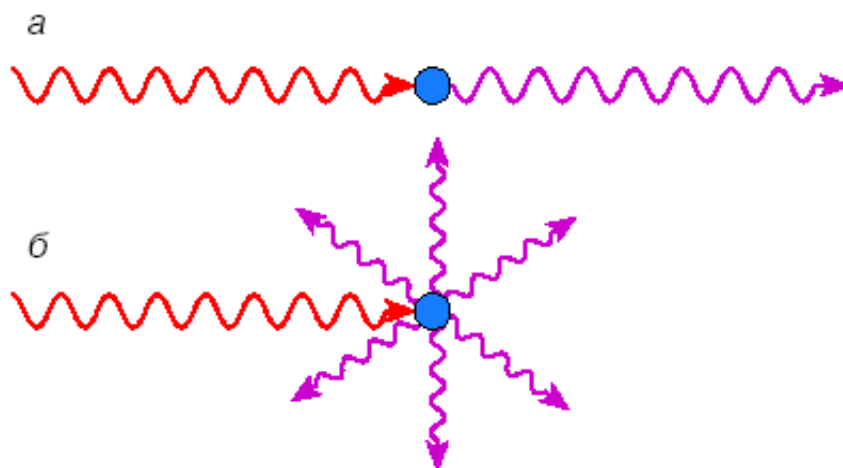
$$\frac{\hbar k}{Mv_T} = \frac{2\pi\hbar}{\lambda\sqrt{2k_B T/M}}. \quad (1)$$

Даже в случае атома водорода при комнатной температуре оптический квант ( $\lambda \sim 0,5$  мкм) при поглощении чрезвычайно слабо меняет состояние его движения: отношение (1) составляет  $\approx 3 \cdot 10^{-4}$ . Разумеется, чем тяжелее атом (молекула), тем слабее это влияние. Тем не менее эффект может накапливаться, если частица вещества испытывает многократные акты поглощения направленного излучения. В обычных условиях, однако, приобретенный частицей импульс через определенное время теряется за счет взаимодействия ее с окружением и эффект остается слабым, если не использовать специальные источники излучения. Источники излучения, существовавшие в прежние времена (в долазерную эпоху), не могли обеспечить яркого проявления эффекта, а для его наблюдения требовалось высокое экспериментальное искусство. Существуют, правда, специфические условия, реализуемые в космических просторах в окрестностях звезд. Здесь атомы и молекулы вещества могут находиться в свободном состоянии (без столкновений друг с другом) исключительно долгое время, да и само излучение здесь достаточно интенсивное. В ряде случаев интенсивность излучения звезды достаточно высока для того, чтобы сила светового давления преодолела силу гравитационного притяжения. Тогда происходит накопление частицей импульса, направленного от звезды, и она с ускорением покидает окрестность звезды. Это хорошо известный и часто наблюдаемый эффект звездного ветра.

Резко усилить эффект светового давления позволило одно из уникальных свойств лазерного излучения — возможность сосредоточивать его энергию в узкой спектральной области и настроить эту область в резонанс с квантовым переходом внутри атома (резонансное световое давление). Величина этой энергии на несколько порядков выше той, которая могла быть достигнута с помощью резонансных спектральных источников, использованных в экспериментах П.Н. Лебедева и более поздних экспериментах С.Э. Фриша. Чтобы эффект светового давления мог накапливаться, используют камеры с высоким вакуумом. Атомы исследуемого элемента либо находятся в малом количестве в этих камерах в обычных (равновесных) условиях, либо поставляются туда (инжектируются) в виде атомного пучка при постоянно действующей вакуумной откачке. В любом случае обеспечиваются условия, при которых атом может пролететь расстояние порядка метра с тепловой скоростью без столкновений с другими атомами. Частота излучения настраивается в резонанс с квантовым переходом атома из основного энергетического состояния (обозначим это состояние индексом 0) в первое возбужденное состояние (ему припишем индекс 1). Все остальные внутренние состояния (уровни энергии) атома не принимают участия в процессе. На такой простой (двухуровневой) модели атома вполне можно уяснить главные черты явления.

Рассмотрим прежде всего случай бегущей монохроматической волны излучения. Примем следующие начальные условия: атом находится в основном состоянии, а частота излучения со близка (в системе координат, связанной с атомом) к частоте  $\omega_{10}$  квантового перехода в атоме.

Механическое действие излучения на атом в данной ситуации развивается по следующему сценарию. Атом поглощает квант излучения и как следствие воспринимает квант импульса  $\hbar k$ . При этом атом оказывается в возбужденном состоянии 1. Акт поглощения фотона происходит тем быстрее, чем выше интенсивность излучения. Дальнейшая судьба атома реализуется альтернативно по одному из двух путей (рис. 3.1): атом испускает либо точно такой же фотон, что и поглощенный ранее, то есть с теми же энергией и импульсом (случай *a* на рис. 3.1), либо фотон с той же энергией, но равновероятно во всех направлениях (случай *б* на рис. 3.1). Первый вариант соответствует процессу вынужденного испускания (этот процесс был теоретически описан А. Эйнштейном и играет определяющую роль в работе всех лазеров), вероятность его тем выше, чем выше интенсивность вынуждающего его внешнего излучения. Во втором случае происходит процесс самопроизвольного (спонтанного) испускания, вероятность которого никак не зависит от интенсивности внешнего излучения (в частности, поэтому равновероятны все направления спонтанно испущенных фотонов). После испускания фотона (как спонтанно, так и индуцированно) атом опять оказывается в основном энергетическом состоянии. На этом завершается единичный цикл, который затем будет повторяться снова и снова.



**Рис. 3.1.** Иллюстрация к процессам вынужденного (*a*) и спонтанного (*б*) испускания. Во втором случае иллюстрацию надо понимать не как испускание многих мелких фотонов, а как испускание одного, но равновероятно во всех направлениях

Отметим принципиальное отличие процессов вынужденного и спонтанного испускания фотона с точки зрения передачи импульса со стороны излучения атому. Если цикл реализовался по каналу поглощение — вынужденное испускание, то импульс атома целиком восстановился, поскольку как перед началом, так и в конце цикла имеем в наличии фотон с одним и тем же импульсом  $\bar{h}k$ . Таким образом, за счет процесса вынужденного испускания атом не получает систематического приращения импульса. Максимум, что может передать излучение атому в данном случае, — это импульс  $\bar{h}k$  в среднем по времени: при достаточно большой интенсивности излучения атом половину времени проводит в возбужденном состоянии, неся в себе при этом импульс фотона  $\bar{h}k$ . Совсем иное дело, когда цикл завершается спонтанным испусканием. В этом случае из-за различия направлений поглощенного и испущенного фотонов в атоме остается ненулевой импульс. При многократном повторении такого типа циклов сумма импульсов отдачи со стороны спонтанно испущенных фотонов в силу изотропности этого процесса близка к нулю, тогда как принимаемый атомом импульс от поглощенных фотонов направленного излучения есть  $N\bar{h}k$ , где  $N$  — число циклов. Это и есть итоговый импульс, который, как видим, способен накапливаться в атоме. Это накопление можно характеризовать средней силой  $F$ , действующей со стороны излучения на атом. Используем соотношение  $F = dp/dt$ : сила есть импульс, передаваемый атому в единицу времени. Последний легко подсчитать, зная число актов спонтанного испускания в единицу времени. Оно есть  $\bar{\omega}\gamma = \bar{\omega}/\tau$ , где  $\bar{\omega}$  — вероятность атому находиться в возбужденном состоянии;  $\gamma$  — так называемая константа спонтанной релаксации, равная обратному времени жизни  $\chi$  возбужденного состояния (это время жизни как раз и обуславливается спонтанным испусканием фотона). Таким образом,

$$F = \frac{dp}{dt} = \bar{h}k\bar{\omega}\gamma .$$

Итак, в бегущей волне излучения основная сила, действующая на атом, возникает благодаря изотропному спонтанному испусканию. Чтобы отличить от других видов сил (о них будет речь дальше), ее стали называть силой спонтанного светового давления, причем она генетически связана с той силой, которую измерял еще П.Н. Лебедев.

Оценим, насколько сильно может повлиять лазерное излучение на состояние движения атома. При относительно слабых интенсивностях излучения величина  $\bar{\omega}$  пропорциональна интенсивности. В поле интенсивного излучения происходит насыщение (величина  $\bar{\omega}$  перестает расти). Максимальное значение, которого в принципе может достигать величина  $\bar{\omega}$ , равно  $1/2$ . Оно реализуется в таких условиях, когда за время порядка  $\tau$  атом много раз попеременно бывает то в основном, то в возбужденном состоянии вследствие вынужденных переходов. Примечателен тот факт, что эти условия легко осуществить за счет фокусировки излучения с помощью даже плохонького (мощностью в несколько милливатт) лазера, если частота излучения настроена

в резонанс с квантовым переходом в атоме. Существующие сейчас лазеры, используемые для подобного типа задач, способны обеспечивать существенно большую мощность в непрерывном режиме (несколько ватт и более).

Итак, максимально возможная сила спонтанного светового давления создает максимальное ускорение

$$a_{\max} = \frac{\hbar k \gamma}{2M}$$

### 3.2. Особенности поглощения светового излучения материалом

Лазерное излучение характеризуется рядом уникальных свойств. К их числу относятся большая интенсивность (плотность тока) электромагнитной энергии, высокая монохроматичность, значительная степень временной и пространственной когерентности. В отличие от других источников электромагнитной энергии лазерное излучение обладает очень узкой направленностью или малой расходимостью луча.

Затухание плотности светового потока в конденсированных средах происходит в соответствии с законом Бугера-Ламберта

$$q = q_0 A \exp\left(-\int \alpha(x) dx\right), \quad (2)$$

где  $q$  и  $q_0$  – плотность потока излучения соответственно на глубине  $x$  от поверхности и на поверхности облучаемого тела;  $A$  – доля поглощенного потока;  $\alpha$  – коэффициент поглощения света.

Для случая изотропной и однородной поглощающей среды выражение (1) приводится к виду.

$$q = A q_0 \exp(-\alpha x).$$

Для металлических материалов поглощение лазерного излучения описывается моделью свободных электронов. Согласно этой модели падающий на поверхность металла световой поток практически полностью поглощается свободными электронами в тонком слое толщиной  $d = \alpha^{-1} = 10^{-5} \dots 10^{-6}$  см. В дальнейшем происходит электрон-электронная релаксация с характерной частотой  $\nu_{ee}$ . Вместе с тем вследствие электрон-фононного взаимодействия с частотой  $\nu_{ep}$  происходит передача энергии от электронного газа к решетке. В результате энергия фононного газа или тепловых колебаний решетки возрастает. Если частота фонон-фононного взаимодействия  $\nu_{pp}$  больше частоты электрон-фононной релаксации  $\nu_{ep}$ , то энергетическая функция распределения фононного газа останется равновесной. Для большей оценки процессов лазерного воздействия на материалы проводят оценку для частоты взаимодействия  $\nu_{ef}$ ,  $\nu_{ee}$ ,  $\nu_{ep}$ ,  $\nu_{pp}$ .

Частоту столкновений электронов с фононами можно оценить с помощью соотношения

$$\nu_{ef} = \frac{\alpha q_0}{h\nu n'}$$

где  $n'$  – концентрация поглощаемых электронов.

Поскольку для лазерного излучения видимого диапазона в ближней инфракрасной области выполняется условие  $h\nu \gg kT_e$  ( $h$  – постоянная Планка;  $\nu$  – частота;  $k$  – постоянная Больцмана;  $T_e$  – температура электрона), то в поглощении будут участвовать лишь электроны с  $E = h\nu$  вблизи уровня Ферми.

Тогда

$$n' = \frac{nh\nu}{E_f} \quad (3)$$

где  $n = 5 \cdot 10^{22} \text{ см}^{-3}$  – число электронов проводимости в единице объема нормальных металлов,  $E_f$  – уровень Ферми. Проводя, например, оценку для излучения лазера на иттий-алюминиевом гранате (вторая гармоника) и принимая  $\alpha = 10^6 \text{ см}^{-1}$ , в соответствии с формулой (3) получаем

$$\nu_{ef} = (1.5 \dots 3.0) 10^2 q_0,$$

где  $\nu_f$  – скорость электрона на поверхности Ферми;  $\sigma_{ee}$  – резерфордовское сечение взаимодействия для рассеяния электронов на электродах;  $kT_e$  – область размытости функции Ферми.

Принимая для начальной и конечной электронных температур значения соответственно 300 К и 3000 К, получаем  $\nu_{ee} = (10^{11} \dots 10^{13}) \text{ с}^{-1}$ .

Частота электрон-фононной релаксации может быть выражена через коэффициент теплообмена между электронами и фононами  $u$ :

$$\nu_{ep} = \frac{u}{\gamma_i C_i} = \frac{\pi^2}{15} \cdot \frac{nk\nu_3^2}{\gamma_i c_i d_0 \nu_f} \quad ,$$

где  $\gamma_i$  – плотность материала;  $c_i$  – теплоемкость;  $d_0$  – постоянная решетки;  $\nu_3$  – скорость звука в металле.

Оценку частоты энергии фононного газа можно выполнить по формуле

$$\nu_{pp} = k\Gamma^2 T_p / (d_0 m_i \nu_3) \quad ,$$

где  $\Gamma$  – параметр Грюнаизена, определяющий коэффициент температурного расширения решетки;  $T_p$  – температура решетки;  $m_i$  – масса иона.

При начальной и конечной температурах решетки 300 К и 3000 К получаем  $\nu_{pp} = (10^{13} \dots 10^{14})c^{-1}$ . Соответствующие времена передачи энергии обратны частотам процессов.

В отличие от других металлов в энергетическом спектре полупроводников и диэлектриков существует запретная зона  $E_d$ , не содержащая энергетических уровней. Поэтому важным фактором является соотношение между энергией кванта  $h\nu$  и шириной запрещенной зоны  $E_d$ . В зависимости от этого соотношения для полупроводниковых материалов выделяют следующие типы механизмов поглощения:

а) металлический [ $h\nu < E_d$ ,  $E_i \leq kT$ , где  $E_i$  - энергия примесных уровней;  $T$  - температура материала, который характерен для легированных полупроводников, когда в зоне проводимости имеется большая концентрация свободных носителей  $n_e \approx \exp[-E_i/(kT)]$ ;

б) примесный ( $h\nu < E_d, h\nu > E_i > kT$ ), который характерен для примесных полупроводников, когда энергия кванта света меньше ширины запрещенной зоны, но превосходит энергию донорных или акцепторных уровней;

в) индуцированный металлический ( $h\nu > E_d$ ), который характерен для поглощения в полупроводниках квантов видимого диапазона частот и существенен, когда вероятности процессов безизлучательной рекомбинации малы, а поглощение происходит на неравновесных свободных носителях, генерируемых этим же излучением;

г) полупроводниковый ( $h\nu > E_d$ ), который характерен для собственных (чистых) полупроводников и наблюдается при относительно невысоких уровнях плотности потока излучения и в дальнейшем переходящий в индуцированный металлический тип поглощения;

д) диэлектрический ( $h\nu \ll E_d$ ), когда поглощение происходит при непосредственном взаимодействии излучения с решеткой (оптической ветвью фотонного спектра).

При мощных коротких импульсах лазерного излучения на полупроводники за времена  $\tau \ll \tau_u$  концентрация свободных электронов повышается до  $10^{20} \dots 10^{21}$  и полупроводник по оптическим свойствам приближается к металлу. Начальный коэффициент поглощения излучения в полупроводниках варьируется в диапазоне от  $10^3$  до  $10^4 \text{ см}^{-1}$ .

Перенос энергии от поверхностного слоя, поглотившего излучение, к объему материала как в случае металлов, так и полупроводников определяется процессами теплопроводности.

Поскольку учет всех параметров и особенностей сложных процессов взаимодействия лазерного излучения с веществом практически невозможен, при их анализе часто используют понятия «критическая интенсивность» или «критическая плотность потока». Эти термины в определенной степени условны, так как связаны с понятием разрушения вещества, также имеющим условный характер. В лазерной обработке (технологии) под началом разру-



шения чаще всего подразумевают плавление поверхности, хотя необратимые изменения большинства материалов происходят и при нагреве ниже температуры плавления. Использование коротких и мощных импульсов лазерного излучения для модификации материалов достаточно перспективно ввиду реализации аномально высоких скоростей нагрева и остывания материала, поглотившего излучение.

Используя критические плотности потока  $q_c^{(i)} = (i=1...4)$ , можно классифицировать технологические процессы и рассматривать их последовательно при переходе от одной критической плотности к другой.

Критическая плотность потока, необходимая для достижения к концу импульса излучения на поверхности тела температуры плавления

$$q_c^{(1)} < \frac{\sqrt{\pi}}{2} \frac{T_{пл} \lambda}{A \sqrt{a \tau_u}} .$$

Критическая плотность потока, соответствующая достижению температуры кипения (без учета фазового перехода)

$$q_c^{(2)} = \frac{\sqrt{\pi}}{2} \frac{T_{кип} \lambda}{A \sqrt{a \tau_u}} .$$

Критическая плотность потока, выше которой процессы испарения преобладают над переносом теплоты в конденсированную среду (режим «чистого» испарения)

$$q_c^{(3)} = \frac{\gamma \Omega}{A} \left( \frac{a}{\tau_u} \right)^{1/2} .$$

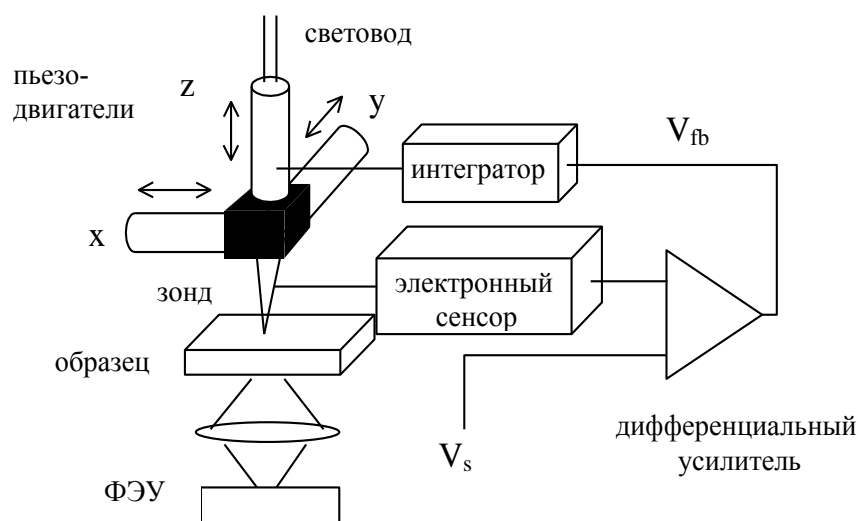
Критическая плотность потока, выше которой в глубь материала распространяется ударная волна, а над поверхностью образуется плазменный факел

$$q_c^{(4)} = \beta \gamma \Omega^{3/2} .$$

В этих формулах :  $T_{пл}$  и  $T_{кип}$  – температура соответственно плавления и кипения;  $\Omega$  – теплота испарения облучаемого вещества;  $\lambda$  – теплопроводность;  $a$  – температуропроводность;  $\tau_u$  – длительность лазерного импульса;  $\beta=0,88$  – коэффициент.

### 3.2.1. Описание и назначение сканирующего ближнепольного оптического микроскопа (СБОМ)

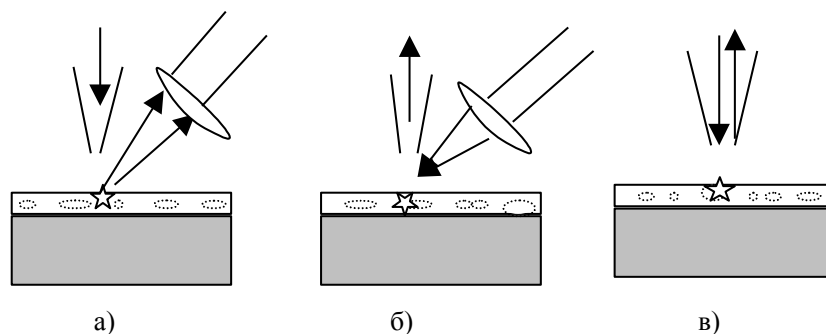
Большинство существующих ближнепольных сканирующих оптических микроскопов реализовано по схеме рис. 3.2. В качестве диафрагмы с малым диаметром отверстия в СБОМ используются оптические волокна (световоды).



*Рис. 3.2. Общая схема сканирующего ближнепольного оптического микроскопа*

В конструкцию СБОМ входят: зонд (заостренный конец световода), фотоэлектронный умножитель (фотоприемник), собирающий прошедший или отраженный от поверхности образца свет, который и формирует СБОМ-изображение, сенсор для измерения расстояния между зондом и образцом, пьезоэлектрические двигатели для перемещения зонда, электронная цепь обратной связи для поддержания расстояния зонд-образец в пределах ближнего поля и компьютер для управления процессом сканирования, получения и обработки изображений.

В стандартной конструкции сканирующие ближнепольные оптические микроскопы позволяют исследовать поглощающую или отражающую способность образца, однако с небольшими изменениями они могут быть использованы для различных целей полупроводниковой наноэлектроники. Наиболее часто используемые схемы экспериментов, используемые в СБОМ, показаны на рис. 3.3.



**Рис. 3.3.** Схемы экспериментов, используемых при исследовании поверхности полупроводников с помощью СБОМ. а) локальное облучение из острия, сбор объективом (линзой); б) фокусировка накачки линзой, сбор острием; в) возбуждение и сбор одним и тем же зондом

При исследовании относительно прозрачных образцов, а также для решения биологических и тестовых задач, обычно используется геометрия "просвечивающего СБОМ". При изучении образцов с непрозрачной для используемого света подложкой наиболее применима схема "на отражение", и ее варианты имеет смысл рассмотреть подробнее.

При возбуждении носителей излучением из острия (рис. 3.3,а) удается достичь больших плотностей накачки, и область взаимодействия образца с накачкой хорошо локализована. Недостаток этой схемы при изучении фотолюминесценции состоит в том, что возбужденные носители в чистом однородном полупроводнике за время жизни до акта рекомбинации могут удалиться от острия на длину диффузии, которая составляет несколько микрон, так что спектр излучаемого сигнала не обязательно будет соответствовать точке, находящейся непосредственно под иглой. Это существенно сужает разрешающую способность метода.

Сбор сигнала острием при облучении ближайшей области поверхности сфокусированной с помощью линз накачкой (рис. 3.3,б) более удобен при регистрации сигнала от заведомо локализованных источников (например, квантовых точек), однако эффективность возбуждения носителей вокруг острия оказывается невысокой, и кроме того, возможные особенности фокусировки излучения накачки в ближнем поле иглы и микронеровностей поверхности, учесть которые сложно, затрудняют интерпретацию наблюдаемых результатов и оставляют сомнения в их однозначности.

Наиболее просты в интерпретации изображения, полученные при локальном облучении поверхности и сборе сигнала с помощью одной и той же иглы (рис. 3.3,в). Единственным недостатком следует считать определенную потерю сигнала при двукратном прохождении света через иглу, что при наличии высокочувствительного фотоприемника не является фатальным.

СБОМ может быть применен в таких областях, как биология, медицина, генетика, для исследования объектов в их естественной среде. Это пре-

дотвращает определенные артефакты, которые могут возникать в процессе приготовления образцов или поведения образца в неестественной среде.

### 3.2.2. Атомный нанозонд с единичным фотоном

Использование фотона для обнаружения отдельного квантового объекта рассматривалось еще Гейзенбергом и фон Нойманом в 20-х годах в "gedanken experiments" в контексте квантовомеханических ограничений воздействия процесса квантового измерения на объем измерения. Развитие оптических и лазерных технологий последних лет сделало возможным проведение экспериментов по исследованию воздействия единичного фотона (локализованного в моде резонатора) на единичный атом. Световое поле в моде высокочастотного резонатора с энергией порядка одного фотона позволяет детектировать отдельные атомы и локализовывать их. Существует ряд предложений по использованию светового поля высокочастотного резонатора для детектирования атома с пространственным разрешением лучше, чем длина волны света. Положение атома в стоячей световой волне высокочастотного резонатора сильно коррелирует с фазой волны, поскольку в присутствии светового поля атом поляризуется, что в свою очередь изменяет фазу световой волны. Измерение изменения фазы световой волны при пролете атомом светового поля дает информацию об относительном (к пучности стоячей волны) положении атома в стоячей волне. Пространственное перекрытие атомного волнового пакета с модой светового поля вносит ограничение на пространственное разрешение подобного метода.

Рассмотрим возможность использования светового поля с энергией порядка единичного фотона для обнаружения единичного атома с пространственным разрешением в нанометровом диапазоне и с наносекундным временным разрешением. На рис.3.4 представлена схема такого *атомного нанозонда*.

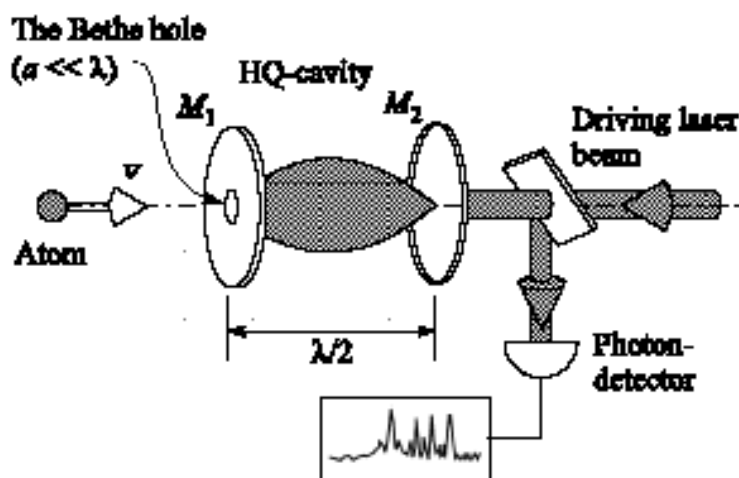


Рис. 3.4. Схема атомного нанозонда

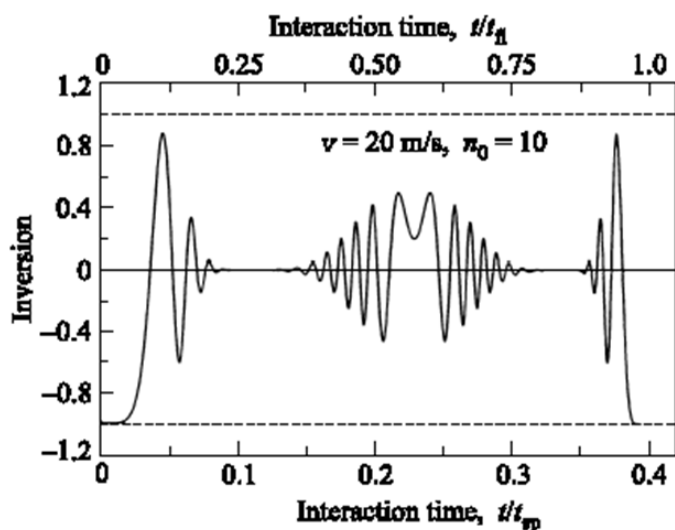
Высокодобротный оптический резонатор образован двумя зеркалами  $M_1$  и  $M_2$ . Лазерное излучение инжектируется в резонатор через зеркало  $M_2$ . Зеркало  $M_1$  имеет отверстие диаметром  $2a$ , существенно меньшим длины волны вводимого в резонатор излучения (получившего название в научной литературе - "the Bethe hole" ). Как будет показано ниже, такое отверстие не изменяет заметно добротности резонатора. Через это отверстие атом может входить внутрь резонатора и взаимодействовать со световым полем моды резонатора. Атом внутри моды резонатора изменяет его резонансные свойства, и часть лазерного излучения отражается от резонатора. Отраженное излучение регистрируется с помощью фотодетектора. Сигнал с фотодетектора несет информацию об атоме внутри моды резонатора. Пространственное разрешение такого атомного зонда определяется размерами отверстия, минимальная величина которого ограничена "размерами" самого атома и характерной длиной межатомного взаимодействия, то есть лежит в нанометровом диапазоне. Очевидно, что временное разрешение атомного нанозонда не хуже времени пролета атома через резонатор. При длине резонатора, равном  $l_r = \lambda/2$ , и тепловых атомных скоростях временное разрешение лежит в наносекундном диапазоне. Объем локализации атома составляет величину  $V = \pi a l^2 \ll \lambda^3$ .

Рис.3.5 и 3.6 иллюстрируют временную эволюцию инверсной заселенности атома и среднего числа фотонов в моде резонатора при пролете атома через резонатор для сравнительно большого числа фотонов в резонаторе  $\langle n_0 \rangle = 10$  и скорости атома  $v = 20$  м/с. Сложная временная динамика поля моды резонатора и инверсной заселенности объясняется зависимостью частоты Раби осцилляции от числа  $n$  фотонов, ведущей к известному эффекту "коллапса" и восстановлению инверсной заселенности атома, отчетливо заметному на рис. 3.5 для достаточного проявления эффекта времени взаимодействия атома при относительно малой его скорости. Появление атома в резонаторе изменяет его свойства. Пропускание резонатора  $T$  и, соответственно, отражение  $R$  становятся зависимыми от времени:

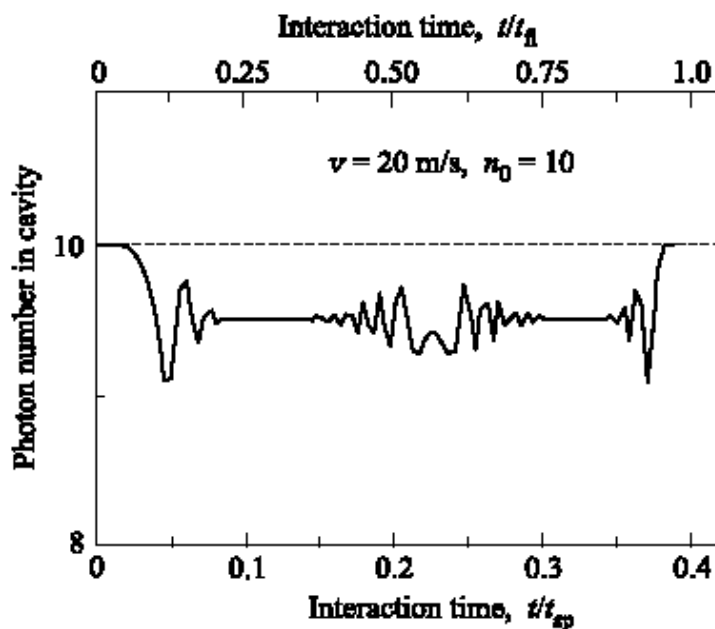
$$T(t) = \left| \frac{X(t)}{Y} \right|^2,$$

где  $X(t)$  - амплитуда поля моды резонатора,  $Y$  - амплитуда лазерного поля, накачивающего резонатор. В рассматриваемой схеме атомного нанозонда информация об атоме заключена в отраженном от резонатора излучении, которое следующим образом зависит от параметров атома и резонатора:

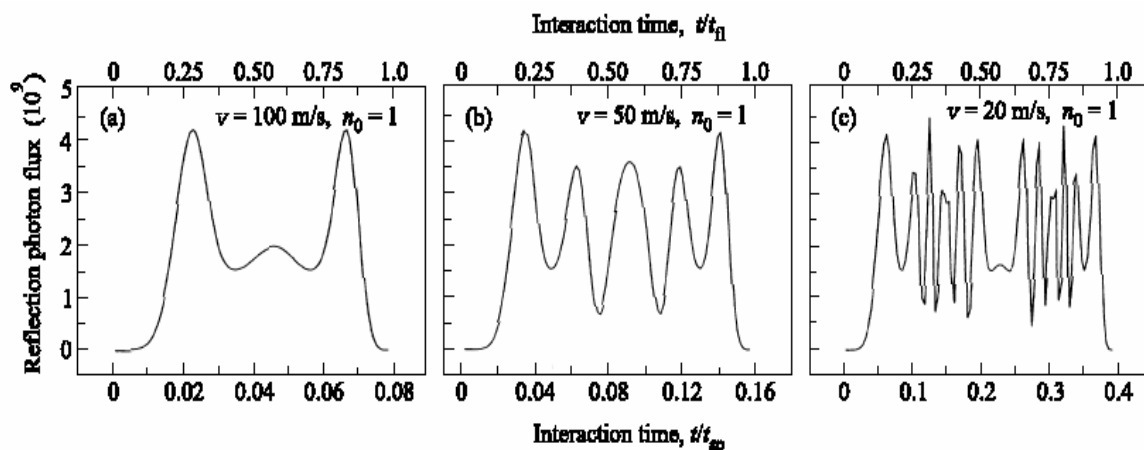
$$|Y_{refl}(t)|^2 = \left( 1 - \pi \frac{\langle n(t) \rangle}{n_{ph} \cdot VF} \right) n_{ph} c S.$$



*Рис. 3.5. Временная эволюция инверсной населенности атома при пролете атома через резонатор. Среднее число фотонов в резонаторе до влета атома  $\langle n_0 \rangle = 10$ , скорость атома  $v = 20$  м/с. На нижней шкале время приведено в единицах времени спонтанного распада верхнего уровня атома  $\tau_{sp}$ , на верхней шкале - время взаимодействия нормировано на время пролета атома через резонатор*



*Рис. 3.6. Временная эволюция среднего числа фотонов в моде резонатора при пролете атома через резонатор*



*Рис. 3.7. Зависимости отраженного от резонатора потока фотонов от времени взаимодействия атома с полем моды резонатора для различных скоростей атома: (a)  $v = 100$  м/с; (b) –  $v = 50$  м/с; (c) –  $v = 20$  м/с. Среднее число фотонов в моде резонатора  $\langle n_0 \rangle = 1$*

На рис. 3.7 показана зависимость отраженного потока фотонов за время пролета атома через резонатор, когда в моде резонатора запасен один фотон  $\langle n_0 \rangle = 1$ , а резкость резонатора  $F = 5 \cdot 10^5$  для трех значений атомных скоростей  $v = 20, 50$  и  $100$  м/с. Осцилляционный характер поведения отраженного потока фотонов определяется быстрым обменом энергией между атомом и электромагнитным полем резонатора. Зависимость интегрального сигнала отраженных фотонов определяется как

$$N_{refl} = \int_0^t |Y_{refl}(t)|^2 dt$$

для различных значений скорости атома  $v = 20, 50$  и  $100$  м/с и  $\langle n_0 \rangle = 1$ . Видно, что даже при достаточно большой скорости атома ( $v = 100$  м/с) отраженное число фотонов достигает величины 5 фотонов. При этом время составляет величину  $t_{per} = -0,08\tau_{sp}$ .

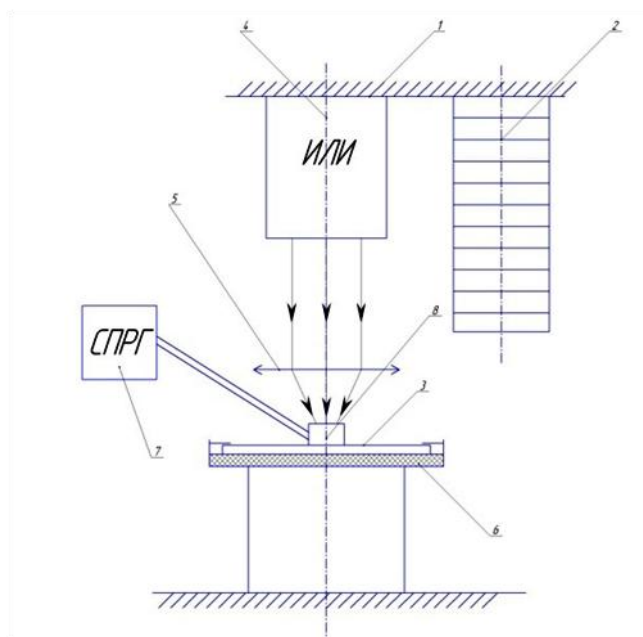
### 3.2.3. Устройство для получения нанометровых отверстий на подложке

В основу разработки положена задача получить нанометровые отверстия на подложке.

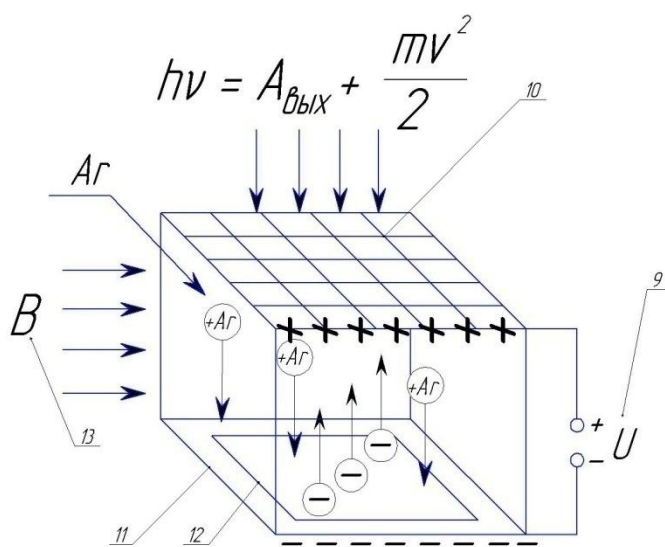
Устройство для получения нанометровых отверстий (рис. 3.8) содержит неподвижное основание 1 с установленным на нем источником лазерного излучения 4, на пьезоприводе 2 жестко закреплена фокусирующая линза 5, а технологический блок 3 установлен на трехкоординатном столе 6, устрой-



ство дополнительно снабжено системой подачи инертного газа 7 и технологическим блоком 8.



**Рис. 3.8.** Устройство для получения нанометровых отверстий на подложке



**Рис. 3.9.** Технологический блок

Технологический блок (рис. 3.9) содержит вакуумную камеру, внутри которой расположены анод и катод, связанные с источником питания 9. Анод выполнен в виде сетки 10, на верхней стенке вакуумной камеры установлена линза с возможностью подачи лазерного потока через сетку-анод на

катод 11, на котором установлена полочка 12, внутри вакуумной камеры расположен электромагнит 13, с возможностью подачи магнитного потока с индукцией  $B$  в анодно-катодный зазор перпендикулярно силовым линиям напряженности  $E$  электрического поля.

Устройство для получения нанометровых отверстий на подложке работает следующим образом

В анодно-катодный зазор подается поток лазерного излучения, что порождает бомбардировку молекул рабочего газа, в следствие чего формируется поток ионов, направленный на сверхтонкую пластину, закрепленную на катоде. Таким образом происходит образование отверстий в сверхтонкой пластине, величиной с ион.

Применение устройства позволяет получать нанометровые отверстия на подложке в производстве изделий электронной техники.

### Заключение

В настоящее время происходят революционные изменения в сфере высоких технологий: электронике, информационных технологиях, микромеханике и других областях человеческой деятельности, связанных с фундаментальными и прикладными исследованиями, конструированием и практическим использованием структур, материалов и устройств, элементы которых имеют размеры менее 100 нм, а также с развитием технологий их изготовления (нанотехнологии) и методов диагностики.

В настоящей концепции под обобщенным понятием наноматериалы понимаются структуры и материалы с качественно новыми свойствами, возникающими благодаря уменьшению в одном или нескольких измерениях размеров структурных элементов до нанометрового масштаба, а технологии их получения и, в определенных случаях, после дующая работа с ними — нанотехнологиями.

### Литература

1. М. Борн, Э. Вольф. Основы оптики. М.: Наука, 1973.
2. В.Г. Дюков, Ю.А. Кудряев. Растровая оптическая микроскопия. М.: Наука, 1992.
3. Г.С. Жданов, М.Н. Либенсон, С.А. Марциновский. Оптика за дифракционным пределом: Принципы, результаты, проблемы. УФН, 1998, №7, с.801.
4. M.A. Paesler, P.J. Moyer. Near-Field Optics: Theory, Instrumentation and Applications. Wiley, 1996.
5. М. Адамс. Введение в теорию оптических волноводов. М.: Мир, 1984.
6. Л.Д. Ландау, Е.М. Лифшиц. Теоретическая физика. Т.П. Теория поля. М.: Физматгиз, 1960.
7. H.A. Bethe. Theory of Diffraction by Small Holes. *Phys. Rev.* 66, 163 (1944).

8. U.Durig, D.W. Rohl, F. Rohner. Near-Field optical-scanning microscopy. *J. Appl Phys.* 59, 3318 (1986).
9. K. Karrai, R.D. Grober. Piezoelectric tip-sample distance control for near field optical microscopes. *Appl. Phys. Lett.* 66, 1842 (1995).
10. Энциклопедия машиностроение. Том III-8 Технологии, оборудование и системы управления в электронном машиностроении, под ред. Ю.В. Панфилова, П.Н. Белянина. М.: Машиностроение, 2000.
11. О.С. Моряков. Электронная обработка. М.: Высшая школа, 1990.
12. К.А. Валиев. Физика субмикронной литографии. М.: Наука, 1990.
13. Физика и технология источников ионов. Сб. под ред. Я. Брауна. М.: Мир, 1998.
14. С. Чу, К.Н. Каэн-Тануджи, У.Д. Филипс. УФН 169, 274 (1999).
15. G. Timp, R.E. Behringer. D.M. Tennat et al. *Phys. Rev. Lett.* 69, 1636 (1992).
16. F.Lison, H.-J. Adams, D. Haubrich et al. *Appl. Phys.* B65, 419 (1997).
17. Н.Б. Делоне. Взаимодействие лазерного излучения с веществом. М.: Наука, 1989.
18. P. Peuser, F. Ames, J. Bohn et al. *Appl. Phys.* B38, 249 (1985).
19. V.K. Mago, B. Lal, A.K. Ray et al. *J. Phys.* B20, 6021 (1987).
20. H.D.V. Bohm, W. Michaelis, and C. Weitkamp. *Opt. Commun.* 26, 177 (1978).
21. Т.А. Ворончев, В.Д. Соболев. Физические основы электровакуумной техники. М.: Высшая школа, 1967.
22. А.П. Казанцев, Г.И. Сурдутович, В.П. Яковлев. Механическое воздействие света на атомы. М.: Наука, 1991.
23. T. Arisawa, Y. Mariyama, Y. Suzuki, and K. Shiba. *Appl. Phys.* B28, 73 (1982).
24. Е.Н.Ивашов и др. Устройство для получения нанометровых отверстий на подложке. Патент РФ на ПМ № 42696.- Опубликовано 10.12.2004. Бюл.№34.
25. Л.Л.Андреева, Е.Н.Ивашов, С.В.Степанчиков. Способ получения электронного луча. Патент РФ на изобретение №2208262.- Опубл.10.07.2003. Бюл. №19.

## **Глава 4. Формирование магнитоориентированных объектов методом туннельно-зондовой нанотехнологии**

### **Введение**

В настоящее время происходят революционные изменения в сфере высоких технологий: электронике, информационных технологиях, микромеханике и других областях науки и техники, связанных с фундаментальными и прикладными исследованиями, конструированием и практическим использованием структур, материалов и устройств, элементы которых имеют размеры менее 100 нм, а также с развитием технологий их изготовления (нанотехнологии) и методом диагностики.

Нанотехнология – это междисциплинарная, образующая технология, позволяющая «технологично» (воспроизводимо, по описанным процедурам) производить исследования, манипуляцию и обработку вещества в диапазоне размеров и с допусками 0,1–100 нм. Нанотехнология объединила в себе все, относящееся к технологии получения устройств, исследованию материалов, созданию материалов с заданными свойствами, созданию отдельных элементов и функциональных сред.

Особым разделом нанотехнологии является нанoeлектроника, которая реализуется на миниатюрных электронных нанообъектах, то есть структурах с линейными размерами порядка нанометра ( $10^{-9}$  м). Формирование таких наноструктур стало возможным с появлением туннельной микроскопии. Туннельный микроскоп – это прибор, который позволяет исследовать поверхность материалов на микронном и даже атомарном уровне. Однако помимо исследования свойств поверхности для нанотехнологии характерен индивидуальный подход, при котором с помощью внешнего управления производится манипуляция отдельными атомами или молекулами, что позволяет создавать абсолютно любые объекты. Более того, с помощью нанотехнологии можно проводить модификацию нанообъектов на атомном уровне. Локальное, с точностью до нанометров, управление атомно-молекулярными реакциями – является важнейшей задачей нанотехнологии.

Традиционная микроэлектроника сегодня уже практически не в состоянии сокращать размеры электронных микроэлементов. Существующие технологии производства приблизились к теоретическим пределам своих возможностей, таким образом, для изготовления еще более компактных микроэлементов необходимо внедрять новые технологии и устройства. Сегодня интенсивно изучаются способы формирования поверхностей модифицированных нанотехнологией, а также исследуются свойства получаемых этими способами материалов. Результаты исследований не заставляют себя ждать – появляются новые перспективные искусственно созданные материалы, на основе которых планируется создавать электронные устройства нанометрового (молекулярного) размера. Именно эти устройства, работая преимущест-

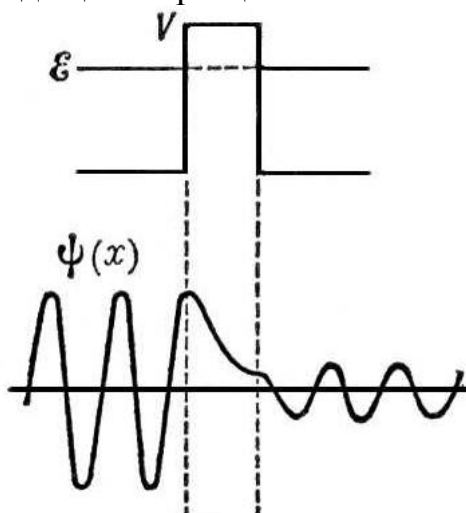
венно на принципах квантовой механики, должны коренным образом изменить жизнь человечества – позволить достичь теоретических пределов в электронике.

#### 4.1. Туннельно-зондовая нанотехнология

Одним из мощных современных методов исследования морфологии и локальных свойств поверхности твердого тела с высоким пространственным разрешением является сканирующая зондовая микроскопия (СЗМ). За последние 10 лет СЗМ превратилась из экзотической методики, доступной лишь ограниченному числу исследовательских групп, в широко распространенный и успешно применяемый инструмент для исследования свойств поверхности. В настоящее время практически ни одно исследование в области физики поверхности и тонкопленочных технологий не обходится без применения методов СЗМ. Развитие СЗМ послужило также основой для развития новых методов в нанотехнологии – технологии создания структур с нанометровыми масштабами.

Сканирующий туннельный микроскоп (СТМ) – первый из семейства зондовых микроскопов – был изобретен сотрудниками швейцарского филиала фирмы ИВМ учеными Гердом Биннигом и Генрихом Рорером [1, 2] в 1981 г. В своих работах они показали, что это достаточно простой и весьма эффективный способ исследования поверхности с пространственным разрешением вплоть до атомарного. Настоящее признание данная методика получила после визуализации атомарной структуры поверхности ряда материалов и, в частности, реконструированной поверхности кремния. В 1986 году за создание туннельного микроскопа Биннигу и Рореру была присуждена Нобелевская премия по физике.

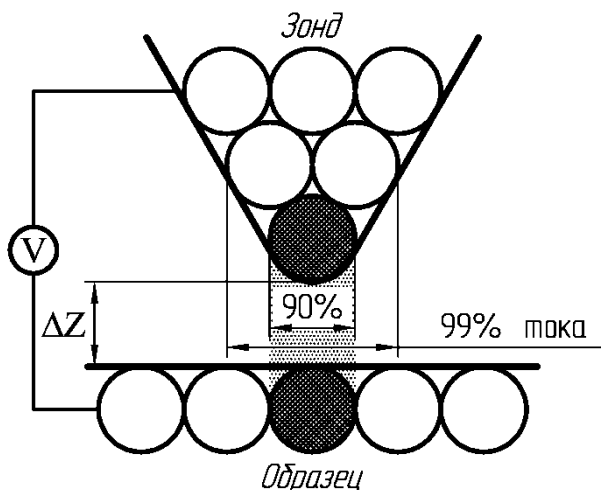
Сканирующий туннельный микроскоп основан на явлении туннелирования электронов через узкий потенциальный барьер (рис. 4.1) между металлическим зондом и проводящим образцом во внешнем электрическом поле.



*Рис. 4.1. Экспоненциальный спад волновой функции электрона внутри потенциального барьера*

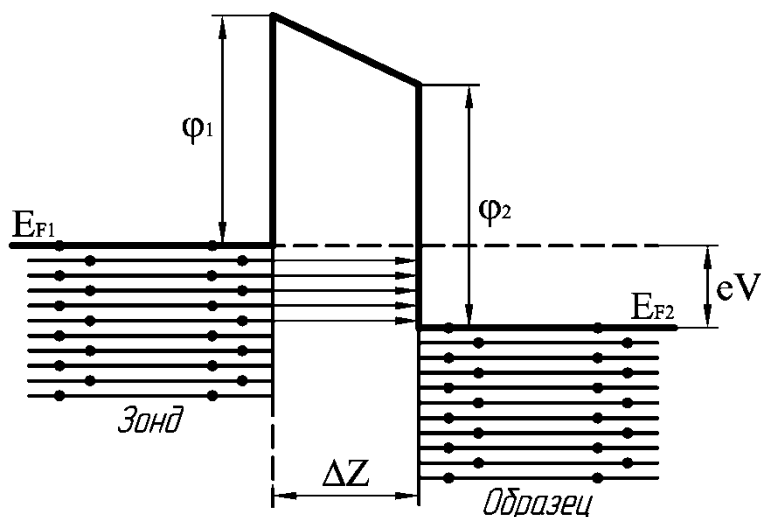
Применение эффекта туннелирования в зазоре между зондом особой конструкции и поверхностью подложки позволяет достигать высокой пространственной разрешающей способности при исследовании поверхности различных проводящих материалов [3].

Принцип работы сканирующего туннельного микроскопа заключается в приложении небольшого напряжения ( $0,01 - 10$  В) между электропроводящим острием и образцом, затем производится регистрация туннельного тока в зазоре (рис. 4.2).



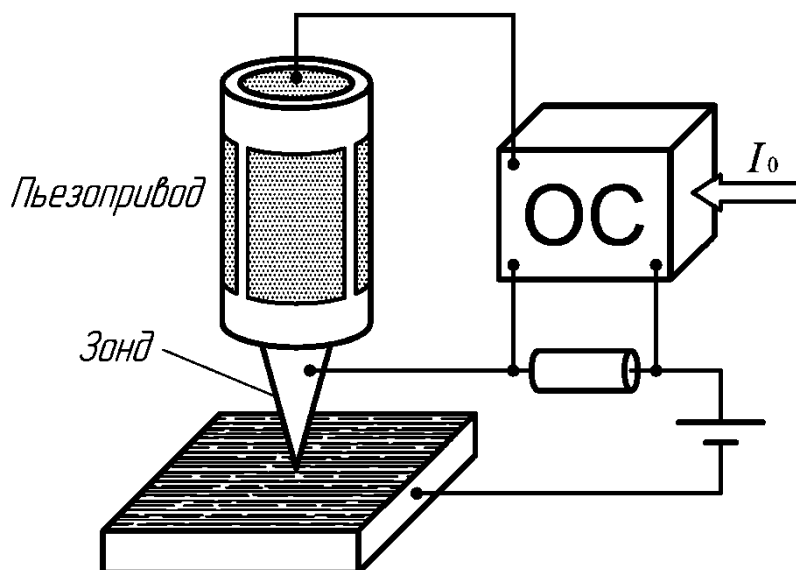
**Рис. 4.2.** Распределение туннельного тока в зазоре

Туннельный ток зависит от свойств и конфигурации атомов на исследуемой поверхности образца и имеет место быть благодаря туннельному эффекту, то есть квантовому переходу электрона через область, запрещенную классической механикой. Здесь этой областью является зазор между кончиком металлической иглы и ближайшей точкой на поверхности образца (рис.4.3).



**Рис. 4.3.** Энергетическая диаграмма туннельного контакта двух металлов:  $E_{F1}$ ,  $E_{F2}$  – энергия уровня Ферми;  $\phi_1$  и  $\phi_2$  – работа выхода электронов из материала зонда и образца соответственно;  $eV$  – напряжение смещения;  $\Delta Z$  – ширина барьера

Обычно используют режим, в котором туннельный ток поддерживается постоянным (1–10 нА) за счет цепей обратной связи, меняющей положение зонда вдоль координаты Z, посредством прецизионного привода движения. Именно сигнал в цепи обратной связи используют для построения изображения поверхности [4, 5].



*Рис. 4.4. Упрощенная схема СТМ измерения рельефа поверхности образца*

Аналитико-технологическим объектом в СТМ может быть только проводящий образец или, хотя бы образец, имеющий проводимость на уровне полупроводникового кристалла.

Разрешающая способность СТМ определяется, главным образом, точностью используемого привода наноперемещений (пьезоэлектрический или магнитострикционный), а так же типом зонда, его качеством и изношенностью. По X и Y разрешение составляет около 1 Å, по Z достигает 0,1 Å и выше [6].

Спустя некоторое время после изобретения СТМ выяснилось еще и то, что с помощью острых проводящих твердотельных зондов можно не только исследовать, но и изменять поверхность с атомным разрешением. Воздействие осуществляется токовыми импульсами и основано на электростатическом либо силовом взаимодействии одиночного выступающего атома зонда с атомом подложки. Однако технология такого формирования столкнулась с существенными проблемами ее практического применения, связанными с огромными сложностями, возникающими при производстве проводящих зондов с острием, состоящим из небольшого кластера атомов, не говоря уже о моноатомном острие. Сказывается и несовершенство сканеров, и сложность возбуждения специфического токового либо силового импульса. Все это пока не позволяет формировать на подложке хоть сколько-нибудь сложные наноструктуры. Сегодня эту задачу пытаются решить с помощью активно развивающейся туннельно-зондовой нанотехнологии (ТЗН) [7]. Туннельно-зондовой нанотехнологией называют сканирующую туннельную микроскопию при рассмотрении ее как метода формирования микро- и нанообъектов.

Таким образом, ТЗН метод имеет следующие преимущества:

1) Легко и быстро достигаются атомные разрешения, за счет того, что в рабочих режимах СТМ туннельный зазор составляет единицы  $\text{Å}$  (обычно  $2-10 \text{ Å}$ ), следовательно, даже на воздухе при атмосферном давлении его, в первом приближении, можно считать вакуумированным (применимо только к высокостабильным материалам, которые не подвержены окислению);

2) Можно получить неискаженное изображение микроскопически плоской поверхности образца на любом его участке;

3) Можно производить качественные и количественные исследования не только топологии поверхности, но и многих других характеристик атомов на ней, другими словами СТМ может исполнять роль спектроскопа;

4) С помощью СТМ можно снимать вольт-амперные характеристики туннельного контакта в различных точках поверхности, что позволяет судить о локальной проводимости образца и изучать особенности локальной плотности состояний в энергетическом спектре электронов;

5) Небольшая сила тока и энергия туннелирующих электронов (примерно  $1 \text{ эВ}$ ) обеспечивает высокую сохранность структуры образца в процессе изучения;

6) Есть возможность производить модификацию поверхности образца на атомном уровне;

7) Можно проводить различные нанотехнологические процессы. В СТМ это легко реализуемо благодаря конструкции СТМ, которая легко может быть дополнена необходимым оборудованием, что скажем, дополнительно создавало бы электромагнитное поле.

Однако ТЗН метод имеет и недостатки:

1) Аналитико-технологический объект должен быть обязательно проводящим, хотя бы на уровне полупроводникового кристалла, а зонд выполнен в виде специального металлического игольчатого электрода;

2) Истинно атомарное разрешение в СТМ достигается лишь при низких температурах (несколько десятков К) и высоком или сверхвысоком вакууме. Тепловой шум, адсорбирование влаги, втягивание молекул воздуха в туннельный промежуток и другие эффекты существенно понижают разрешение при работе в обычных условиях.

Вслед за сканирующим туннельным микроскопом в течение короткого времени были созданы атомно-силовой микроскоп, магнитно-силовой микроскоп, электросиловой микроскоп, сканирующий ближнепольный оптический микроскоп и многие другие приборы, имеющие сходные принципы работы, а также общее название – сканирующие зондовые микроскопы.

Стоит отметить, что возможность изменения поверхности была реализована и другими видами СЗМ. Однако в этой работе СТМ способу исследования и модификации поверхностей материалов уделяется больше внимания, чем остальным. Это связано с тем, что в других методах СЗМ значительно ограничено проведение различных нанотехнологических процессов из-за их специфической конструкции, которую достаточно сложно дополнить необ-



ходимым оборудованием, в частности, магнитной системой, создающей однородное магнитное поле, пронизывающее подложку.

## 4.2. Магнитные свойства наноструктур

Магнитные свойства наноструктур обладают большим разнообразием и значительно отличаются от массивного материала. Основной вклад здесь вносят размерные эффекты, влияние поверхности, образующих наноструктуру кластеров, межкластерные взаимодействия и межкластерная организация. К числу наиболее характерных и впечатляющих свойств нанокластеров и наноструктур следует отнести прежде всего магнитную однодоменность наноструктур вплоть до 20 нм и суперпарамагнетизм кластеров размером 1–10 нм.

### 4.2.1. Нанокристаллические материалы

Физическое материаловедение успешно занимается детальным изучением структуры и свойств макроскопических объемов материалов, содержащих до  $10^{21}$ – $10^{25}$  атомов. Как правило, такие материалы состоят из макроскопических (5–50 мкм) зерен или являются монокристаллами. Однако тепловые, магнитные, сверхпроводящие, полупроводниковые, механические, коррозионные и другие свойства для материалов, находящихся в макроскопическом объеме, достигли своих максимальных значений и вряд ли можно добиться улучшения каких-то характеристик за счет более тщательной технологии или изменения концентрации составляющих элементов.

Одним из наиболее перспективных научных направлений является создание и изучение структуры и свойств материала, который сконденсирован из кластеров, малых кристалликов, имеющих объем, равный  $10^2$ – $10^4$  атомов. Материал с такой структурой можно считать наноструктурным композитом. Наноструктурное состояние может быть реализовано в материалах с дальним порядком расположения атомов, путем создания нанокристаллической структуры, где размер кристаллических зерен не превышает 10–15 нм.

Достаточно давно известно, что свойства малых частиц вещества в значительной степени отличаются от свойств массивного материала. Перспективность нанокристаллических материалов (НКМ) заключается именно в том, что уменьшение размера кристалликов в материале может приводить к существенному изменению их свойств [8]. Такие изменения появляются, когда средний размер кристаллических зерен не превышает 100 нм, и наиболее заметны при размере зерен менее 10 нм.

Малый размер зерна приводит к большим изменениям структуры твердого тела. Если считать в грубом приближении форму зерна сферической с диаметром  $d$  и толщиной границы раздела  $h$ , то доля межзеренных границ раздела в общем объеме зерна равна:

$$\frac{\Delta V}{V} = \left( \frac{\pi}{6} d^3 - \frac{\pi}{6} (d - 2h)^3 \right) \left( \frac{\pi}{6} d^3 \right)^{-1} \cong \frac{6h}{d}.$$

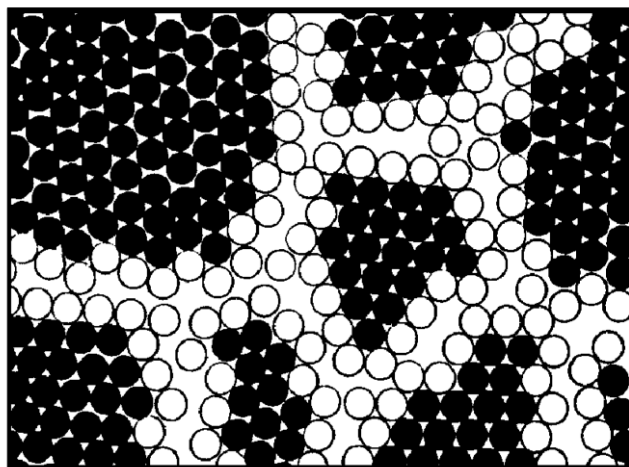
При толщине границы  $h$ , равной 3–4 атомным монослоям (0,5–1,5 нм), на поверхностный слой приходится до 50 % всего вещества. Кроме того, существенно изменяется и структурное состояние самого зерна. Дефекты, типичные для крупных зерен, часто не могут удержаться внутри зерна и выходят на поверхность, делая структуру зерна бездефектной. С другой стороны, межзеренная граница является нестабильной структурой и поэтому НКМ отличаются высокой диффузионной подвижностью атомов, которая на 5–6 порядков выше, чем в обычных поликристаллических материалах.

Учитывая вышесказанное, можно считать, что уменьшение размера зерна является эффективным средством изменения свойств твердого тела. Таким образом, НКМ – особый тип конденсированного состояния вещества, в котором необычные свойства обусловлены как особенностями малых кристаллитов так и коллективным поведением, зависящим от характера взаимодействия между наночастицами.

Типичные НКМ представляют собой однофазные или многофазные поликристаллы с размером зерна от 1 до 25 нм. В таких материалах от 2 до 50 процентов может приходиться на межзеренные или межфазные границы, что значительно больше, чем у поликристаллов ( $10^{-9}$ – $10^{-4}$  всего объема). Глейтер Ш. сделал предположение о том, что изменения физических параметров в наноматериалах связаны с большим объемом межкристаллитной фазы, которая имеет более низкую плотность по сравнению с кристаллической фазой. Некоторые ученые считают [9-11], что структурное состояние атомов, составляющих межзеренные или межфазные границы, отличается от пространственного расположения атомов как кристаллических, так и аморфных твердых тел. Разупорядоченная структура границ зерен или межфазных границ, для которой даже ближний порядок не является обязательным, аналогична «газообразному» беспорядку расположения атомов в пространстве.

Схематическое (модель тяжелых сфер) расположение атомов в двумерном измерении НКМ, взятое из работы [9], представлено на рис. 4.5.

В [10] показано, что НКМ может быть разделен на две структурные компоненты: кристаллическую, которая включает в себя атомы, локализованные внутри кристаллитов, и межкристаллитную или межфазную, формируемую из всех атомов, расположенных на границах кристаллитов. В кристаллитах имеется дальний порядок расположения атомов, а межкристаллитная компонента не является структурой с ближним порядком расположения атомов. Расположение в пространстве «граничных» атомов отличается от расположения «решеточных» атомов.



*Рис. 4.5. Схематическое двумерное представление структуры НКМ. Зачерненные и светлые кружочки – атомы кристаллитов и междоменных границ соответственно*

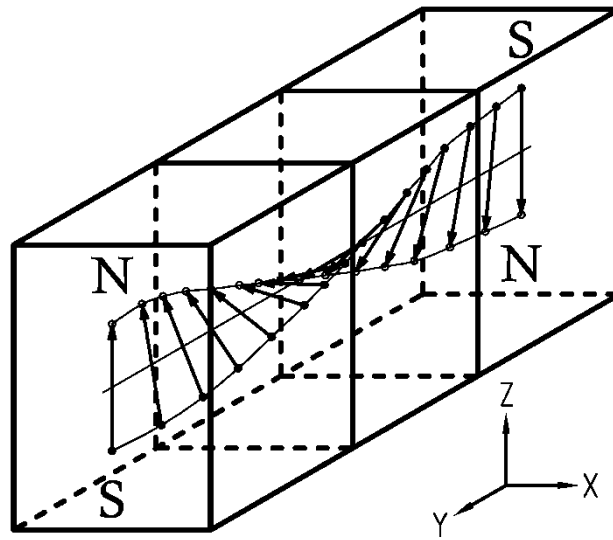
Атомная структура границ не является простой и зависит от многих параметров, и в первую очередь от ориентационных соотношений между кристаллитами. Границы в НКМ составляют порядка  $10^{19}$  в  $1 \text{ см}^3$ . Следовательно, междоменная компонента представляет собой участки беспорядочно расположенных атомов в пространстве, ограниченных различно ориентированными кристалликами.

Предполагается [12], что границы между зернами имеют ту же плотность, что и само зерно, а более низкая плотность наноматериалов обусловлена пористостью, которая возникает на стыках трех и более зерен.

Магнитные кластеры представляют собой звено, соединяющее микроскопический магнетизм индивидуальных атомов и макроскопический магнетизм кристаллических и аморфных тел. Поэтому их называют иногда мезоскопическими магнитами. Термины “мезоскопический” и “мезоскопика” происходят от греческого слова *mesos* – средний, промежуточный. Принципиальное значение изучения мезоскопических элементов состоит в том, что она дает возможность понять особенности перехода от микроскопических закономерностей к макроскопическим; эти исследования важны также для изучения и установления предельных свойств мезообъектов и, значит, определения достижимых границ миниатюризации электронных приборов и элементов памяти.

#### **4.2.2. Однодоменные ферромагнитные частицы**

Между доменами существуют переходные слои конечной толщины, в которых вектор намагниченности  $\mathbf{J}_s$  непрерывно меняет свое направление (рис.4.6).



*Рис. 4.6. Блоховская стенка между доменами с противоположной ориентацией магнитных моментов*

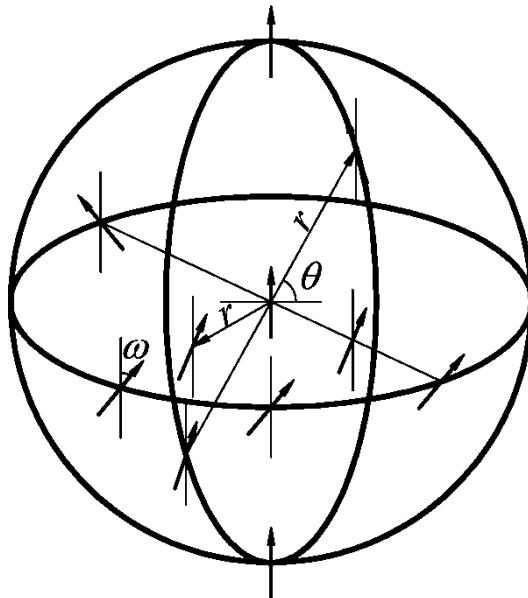
На образование этих слоев затрачивается положительная энергия, но она меньше энергии поля  $H_0$ , которая возникла бы в отсутствие доменов. При уменьшении размеров ферромагнетика до некоторой критической величины его разбиение на домены может стать энергетически невыгодным. В этом случае такие мелкие ферромагнитные частицы оказываются однородно намагниченными, то есть происходит образование так называемой однодоменной структуры, при которой каждая ферромагнитная частица представляет собой один домен.

Однодоменные ферромагнитные частицы – это малые частицы ферромагнитного вещества, являющиеся одиночными доменами. При уменьшении размеров многодоменных частиц они могут стать однодоменными, если их свободная энергия в этом состоянии меньше, чем в многодоменном. Критический размер  $D_{кр}$ , ниже которого частица обладает однодоменной структурой, зависит в первую очередь от значения намагниченности насыщения  $J_S$  и коэффициента магнитной анизотропии  $K$ . Величина  $D_{кр}$  возрастает с увеличением  $K$  и уменьшением  $J_S$ . Для большинства ферромагнетиков  $D_{кр} \sim 10^{-4} - 10^{-6}$  см. Однодоменные ферромагнитные частицы характеризуются некоторыми особенностями магнитных свойств. Их остаточная намагниченность  $J_r = J_S$ . Перемагничивание однодоменной ферромагнитной частицы происходит необратимым вращением вектора намагниченности  $J_S$ . Если однодоменные ферромагнитные частицы обладают высокой магнитной анизотропией (например, за счет анизотропии формы частицы или магнитной кристаллографической анизотропии), то процесс вращения  $J_S$  затруднен и коэрцитивная сила  $H_C$  велика. В магнитно-одноосных однодоменных ферромагнитных частицах  $J_r$  и  $H_C$  имеют максимальное значение вдоль оси легкого намагничивания, так как поворот вектора  $J_S$  из этого направления

требует наибольшей работы магнитного поля.

Для намагничивания до насыщения ферромагнитного образца, состоящего из суспензии однодоменных ферромагнитных частиц, вкрапленных в диамагнитную матрицу, требуется поле значительно меньшей напряженности, чем в случае многодоменных частиц. Создание в подобном конгломерате однодоменных ферромагнитных частиц магнитной текстуры, при которой оси легкого намагничивания частицы выстроены вдоль одного направления (оси текстуры), приводит к увеличению  $J_r$  и  $H_C$  вдоль этой оси. С увеличением концентрации однодоменных ферромагнитных частиц в конгломерате растет значение остаточной намагниченности  $J_r$  [13, 14].

Следует отметить, что повышенный интерес к ферромагнитным нанокристаллическим материалам связан с тем, что магнитные кластеры в них находятся в однодоменном состоянии. Существование такого состояния можно пояснить с помощью модели распределения магнитного момента в кластере сферической формы (рис. 4.7).



*Рис. 4.7. Распределение магнитного момента в кластере сферической формы*

Полная магнитная энергия одного кластера складывается из основных трех членов:

$$E_m = E_{\text{обм}} + E_{\text{дип}} + E_{\text{ан}},$$

где  $E_{\text{обм}}$  – обменная энергия;  $E_{\text{дип}}$  – энергия магнитного дипольного взаимодействия;  $E_{\text{ан}}$  – энергия анизотропии.

Обменное взаимодействие выстраивает спины атомов параллельно друг другу. Однако с увеличением размера кластера возрастает потеря энергии за счет магнитного дипольного взаимодействия:

$$E_{\text{дип}} \sim \frac{1}{8\pi} \int H^2 dV,$$

где  $H$  – напряженность магнитного поля.

Отсюда следует, что энергия размагничивания пропорциональна объему кластера  $V$ . Как известно энергетически выгодным является такое распределение спинов, которое обеспечивает минимум магнитной энергии. Это приводит для массивного вещества к разбиению на домены с противоположно направленными суммарными магнитными моментами. Для кластера с размерами менее некоторого критического радиуса:

$$R_{cr} = \frac{2,085\sqrt{3I}}{4\pi},$$

где  $I$  – обменный интеграл, характерно состояние с однородным намагничиванием, т.е. постоянным по всему объему магнитным моментом  $M(r, \theta, \omega) = M_0$ . При  $R > R_{cr}$  нарушение однородности по объему кластера происходит при непрерывном изменении направления магнитного момента на угол  $\omega$  с первоначально выбранной осью по кругу. Угол наклона является функцией  $r$  и  $\theta$ , а от переменной  $\omega$  не зависит [15].

### 4.2.3. Суперпарамагнетизм

Суперпарамагнетизм – это квазипарамагнитное поведение веществ, состоящих из очень малых ферромагнитных частиц (кластеров), слабо взаимодействующих друг с другом. Это обусловлено тем, что очень малые частицы (с линейными размерами примерно 10–100 Å и меньше) ниже точки Кюри или Нееля теряют свое однодоменное ферромагнитное состояние, то есть такое состояние, при котором по всей частице намагниченность однородна. Это можно объяснить тем, что направление намагниченности таких частиц вследствие тепловых флуктуаций хаотически изменяется подобно тому, как меняется под воздействием теплового движения направление магнитных моментов атомов или ионов в парамагнетике. Суперпарамагнетизм можно уподобить броуновскому вращению или движению магнитного момента кластера как целого, которое было впервые отмечено Неелем [16], а подобное состояние названо Бином [17] термином «суперпарамагнетизм».

Вследствие суперпарамагнетизма система малых частиц ведет себя в магнитных полях и при изменении температуры подобно парамагнитному газу из  $N$  атомов ( $N$  – число однодоменных частиц, каждая из которых обладает магнитным моментом  $M$ ). Для такой системы выполняется закон Кюри в слабых магнитных полях и применима формула Ланжевена для намагниченности в области магнитного насыщения.

Намагниченность суперпарамагнетиков может быть во много раз больше намагниченности обычных парамагнетиков. Чтобы векторы намагниченности частиц хаотически меняли свою пространственную ориентацию, энергия теплового движения  $kT$  должна быть больше или порядка энергии магнитной анизотропии частицы  $KV$ , где  $K$  – константа анизотропии,  $V$  – объем частицы. Для этого при температурах  $\sim 100$  К размер частиц должен быть меньше 100 Å. Типичными представителями суперпарамагнитных систем являются малые частицы  $Co$ , выделяющиеся при распаде твердого рас-

твора Cu – Co (2% Co), мелкие выделения Fe в  $\beta$ -латуни(0,1% Fe), Cu в Mn, Ni в Au, а также некоторые антиферромагнитные окислы.

В табл. 4.1 приведены значения критических размеров и температур блокирования характерных для типичных ферромагнетиков (железа и никеля) [13, 18].

**Таблица 4.1**

Металл	Средний размер частицы $d$ , нм	Температура суперпарамагнетизма ( $T_\delta$ ), К	Теоретический критический радиус домена, нм	Экспериментальный критический радиус домена, нм
Fe	9,0	310	9,8	10,6
	6,5	170		
	6,0	150		
Ni	11,0	300	21,2	16,0
	7,6	200		
	5,5	150		

### 4.3. Магнитоориентированные объекты

Нанообъекты представляют собой системы, где движение частиц ограничено в одном направлении (тонкие пленки, сверхрешетки), двух направлениях (нано-проводами) или во всех трех направлениях (квантовые точки). Эти системы демонстрируют целый ряд новых фундаментальных свойств благодаря как пониженной размерности, так и мезоскопическому эффекту, причем путь от открытия новых свойств до промышленного использования здесь оказывается рекордно коротким.

В данной работе большее внимание уделено системам, где движение частиц ограничено в одном направлении, а именно тонким пленкам. Рассмотрим ферромагнитную нанокристаллическую тонкую пленку, которая представляет собой конгломерат магнитных нанокластеров, находящихся в однодоменном состоянии. Однако назвать такую пленку магнитоориентированным объектом некорректно, так как ее нанокластеры не имеют какой-либо упорядоченности в направлении магнитных моментов. Таким образом, интересующий нас магнитоориентированный объект есть ничто иное, как ферромагнитный нанокристаллический материал, у которого магнитная ориентация всех кластеров параллельна и совпадает с направлением их легких осей. Магнитные нанокластеры особенно интересны, из-за наличия у них внутренней дополнительной степени свободы – магнитного момента, который придает большое разнообразие их свойствам и позволяет управлять их состоянием.

С точки зрения приложений магнитоориентированные объекты интересны, тем, что их состоянием удается воспроизводимым образом управлять. Этого можно добиться при помощи внешнего магнитного, электрического

либо силового воздействия.

Использование магнитоориентированных объектов особенно востребовано в области прецизионных измерений, а также сверхплотного хранения информации. Магнитоориентированные объекты, а точнее структуры из магнитных кластеров целесообразно использовать в качестве основы для активной области сверхчувствительных датчиков, а также в качестве пластин перспективных энергонезависимых магнитных запоминающих устройств. Рассмотрим преимущества, которыми обладают устройства, построенные на базе магнитоориентированных структур.

#### **4.3.1. Применение магнитоориентированных объектов в датчиках**

Магнитные наноструктуры находят свое широкое применение в миниатюрных сверхчувствительных датчиках информационно-измерительных систем. Датчик, выполненный из наноструктур с определенной магнитной ориентацией, гораздо чувствительнее к внешним воздействиям, чем датчик, в котором используется спонтанно ориентированная структура этого же вещества. Таким образом, анализируя характер восприятия наноструктурой производимого воздействия, можно получить гораздо более точные сведения об этом воздействии.

Каким же образом магнитная структура активной области сверхчувствительного датчика эффективно регистрирует то или иное воздействие? Дело в том, что от магнитного момента вещества зависят многие свойства ферромагнетика, например электрическое сопротивление, теплопроводность, механическая деформация, обратная зависимость тоже имеет место быть. Таким образом, связь ферромагнетизма со многими немагнитными свойствами позволяет по данным измерений магнитного момента вещества получить информацию о различных внешних воздействиях. Регистрировать намагниченность активной области датчика позволяет, например, эффект Баркгаузена [19, 20]. С его помощью по току, индуцируемому в измерительной катушке, надетой на ферромагнитный датчик, можно судить о намагниченности ферромагнитного вещества при непрерывном изменении внешних условий. Импульсы тока в катушке, обусловленные скачкообразным изменением намагниченности активной области датчика, позволяют с очень высокой точностью контролировать тот или иной параметр. К таким параметрам можно отнести магнитное и электрическое поле, температуру, вибрацию, разнообразные давления и напряжения, шумы и даже косвенно атомную магнитную структуру. Благодаря тому, что магнитные наночастицы обладают сверхмалыми линейными размерами, они могут использоваться и в сенсорных панелях фото- и видеотехники. Подбирая необходимые материалы можно добиться высокой стабильности и механической прочности датчиков, что позволяет использовать их и как приемники сигнала от высокоэнергетических источников заряженных частиц или мощных лазеров.



Необходимо отметить, что активную область ферромагнитного датчика следует изготавливать преимущественно из магнитно-мягких материалов, так как процесс вращения вектора намагниченности в них не требует приложения сильных полей, а значит, увеличивается точность контролируемых параметров. Исключением являются датчики, используемые для регистрации высокоэнергетических заряженных частиц или излучений мощных лазеров, где поля значительно сильнее.

Ввиду применения магнитно-мягких материалов проблема суперпарамагнетизма для датчиков стоит особенно остро, так как не позволяет использовать нанокластеры с размером менее критического. Это сказывается на точности измерений, которая напрямую связана с количеством измерительных элементов приходящихся на единицу площади поверхности. Вторым негативным последствием суперпарамагнитного эффекта является невозможность применения магнитоориентированных датчиков в условиях повышенной температуры.

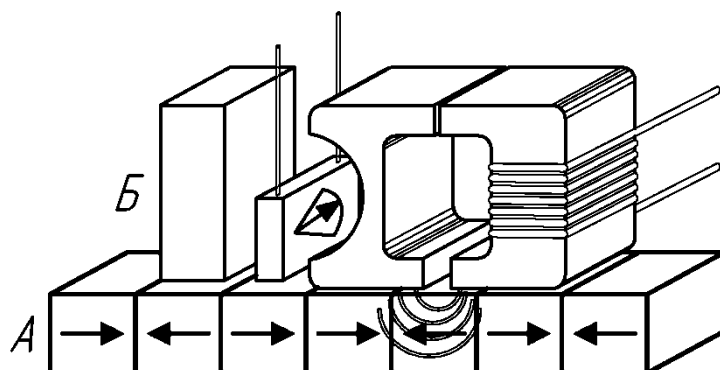
#### **4.3.2. Применение магнитоориентированных объектов в магнитных запоминающих устройствах**

Применение магнитоориентированных объектов в качестве пластин перспективных энергонезависимых магнитных запоминающих устройств обусловлено тем, что в однодоменном состоянии кластеры НКМ с одноосной симметрией имеют всего два предпочтительных направления ориентации магнитных моментов, разделенных энергетическим барьером. Намагниченная частица занимает положение в одном из стабильных состояний с наименьшей энергией, то есть под углом 0 или 180 градусов относительно оси легкого намагничивания. Для представления информации эти положения принимаются за логический ноль или единицу, что позволяет использовать кластеры магнитоориентированных объектов как биты хранения информации. Преимуществом такой записи информации служит магнитный порядок частиц, который остается стабильным долгое время при отсутствии суперпарамагнетизма. Его суть, применительно к магнитным носителям информации, заключается в том, что ферромагнитные домены, составляющие плоскость магнитных дисков, из-за чрезмерного уменьшения, не в состоянии стабильно удерживать направление магнитного момента в течение длительного времени. Проблема суперпарамагнитного эффекта для накопителей магнитных жестких дисков является одной из актуальных на сегодняшний день, из-за того, что она встала на пути прогресса в области хранения информации.

#### **4.3.3. Трудности связанные с увеличением плотности записи информации магнитных запоминающих устройств**

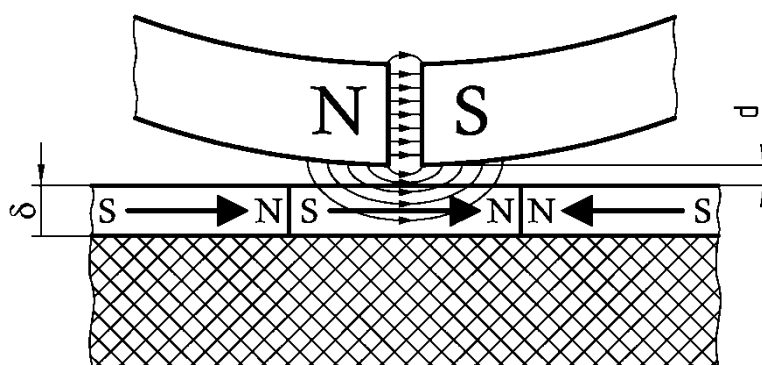
Рассмотрим подробнее трудности, связанные с увеличением удельной

плотности продольной технологии магнитной записи (рис. 4.8), которая уже более 50 лет применяется как в дисковых, так и в ленточных накопителях.



**Рис. 4.8.** Схема продольной технологии магнитной записи: А – магнитный слой, В – блок головок записи и чтения

В этой технологии намагничивание участков носителя информации в процессе записи осуществляется магнитным полем, возникающим под рабочим зазором индуктивной головки записи при протекании электрического тока через ее обмотку (рис. 4.9).



**Рис. 4.9.** Рабочий зазор головки записи

При переходе от одной магнитной ячейки к соседней происходит изменение направления намагниченности, причем не скачком, а постепенно в пределах некоторого намагниченного перехода определенной ширины. Эта доменная граница называется блоховской стенкой и является важным фактором, ограничивающим плотность записи, так как на границе намагниченных доменов возникает поле рассеяния, которое забирает энергию у магнитных полей обоих доменов. В результате смежные доменные частицы теряют часть магнитного заряда и становятся менее стабильными. Ширину доменной границы можно оценить по формуле:

$$a \equiv \sqrt{\frac{4M_r \delta (d + \delta / 2)}{QH_c}},$$

где  $M_r$  – остаточная намагниченность;  $\delta$  – толщина ферромагнитного слоя;  $d$  – расстояние головка-носитель;  $Q$  – величина, связанная с градиентом поля индуктивной головки;  $H_c$  – коэрцитивная сила носителя. Из формулы

видно, что для получения минимальной ширины доменной границы величина коэрцитивной силы  $H_C$  должна иметь как можно большее значение по сравнению с произведением  $4M_r\delta$ , а расстояние  $d$  между поверхностью пластины и головкой записи должно быть уменьшено.

Высокая удельная плотность записи информации на поверхности пластин жесткого диска, зависит и от размера самих битовых ячеек, чьи геометрические размеры определяются множеством факторов, среди которых следует выделить два наиболее важных:

1. Эффект размагничивания. Смещение доменных границ отдельных магнитных ячеек при последовательной магнитной записи происходит во внешних магнитных полях с напряженностью  $H_{см}$ , меньше величины коэрцитивной силы  $H_C$ . При уменьшении длины битовой ячейки величина напряженности магнитного поля  $H_{см}$  еще более уменьшается по сравнению с коэрцитивной силой  $H_C$ . Поэтому устойчивость битовой ячейки к внешним магнитным полям снижается. Из этого следует, что ячейку можно уменьшать до некоторого минимального размера, при котором размагничивающиеся поля становятся настолько большими, что ячейка саморазмагничивается, что влечет самопроизвольное стирание информации. Минимальные размеры ячейки, то есть размеры, при которых она еще устойчива при отсутствии внешнего магнитного поля, зависят от параметров магнитного покрытия носителя, и большую роль играет коэрцитивная сила.

2. Тепловая энергия частицы. Важным фактором магнитной стабильности битовой ячейки является температура. Увеличение температуры вызывает небольшие случайные флуктуации намагниченности частиц, аналогично тому, как оно вызывает Броуновское движение небольших частиц. Если общая энергия анизотропии однодоменной частицы  $KV$  будет порядка температурной энергии  $kT$ , то намагниченность может самопроизвольно изменяться во времени. Критический объем магнитной частицы, ниже которого наблюдается этот эффект находят по формуле:

$$V_{cr} = \frac{kT}{K} \ln(2\tau_m f_0),$$

где  $k$  – константа Больцмана;  $T$  – абсолютная температура;  $K$  – константа плотности энергии анизотропии;  $\tau_m$  – период наблюдения;  $f_0$  – частота Лармора.

Для хранения данных на магнитных носителях желательно, чтобы информация надежно хранилась много лет. Для периода хранения в 100 лет, с точки зрения температурной стабильности отношение  $KV/kT$  должно быть равно 43. В качестве примера, для периода хранения данных на протяжении 5-ти лет, это отношение должно быть равным 40. Поэтому в различной литературе отношение  $KV/kT > 43$  используется как условие стабильности магнитной ячейки [15, 21].

В реальных носителях ситуация несколько отличается от теоретических расчетов и оказывается еще более сложной. Во-первых, магнитные зер-

на в битовой ячейке имеют различные размеры и формы, а не являются идеально круглыми. Во-вторых, зерна взаимодействуют друг с другом и с внешним магнитным полем. В-третьих, при традиционном методе продольной записи на диск магнитные частицы располагаются магнитными моментами параллельно плоскости диска. А, как известно, два постоянных магнита, расположенных одинаковыми полюсами друг к другу, отталкиваются, а разными – притягиваются. Значит, между ними тоже происходит энергетическое взаимодействие. Все эти факторы приводят к тому, что критический объем, ниже которого битовые ячейки становятся нестабильными, увеличивается по сравнению с рассчитанным по формуле.

Современные оценки показывают, что суперпарамагнитный предел ограничивает плотность продольной записи на уровне 150 Гбит/дюйм<sup>2</sup>, а перпендикулярной – на уровне 500 Гбит/дюйм<sup>2</sup>. Следует отметить, что эти цифры постоянно корректируются в сторону увеличения и отражают лишь современный уровень развития технологий создания магнитных носителей. Тем не менее, суперпарамагнитный предел объективно существует, ограничивая максимальный размер магнитных доменов, что автоматически накладывает определенные ограничения на максимальную величину удельной плотности магнитной записи.

#### **4.3.4. Пути увеличения плотности магнитной записи информации**

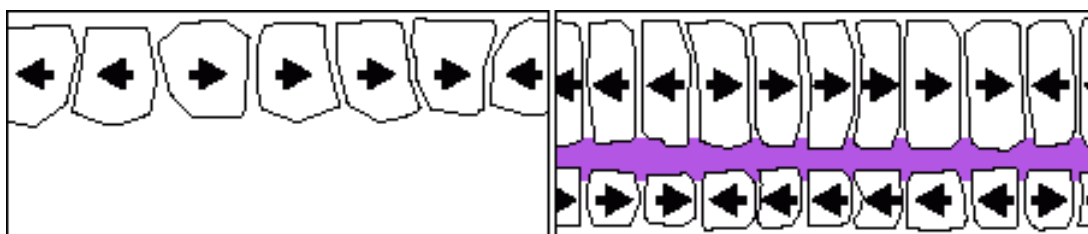
До недавнего времени активно исследовались лишь два пути предотвращения внезапной остановки на пути прогресса. Одни специалисты пытались улучшить обработку сигнала и усовершенствовать коды для исправления ошибок, неизбежно возникающих при ослаблении сигналов ввиду уязвимости к тепловым флуктуациям, которые начинают проявляться при достижении критического размера ферромагнитных доменов. Другие ученые разрабатывали особо устойчивые ферромагнитные сплавы, частицы которых надежно сохраняют магнитную ориентацию за счет низкой чувствительности к тепловому разупорядочению магнитных моментов. Поиск таких сплавов обусловлен тем, что увеличение плотности записи информации требует уменьшения размеров битов, вплоть до критического размера, ниже которого они могут спонтанно перемагничиваться и для того, чтобы этого не происходило, следует использовать среды с большой коэрцитивной силой. Разумеется, для изменения полярности магнитных частиц таких сред необходимы большие напряженности полей, которые не могут быть получены индуктивными головками записи ввиду их малых линейных размеров. Ясно, что увеличивать плотность записываемой информации только за счет этих ухищрений в дальнейшем будет все труднее и труднее.

Таким образом, используемые в настоящее время технологии продольной и перпендикулярной магнитной записи практически исчерпали свои возможности и не в состоянии обеспечить необходимые темпы роста потребно-

сти в хранении все большего и большего количества информации. Именно это вынуждает разработчиков искать новые технологии увеличения емкости магнитных накопителей, позволяющие не только вплотную подойти к суперпарамагнитному пределу, но и отодвинуть его в область более высокой плотности записи информации.

В качестве примера можно привести следующие технологии:

1) IBM и Fujitsu предлагают использовать в качестве магнитного покрытия антиферромагнитно-сопряженный носитель (antiferromagnetically-coupled или AFC) – систему из двух слоев ферромагнетика вместо традиционного одного. В таких системах магнитный момент в сопряженных ячейках противоположен. При этом ферромагнитные слои разделены слоем рутения толщиной в три атома. Сверхтонкий слой рутения – редкого металла сходного по свойствам с платиной и не обладающего магнитными свойствами – заставляет смежные слои приобретать противоположную магнитную ориентацию.



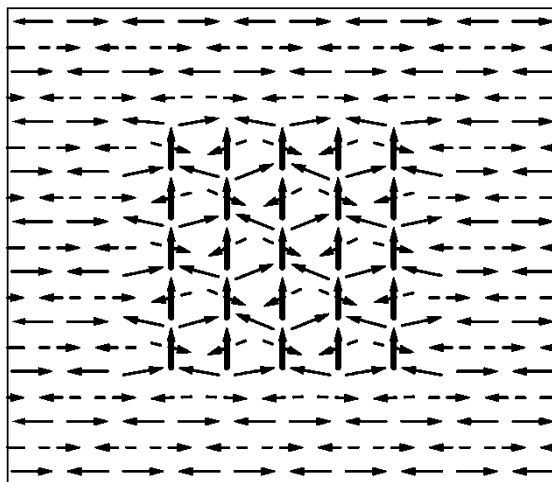
*Рис. 4.10. Схема магнитного покрытия, выполненная из одного ферромагнитного слоя (слева) и система на базе антиферромагнитно-сопряженных носителей (справа)*

Благодаря тонкому рутениевому слою, технология получила название «*prairie dust*» (дословно «пыльца феи»). Многослойная структура позволяет одновременно повысить плотность и надежность записи, так как намагниченность покрытия надежно удерживается благодаря общей толщине структуры. Плотность записи этой технологии ограничена примерно 300 Гбит/дюйм<sup>2</sup>.

2) Интернациональная группа европейских исследователей и в частности Скумриев В. и Енсен П. предлагают для повышения стабильности наносить ферромагнитные наночастицы не на парамагнитную, а на антиферромагнитную подложку. В такой системе за счет взаимодействия ферромагнитного кластера с антиферромагнитной подложкой барьер между двумя стабильными магнитными состояниями может быть увеличен в несколько раз (в зависимости от выбора материала подложки), поскольку в подобном случае необходимо изменить ориентацию спинов не только в кластере, но и в подложке.

Были проведены эксперименты, в которых изучаемые системы состояли из наночастиц кобальта (Co) диаметром 4 нм, встроенных в одном случае в парамагнитную матрицу (C или Al<sub>2</sub>O<sub>3</sub>), а в другом – в антиферромагнитную матрицу (CoO). В то время как в первой системе частицы кобальта теряли

первоначальную ориентацию своих магнитных моментов уже при 10 К, наночастицы на антиферромагнитной подложке оставались ферромагнитными вплоть до температуры 290К. Такое поведение является результатом обмен -



*Рис. 4.11. Спиновая конфигурация, возникающая при нанесении ферромагнитного кластера на антиферромагнитную подложку. Жирными стрелками обозначены спины ферромагнетика, сплошными стрелками – спины первого атомного слоя антиферромагнетика, итрихованными стрелками – спины второго атомного слоя антиферромагнетика*

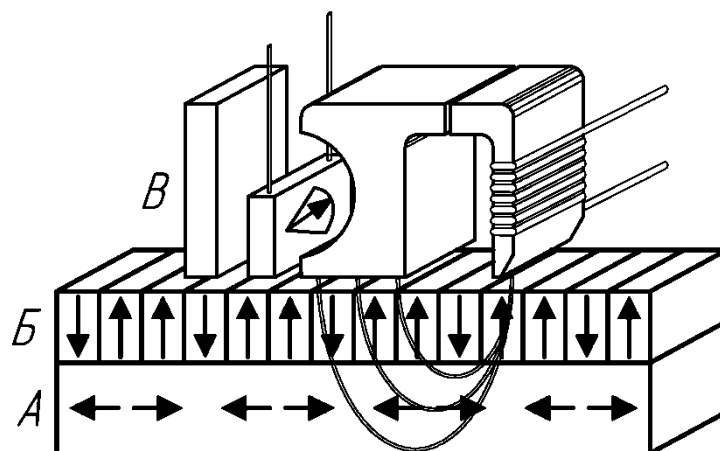
ного взаимодействия ферромагнитных наночастиц с антиферромагнитной матрицей, которое приводит к дополнительному эффективному источнику анизотропии. При правильном выборе ферромагнитных и антиферромагнитных компонентов, обменное анизотропное взаимодействие позволит добиться магнитной стабильности элементов размером всего в несколько нанометров, что позволит достичь плотности записи более чем 1 Тбит/дюйм<sup>2</sup>.

3) Seagate предлагает использовать так называемую перпендикулярную схему записи (Perpendicular recording), в которой магнитные домены ориентированы перпендикулярно поверхности диска. Технология перпендикулярной записи позволяет реализовать намагничивание ячейки данных по нормали к плоскости носителя, а не параллельно, как в случае с традиционной продольной записью. Этого можно добиться с помощью записывающей головки G-образной формы и специальной подложки, которая создает магнитный образ записывающей головки.

Основной полюс сердечника G-образной головки создает магнитное поле, силовые линии которого выходят из него перпендикулярно магнитной поверхности диска и замыкаются через нижний магнитно-мягкий подслой диска на вспомогательном полюсе сердечника. При этом частицы записывающего слоя намагничиваются вертикально, а частицы подслоя – горизонтально. Таким образом, магнитно-мягкий подслой является своего рода зеркальным отображением головки записи, что увеличивает проникновение магнитного поля.

Эта технология позволяет практически на порядок увеличить плот-

ность записи информации и уменьшает проблемы с магнитным влиянием (интерференцией). При этом соседние и различающиеся домены уже не «глядят» друг на друга одноименными полюсами, которые, как известно, оттал-



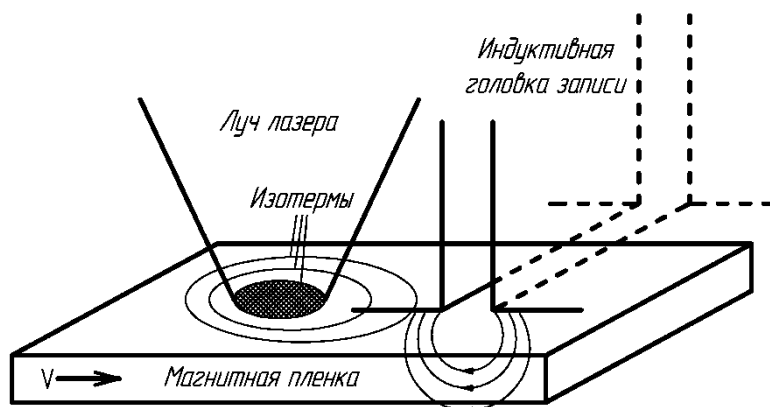
**Рис. 4.12.** Схема перпендикулярной технологии магнитной записи: А – магнитномягкий подслой; Б – магнитный слой; В – блок головок записи и чтения

квиваются. Это позволяет уменьшить размер междоменного пространства, по сравнению с классической технологией записи, что так же увеличивает емкость жестких дисков. Идея перпендикулярной записи появилась еще 20 лет назад, но в отличие от традиционной продольной записи, новая технология пока не может похвастать надежностью построенных с ее применением накопителей. К тому же стоимость винчестеров нового поколения выше из-за затрат по переводу производственных линий носителей и головок чтения/записи на новую технологию. Предполагается, что плотность записи этой технологии достигнет 1 Тбит/дюйм<sup>2</sup>.

4) Seagate предлагает использовать технологию «магнитной записи с нагревом носителя» HAMR (Heat Assisted Magnetic Recording), которая позволяет записывать данные на высокостабильном магнитном носителе, обладающем намного большей коэрцитивной силой по сравнению с используемыми в настоящее время ферромагнитными материалами. Такие материалы требуют для перемагничивания куда более сильных полей, нежели те, которые генерируют современные магнитные головки. HAMR успешно решает эту проблему путем точечного кратковременного (порядка 1 пс) нагрева лазерным лучом магнитного носителя в месте магнитной записи бита информации. При повышении температуры участка рабочего слоя происходит существенное изменение его магнитных свойств, например, уменьшение коэрцитивной силы в 3–4 раза. Таким образом, нагревание носителя облегчает запись, а быстрое последующее охлаждение позволяет стабилизировать магнитную ориентацию носителя, а значит и записанную информацию.

Используя запись с нагревом носителя, можно во много раз увеличить плотность записи, надежность и производительность накопителей на жестких магнитных дисках. Решения, в которых будет реализована технология, поя-

вятся на рынке не раньше 2012 года. По предварительным расчетам, плотность записи составит порядка 1–5 Тбит/дюйм<sup>2</sup>.



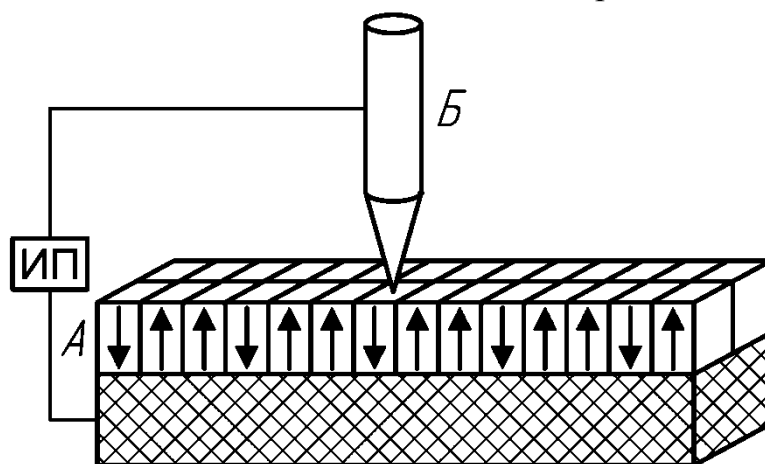
**Рис. 4.13.** Схема технологии магнитной записи с нагревом носителя

5) Seagate предлагает еще одну перспективную технологию записи – технологию самоорганизующихся магнитных решеток (Self-Organized Magnetic Array, SOMA). Данная технология предусматривает формирование на поверхности диска монодисперсного слоя «самоорганизующихся магнитных массивов» из мельчайших однородных железно-платиновых конгломератов размером около 3 нм (3 нм – это 10–15 атомов твердого вещества, выложенных в ряд). Применение этой технологии позволит существенно снизить уровень нестабильности отдельных магнитных зерен. Основная особенность данных решеток заключается в том, что магнитные зерна в них строго упорядочены и располагаются в узлах самой решетки. В перспективе это даст возможность использовать для хранения одного бита информации всего лишь одно зерно, что позволит резко увеличить плотность записи информации и в купе с HAMR достичь 40–50 Тбит/дюйм<sup>2</sup>.

По мнению ученых, революции в области хранения информации не миновать. Предпосылкой к этому является, главным образом то, что используемый в настоящее время способ продольной магнитной записи практически исчерпал свои возможности и не в состоянии обеспечить необходимые темпы роста производительности накопителей. Именно это вынуждает разработчиков искать новые технологии, которые увеличили бы емкость магнитных накопителей. Одной из таких технологий можно назвать наноманитную память. Приставка «нано» в данном случае говорит о том, что интерпретаторы двоичной системы счисления будут функционировать на базе частиц размером порядка 10 нм. Над наноманитными технологиями ведут работы многие исследователи, в том числе и в компаниях, производящих жесткие диски. Так, например, исследователи компании Hitachi Data Systems совместно с командой ученых из технологического университета Tohoku (Япония) совместно разработали метод, способный повысить плотность записи жестких дисков на несколько порядков. Предлагаемый ими метод увеличения плотности записи заключается в создании на подложке большого количества ячеек магнитной памяти и пропускании через них электрического тока



путем прикосновения тонкой иглой для смены полярности магнитов.



*Рис. 4.14. Технология пропускания электрического тока через ячейки магнитной памяти (А) путем прикосновения тонкой иглы (Б) для изменения направления магнитного момента*

В ходе проведенных экспериментов разработчики сумели добиться изменения полярности магнитных наночастиц размером 100 нм. Дополнительно к стократному увеличению плотности записываемой информации новая технология позволит серьезно снизить энергопотребление накопителей, созданных на ее основе. Предполагают, что решения, в которых будет реализована технология, появятся на рынке уже к 2010 году. Плотность записи обещает быть порядка 10 Тбит/дюйм<sup>2</sup>. Однако для реализации этой (и многих других) технологии увеличения плотности записи информации необходимо использовать магнитоориентированные объекты.

#### **4.3.5. Методы получения магнитоориентированных объектов**

В микроэлектронике метод получения магнитоориентированных объектов заключается в намагничивании нанокристаллических магнитнотвердых тел, которое состоит в переориентации векторов намагниченности кластеров в направлении приложенного внешнего магнитного поля. Оптимальным магнитоориентированным объектом является нанокристаллический материал, у которого магнитная ориентация всех кластеров параллельна и совпадает с направлением их легких осей. Для этого применяют наложение магнитного поля во время компактирования гранул нанокристаллического материала, полученных из газообразного или жидкого агрегатного состояния массивного ферромагнетика. Это обусловлено тем, что НКМ полученные из этих фаз обычно являются нанокристаллическими порошками сферической формы. Таким образом, однодоменные гранулы закрепляются с наименьшей внутренней энергией, то есть так, что их оси легкого намагничивания оказываются ориентированными в направлении силовых линий внешнего магнитного поля.

В случае наложения магнитного поля во время рекристаллизации или

пластической деформации обычно изменяется кинетика накопления рекристаллизованной фазы. К сожалению, при производстве нанокристаллического материала из твердого агрегатного состояния скорость рекристаллизации и без того очень велика, поэтому ее дополнительное увеличение на начальных стадиях недопустимо. Это обстоятельство не позволяет эффективно создавать магнитоориентированные объекты из твердого агрегатного состояния массивного ферромагнетика.

При наложении внешнего магнитного поля во время ионно-плазменной конденсации нанокристаллического материала на немагнитной подложке, намагничивание ограничено. Это связано с током, протекающим преимущественно вдоль магнитного поля, что приводит к формированию суженного цилиндрического плазменного канала [22].

Таким образом, получение магнитоориентированных объектов, у которых магнитная ориентация всех кластеров параллельна и совпадает с направлением их легких осей реализовано без особых ограничений только при компактировании из гранул нанокристаллического материала. Однако объекты, полученные таким методом, обладают высокой пористостью. Для эффективного формирования беспористых магнитоориентированных объектов требуются новые конструкторские решения.

#### **4.4. Варианты технических устройств для формирования магнитоориентированных объектов**

Рассмотрим несколько вариантов технических устройств позволяющих выполнять формирование магнитоориентированных объектов из тонких нанокристаллических пленок. Все эти устройства базируются на методе туннельно-зондовой нанотехнологии, основанной на применении сканирующего туннельного микроскопа – прибора, который позволяет исследовать и модифицировать поверхность материалов на микроном и даже атомарном уровне. Следует отметить, что именно сканирующий туннельный микроскоп позволяет проводить различные нанотехнологические процессы над отдельными магнитными кластерами. Это реализуется благодаря его конструкции, которая может быть легко дополнена необходимым оборудованием, например, магнитной системой, с помощью которой и реализуется управление магнитными моментами кластеров.

##### **4.4.1. Устройство для ориентированной сборки наноструктур**

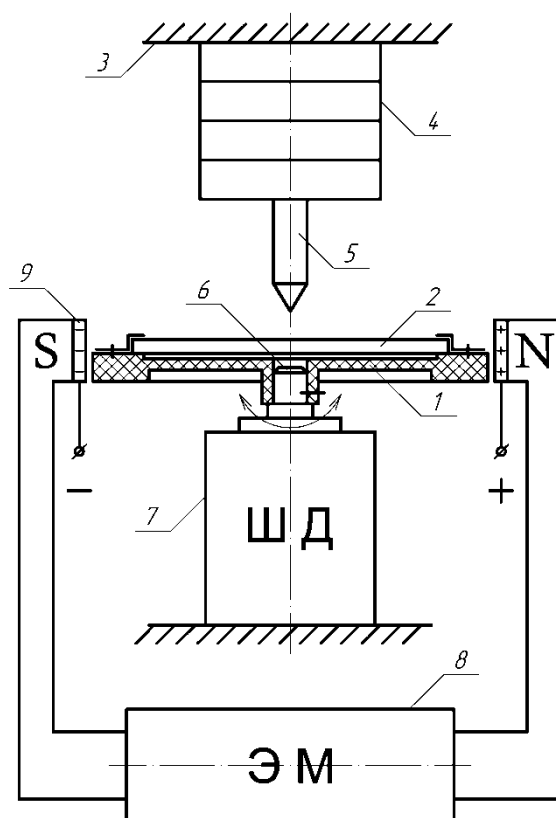
В основу полезной модели положена задача обеспечить синтез наноструктур с необходимой магнитной и дипольной ориентацией посредством предложенного устройства для ориентированной сборки наноструктур.

Эта задача решается тем, что устройство для ориентированной сборки наноструктур, содержащее подложкодержатель для размещения подложки,

неподвижную платформу, жестко связанный с ней пьезопривод, на торце которого закреплен зонд, согласно предложенной полезной модели, снабжено источником магнитного поля и источником электрического поля, размещенные таким образом, что силовые линии магнитного и электрического полей параллельны друг относительно друга и плоскости подложкодержателя, шаговым электродвигателем, на валу которого жестко закреплен подложкодержатель.

Выполнение устройства для ориентированной сборки наноструктур в виде магнита совмещенного с конденсаторными пластинами обеспечивает параллельность силовых линий магнитного и электрического полей относительно друг друга и плоскости подложкодержателя, вращающегося посредством шагового электродвигателя, что в свою очередь позволяет синтезировать электрически- и магнитоориентированные наноструктуры.

Сущность полезной модели поясняется рис. 4.15, где показано устройство для ориентированной сборки наноструктур.



*Рис. 4.15. Устройство для ориентированной сборки наноструктур*

Устройство для ориентированной сборки наноструктур содержит подложкодержатель 1 для размещения подложки 2, неподвижную платформу 3, жестко связанный с ней пьезопривод 4, на торце которого закреплен зонд 5. Подложкодержатель 1 закреплен на валу 6 шагового электродвигателя 7, жестко связанного с неподвижной платформой 3. Устройство снабжено источником магнитного поля, выполненным, например, в виде электромагнита 8 и источником электрического поля, выполненным, например, в виде конденса-

торных пластин 9, причем силовые линии магнитного и электрического полей параллельны друг относительно друга и плоскости подложкодержателя 1.

Устройство для ориентированной сборки наноструктур (см. рис. 4.15) работает следующим образом.

При приложении разности потенциалов между зондом 5 и подложкой 2, возникающее между ними неоднородное электрическое поле удерживает обрабатываемую наночастицу в промежутке зонд-подложка. После этого включается электромагнит 8 и прикладывается разность потенциалов к пластинам конденсатора 9. Возникающие магнитное и электрическое поля заставляют магнитные и дипольные моменты наночастиц ориентироваться в направлении параллельных силовых линий соответствующих полей, пронизывающих подложку 2 в продольном направлении. Затем включается шаговый электродвигатель 7, который ориентирует подложку 2, а следовательно магнитные и дипольные моменты обрабатываемой наночастицы, относительно силовых линий соответствующих полей. После необходимой ориентации обрабатываемой наночастицы в однородном магнитном и электрическом полях, производится контактная фиксация наночастицы на подложке 2, посредством зонда 5 и пьезопривода 4. Описанный выше процесс повторяется для необходимого числа наночастиц на подложке 2.

Применение предложенного устройства для ориентированной сборки наноструктур позволяет осуществлять синтез наноструктур с необходимой магнитной и дипольной ориентацией.

#### **4.4.2. Многозондовое устройство для ориентированной сборки магнитных наноструктур**

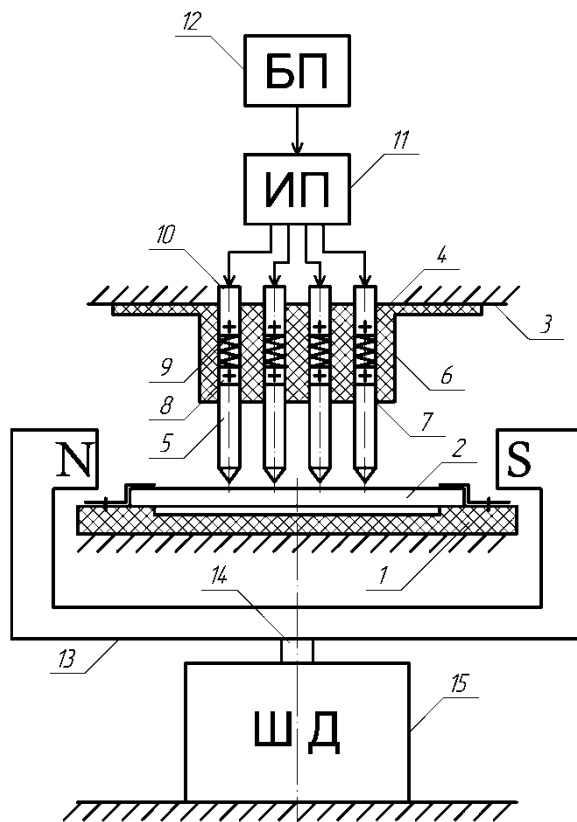
В основу полезной модели положена задача повысить быстродействие при создании структур, а также реализовать возможность ориентированной сборки магнитных наноструктур посредством предложенного многозондового устройства для ориентированной сборки магнитных наноструктур.

Эта задача решается тем, что известное устройство для сборки наноструктур, содержащее подложкодержатель для размещения подложки, неподвижную опору, жестко связанный с ней подложкодержатель и многозондовый привод, связанный с пакетом зондов, согласно предложенной полезной модели, снабжено многозондовым приводом, выполненным в виде диэлектрической кассеты с отверстиями, внутри которых, с возможностью перемещения, установлены зонды, на торцах которых расположены электретные шайбочки, выполненные, например, из фторопласта, причем торцы электретных шайбочек подпружинены посредством диэлектрических пружин, каждая из которых связана с электродом, установленным в отверстии с натягом, причем каждый электрод независимо связан с источником потенциала, который в свою очередь, связан с блоком питания, а подложка установлена между полюсами магнита, жестко связанного с валом шагового электродвигателя,

симметрично относительно полюсов магнита.

Введение в известное мнгозондовое устройство для сборки наноструктур многозондового привода, выполненного в виде диэлектрической каскады с отверстиями, внутри которых, с возможностью перемещения, установлены зонды, на торцах которых расположены электретные шайбочки, причем торцы электретных шайбочек подпружинены посредством диэлектрических пружин, каждая из которых связана с электродом, установленным в отверстии с натягом, причем каждый электрод независимо связан с источником потенциала, который в свою очередь, связан с блоком питания, а подложка установлена между полюсами магнита, жестко связанного с валом шагового электродвигателя, симметрично относительно полюсов магнита, за счет электростатического взаимодействия между электретными шайбочками и электродами, обеспечивает высокую скорость синтеза магнитоориентированных наноструктур на подложке, что позволяет повысить быстродействие при создании наноструктур на подложке, а также реализовать возможность ориентированной сборки магнитных наноструктур посредством предложенного многозондового устройства для ориентированной сборки магнитных наноструктур.

Сущность полезной модели поясняется рис. 4.16, где показано многозондовое устройство для ориентированной сборки магнитных наноструктур.



**Рис. 4.16.** Многозондовое устройство для ориентированной сборки магнитных наноструктур

Многозондовое устройство для ориентированной сборки магнитных наноструктур содержит подложкодержатель 1 для размещения подложки 2,

неподвижную опору 3, жестко связанный с ней подложкодержатель 1 и многозондовый привод 4, связанный с пакетом зондов 5. Многозондовый привод 4 устройства выполнен в виде диэлектрической кассеты 6 с отверстиями 7, внутри которых, с возможностью перемещения, установлены зонды 5, на торцах которых расположены электретные шайбочки 8. Торцы электретных шайбочек 8 подпружинены посредством диэлектрических пружин 9, каждая из которых связана с электродом 10, установленным в отверстии 7 с натягом. Каждый электрод 10 независимо связан с источником потенциала 11, который в свою очередь, связан с блоком питания 12. Подложка установлена между полюсами магнита 13, жестко связанного с валом 14 шагового электродвигателя 15, симметрично относительно полюсов магнита 13.

Многозондовое устройство для ориентированной сборки магнитных наноструктур (см. рис. 4.16) работает следующим образом.

Однородное магнитное поле между полюсами магнита 13 заставляет магнитные моменты наночастиц на поверхности подложки 2 ориентироваться в направлении магнитных силовых линий, пронизывающих подложку 2 в продольном направлении. Затем включается шаговый электродвигатель 15, который ориентирует магнит 13, а, следовательно, и магнитные моменты наночастиц на подложке 2, относительно силовых линий магнитного поля. После необходимой ориентации обрабатываемых наночастиц в однородном магнитном поле производится их контактная фиксация на подложке 2, посредством пакета зондов 5 многозондового привода 4, которые срабатывают за счет электростатического взаимодействия между соответствующими электретными шайбочками 8 и электродами 10. Это взаимодействие возникает из-за разности потенциалов на взаимодействующих поверхностях электретных шайбочек 8 и электродов 10 благодаря источнику потенциалов 11, питающегося от блока питания 12. Диэлектрические пружины 9 между электретными шайбочками 8 и электродами 10 необходимы для удержания зондов 5 при отсутствии электростатического взаимодействия между соответствующими электретными шайбочками 8 и электродами 10.

Зонды устройства не обязательно производят контактную фиксацию наночастиц на подложке, срабатывая одновременно, так как снабжены независимыми подводами электрического сигнала от источника потенциала к соответствующим электродам зондов, тем самым, обеспечивая высокую избирательность сборки магнитных наноструктур посредством предложенного устройства.

Применение предложенного многозондового устройства для ориентированной сборки магнитных наноструктур позволяет повысить быстродействие при создании структур, а также реализовать возможность ориентированной сборки магнитных наноструктур.

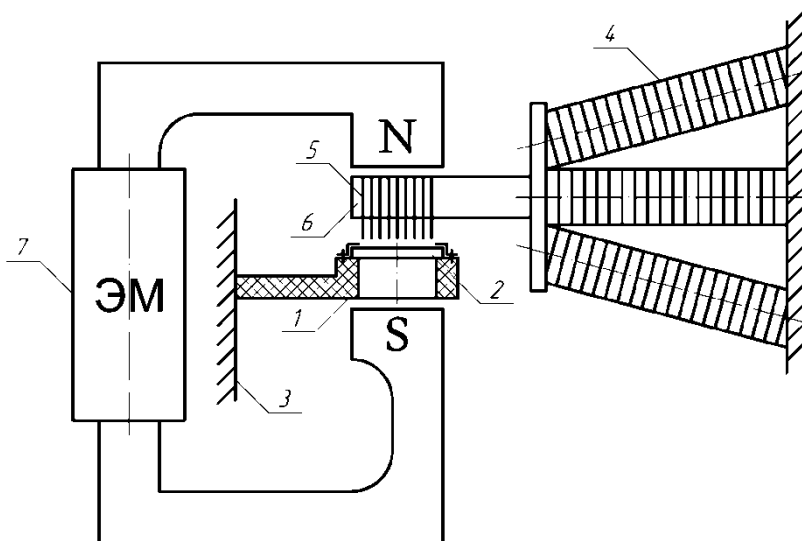
#### 4.4.3. Многозондовое устройство для формирования магнитоориентированных наноструктур

В основу полезной модели положена задача повысить быстродействие нанотехнологических операций при создании структур, а также реализовать возможность формирования магнитоориентированных наноструктур посредством предложенного многозондового устройства для формирования магнитоориентированных наноструктур.

Эта задача решается тем, что известное многозондовое устройство для формирования наноструктур, содержащее подложкодержатель для размещения подложки, неподвижную стенку, жестко связанный с ней подложкодержатель и пьезопривод, связанный с пакетом зондов, согласно предложенной полезной модели, имеет пьезопривод выполненный трехкоординатным, с возможностью перемещения пакета зондов по координатным осям X, Y, Z, пакет зондов, установленный в немагнитной обойме, связанной с трехкоординатным пьезоприводом, причем пакет зондов в немагнитной обойме и подложкодержатель установлены между полюсами электромагнита, а каждый зонд выполнен из магнитострикционного материала, например, никеля.

Введение в известное многозондовое устройство для формирования наноструктур электромагнита, трехкоординатного пьезопривода и немагнитной обоймы, снабженной пакетом зондов, выполненных из магнитострикционного материала за счет магнитострикционного эффекта обеспечивает высокую скорость сборки магнитных наноструктур на подложке, что позволяет повысить быстродействие нанотехнологических операций при создании структур, а также реализовать возможность формирования магнитоориентированных наноструктур посредством предложенного многозондового устройства для формирования магнитоориентированных наноструктур.

Сущность полезной модели поясняется рис. 4.17, где показано многозондовое устройство для формирования магнитоориентированных наноструктур.



*Рис. 4.17. Многозондовое устройство для формирования магнитоориентированных наноструктур*

Многозондовое устройство для формирования магнитоориентированных наноструктур содержит подложкодержатель 1 для размещения подложки 2, неподвижную стенку 3, жестко связанный с ней подложкодержатель 1 и трехкоординатный пьезопривод 4, связанный с пакетом зондов 5. Пакет зондов 5 установлен в немагнитной обойме 6, связанной с трехкоординатным пьезоприводом 4, причем пакет зондов 5 в немагнитной обойме 6 и подложкодержатель 1 установлены между полюсами электромагнита 7, а каждый зонд 5 выполнен из магнитострикционного материала.

Многозондовое устройство для формирования магнитоориентированных наноструктур (см. рис. 4.17) работает следующим образом.

При приложении разности потенциалов между пакетом зондов 5 и подложкой 2, возникающее между ними неоднородное электрическое поле удерживает обрабатываемые наночастицы в промежутке зонд-подложка. После этого включается электромагнит 7. Возникающее между полюсами электромагнита 7 однородное магнитное поле заставляет магнитные моменты наночастиц ориентироваться в направлении силовых линий этого поля, пронизывающего подложку 2 в поперечном направлении. Одновременно с этим, выполненные из магнитострикционного материала зонды 5, за счет возникновения в них магнитострикционного эффекта, увеличивают свои линейные размеры по координате  $Z$ , тем самым реализуя контактную фиксацию наночастиц на подложке 2.

В предложенной полезной модели, как видно из рис. 4.3, магнитная ориентация обрабатываемых наночастиц перпендикулярна плоскости подложки, соответственно, в ней реализуется выполнение группового «вертикального» синтеза магнитоориентированных наноструктур на подложке.

Применение предложенного многозондового устройства для формирования магнитоориентированных наноструктур позволяет повысить быстродействие нанотехнологических операций при создании структур, а также реализовать возможность формирования магнитоориентированных наноструктур.

### **Заключение**

На сегодняшний день исследования в области формирования магнитоориентированных объектов вполне актуальны. Интерес к таким объектам связан с возможностью их широкого применения в электронных устройствах будущего, работающих преимущественно на принципах квантовой механики – это сверхчувствительные датчики и перспективные магнитные запоминающие устройства. Выявлено, что вопрос увеличения удельной плотности магнитной записи стоит особенно остро.

Проведенные обзорно-аналитические исследования показали, что сейчас активно разрабатываются перспективные технологии магнитной записи, позволяющие реализовать накопители с довольно высокой информационной плотностью. Основой таких накопителей являются магнитоориентированные



объекты. В ходе проведенных исследований предложено несколько полезных моделей, реализующих эффективное формирование магнитоориентированных объектов из тонких нанокристаллических пленок.

Кроме чисто прикладного интереса магнитоориентированные объекты открывают широкое поле фундаментальных исследований, поэтому в ближайшем будущем следует ожидать бурного развития в этой области, что в итоге приведет к ускорению темпов развития магнитных технологий.

### Литература

1. Binnig G., Rohrer H. Scanning tunneling microscopy // *Helv. Phys. Acta.* – 1982. – Vol. 55, № 6. – P. 726-735.
2. Binnig G., Rohrer H., Gerber Ch., Weibel E. Tunneling through a controllable vacuum gap // *Appl. Phys. Lett.* – 1982. – Vol. 40. – P. 178.
3. Simons J.G. Electric tunnel effect between dissimilar electrodes separated by a thin insulating film // *J. Appl. Phys.* – 1963. – Vol. 34. – P. 2581.
4. Хэммонд П.Х. Теория обратной связи и ее применения: Пер. с англ. – М.: Наука, 1961. – 216 с.
5. Миронов В.Л. Основы сканирующей зондовой микроскопии. – Нижний Новгород, 2004. – 110 с.
6. Tersoff J., Hamann D.R. Theory of the scanning tunneling microscope // *Phys.* – 1985. – Rev. B, vol. 31 (2). – P. 805-813.
7. Неволин В.К. Основы туннельно-зондовой нанотехнологии: Учебное пособие. – М.: МГИЭТ (ТУ), 1996. – 91 с.
8. Гусев А.В. // *УФН.* – 1998. – Т. 168, № 1. – С. 55.
9. Birringer R. // *Mater. Science and Engin.* – 1989. – Vol. A117. – P. 33.
10. Gleiter H. // *Phys. B1.* – 1991. – Vol. B47, № 8. – P. 753.
11. Nieman G.W., Weerman J.R., Siegel R.W. // *Nanostructured Materials.* – 1992. – Vol. 1. – P. 185.
12. Jshida Y., Kizuka T., Xu B.S., Jchinose H. // *Ann. Chim. Fr.* – 1993. – Vol. 18. – P. 415.
13. Вонсовский С.В. Магнетизм. – М.: Наука, 1971.
14. Ландау Л.Д., Лифшиц Е. М. Теоретическая физика том 3 – Квантовая механика. – М.: Физматлит, 2001. – 804 с.
15. Суздаев И.П. Нанотехнология: физико-химия нанокластеров, наноструктур и материалов. – М.: КомКнига, 2006. – 592 с.
16. Neel L. // *Compt. rend.* – 1949. – Vol. 228. – P. 664; *Ann. Geophys.* – 1949. – Vol. 5. – P. 99.
17. Bean C.P. // *J. Appl. Phys.* – 1955. – Vol. 26. – P. 1381.
18. Вонсовский С.В. Суперпарамагнетизм // *Физический энциклопедиче-*

ский словарь, т. 5. – М., 1966. – С. 103.

19. Barkhausen H. Zwei mit Hilfe der neuen Verstärker entdeckte Erscheinungen // *Physische Zeitschrift*. – 1919. – Jg. 20, № 17.

20. Рудняк В.М. Эффект Баркгаузена // *УФН*. – 1970. – Т. 101, в. 3. – С. 429.

21. Коженевский С. Взгляд на жесткий диск «изнутри». HDD – Все о жестких дисках. – <http://hdd.kulichki.com/page.php?id=11> (25.09.2005).

22. Золотухин И.В. и др. Новые направления физического материаловедения: Учебное пособие. – Воронеж: Издательство Воронежского государственного университета, 2000. – 360 с.

23. Вишневский А.С., Ивашов Е.Н., Лучников А.П., Степанчиков С.В.. Метод записи информации на электронном носителе/ Опто-, наноэлектроника, нанотехнологии и микросистемы. – Ульяновск: УлГУ, 2008.- С.36

## Глава 5. Формирование искусственных нейронных сетей в туннельно-зондовой нанотехнологии

### Введение

Одним из наиболее мощных средств для исследования и проектирования технических систем является моделирование. Использование моделирования, начиная с ранних стадий, и постепенное накопление информации за счет уточнения и детализации модели позволяет говорить о расширяемой адаптивной модели всего цикла проектирования. Соответственно, при анализе различных свойств объекта проектирования (ОП) модельное представление должно формироваться наиболее подходящим для этой цели образом, независимо от конкретного процесса или этапа проектирования, и сохранять все требуемые свойства проектируемого объекта.

В словах «искусственные нейронные сети» слышатся отзвуки фантазий об андроидах и бунте роботов, о машинах, заменяющих и имитирующих человека. Эти фантазии интенсивно поддерживаются многими разработчиками нейросистем: рисуется не очень отдаленное будущее, в котором роботы осваивают различные виды работ, просто наблюдая за человеком, а в более отдаленной перспективе – человеческое сознание и личность перегружаются в искусственную нейронную сеть – появляются шансы на вечную жизнь.

Если перейти к прозаическому уровню повседневной работы, то нейронные сети – это всего-навсего сети, состоящие из связанных между собой простых элементов – формальных нейронов. Значительное большинство работ по нейроинформатике посвящено переносу различных алгоритмов решения задач на такие сети.

Ядром используемых представлений является идея о том, что нейроны можно моделировать довольно простыми автоматами, а вся сложность мозга, гибкость его функционирования и другие важнейшие качества определяются связями между нейронами. Каждая связь представляется как совсем простой элемент, служащий для передачи сигнала. Предельным выражением этой точки зрения может служить лозунг: «Структура связей – все, свойства элементов ничто».

Совокупность идей и научно-техническое направление, определяемое описанным представлением о мозге, называется коннекционизмом (по-английски connection – связь). Как все это относится к реальному мозгом? Так же, как карикатура и шарж со своим прототипом-человеком – это весьма ус-

ловно. И это нормально: важно не буквальное соответствие живому прототипу, а продуктивность идеи.

Существует большой класс задач: нейронные системы ассоциативной памяти, статистической обработки, фильтрации и др., для которых связи формируются по явным формулам. Но еще больше (по объему существующих приложений) задач требует неявного процесса. По аналогии с обучением животных или человека этот процесс называют обучением.

Обучение обычно строится так: существует задачник, набор примеров с заданными ответами. Эти примеры предъявляются системе. Нейроны получают по входным связям сигналы, "условия примера", преобразуют их, несколько раз обмениваются преобразованными сигналами и, наконец, выдают ответ - также набор сигналов, отклонение от правильного ответа штрафует. Обучение состоит в минимизации штрафа как (неявной) функции связей. Технике такой оптимизации и возникающим при этом задачам посвящена обширная литература.

Неявное обучение приводит к тому, что структура связей становится "непонятной", не существует иного способа ее прочесть, кроме как запустить функционирование сети. Становится сложно ответить на вопрос: "Как нейронная сеть получает результат?". То есть построить понятную человеку логическую конструкцию, воспроизводящую действия сети. В последнее время все большее внимание привлекает вопрос: как сделать работу нейронной сети понятной - доступной явному описанию.

Среди задач, возникающих при изучении искусственных нейронных сетей, особое место занимает вопрос о пределах возможности этих систем: какие задачи они могут решать? Вопросы такого рода встают всегда при появлении новых моделей вычисления, универсальных по своему замыслу. Существует несколько вариантов утверждений о том, что нейронная сеть может вычислять любую функцию с произвольной наперед заданной точностью. В данной работе вопрос о аппроксимационных возможностях нейронных сетей решается в более общем контексте и доказана теорема, единообразно охватывающая широкий спектр задач аппроксимации - от равномерного приближения функций многочленами до их аппроксимации искусственными нейронными сетями.

## 5.1 Особенности искусственных нейронных сетей

Подобно биологической нейронной системе ИНС является вычислительной системой с огромным числом параллельно функционирующих простых процессоров с множеством связей. Модели ИНС в некоторой степени воспроизводят «организованные» принципы, свойственные мозгу человека. Моделирование биологической нейронной системы с использованием ИНС может также способствовать лучшему пониманию биологических функций. Такие технологии производства, как VLSI (сверхвысокий уровень интеграции) и оптические аппаратные средства, делают возможным подобное моделирование.

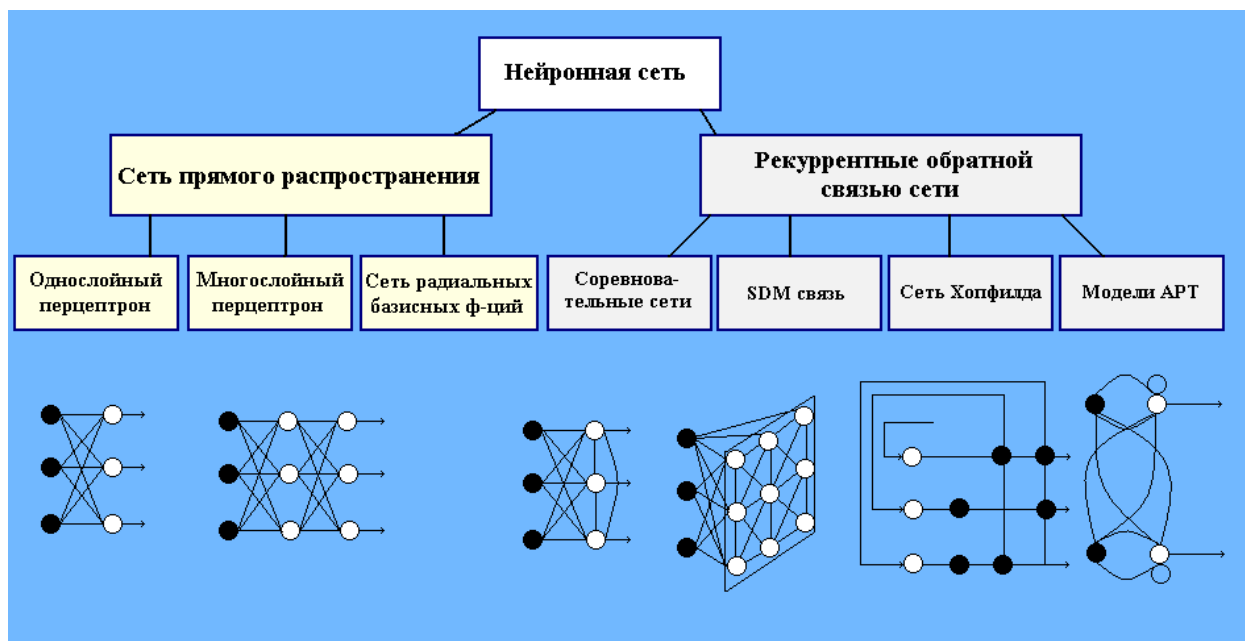
### 5.1.1. Модель технического нейрона

МакКаллок и Питтс [4] предложили использовать бинарный пороговый элемент в качестве модели искусственного нейрона. Этот математический нейрон вычисляет взвешенную сумму  $n$  входных сигналов  $x_j$ ,  $j=1,2,\dots,n$ , и формирует на выходе сигнал величины 1, если эта сумма превышает определенный порог  $u$ , и 0 – в противном случае.

Часто удобно рассматривать  $u$  как весовой коэффициент, связанный с постоянным входом  $x_0=1$ . Положительные веса соответствуют возбуждающим связям, а отрицательные – тормозным. МакКаллок и Питтс доказали, что при соответствующим образом подобранных весах совокупность параллельно функционирующих нейронов подобного типа способна выполнять универсальные вычисления. Здесь наблюдается определенная аналогия с биологическим нейроном: передачу сигнала и взаимосвязи имитируют аксоны и дендриты, веса связей соответствует синапсам, а пороговая функция отражает активность сомы.

### Архитектура нейронной сети

ИНС может рассматриваться как направленный граф со взвешенными связями, в котором искусственные нейроны являются узлами. По архитектуре связей ИНС могут быть сгруппированы в два класса (рис. 5.1): сети прямого распространения, в которых графы не имеют петель, и рекуррентные сети, или сети с обратными связями. В наиболее распространенном семействе сетей первого класса, называемых многослойным перцептроном, нейроны расположены слоями и имеют однонаправленные связи между слоями.



*Рис.5.1. Систематизация архитектур сетей прямого распространения и рекуррентных (с обратной связью)*

На рис. 5.1 представлены типовые сети каждого класса. Сети прямого распространения являются статическими в том смысле, что на заданный вход они вырабатывают одну совокупность выходных значений, не зависящих от предыдущего состояния сети. Рекуррентные сети являются динамическими, так как в силу обратных связей в них модифицируются входы нейронов, что приводит к изменению состояния сети.

## Обучение

Способность к обучению является фундаментальным свойством мозга. В контексте ИНС процесс обучения может рассматриваться как настройка архитектуры сети и весов связей для эффективного выполнения специальной задачи. Обычно нейронная сеть должна настроить веса связей по имеющейся обучающей выборке. Функционирование сети улучшается по мере итеративной настройки весовых коэффициентов. Свойство сети обучаться на примерах делает их более привлекательными по сравнению с системами, которые следуют определенной системе правил функционирования, сформулированной экспертами.

Для конструирования процесса обучения, прежде всего, необходимо иметь модель внешней среды, в которой функционирует нейронная сеть - знать доступную для сети информацию. Эта модель определяет парадигму обучения [3]. Во-вторых, необходимо понять, как модифицировать весовые параметры сети - какие правила обучения управляют процессом настройки. Алгоритм обучения означает процедуру, в которой используются правила обучения для настройки весов.

Существуют три парадигмы обучения: "с учителем", "без учителя" (самообучение) и смешанная. В первом случае нейронная сеть располагает правильными ответами (выходами сети) на каждый входной пример. Веса настраиваются так, чтобы сеть производила ответы как можно более близкие к известным правильным ответам. Усиленный вариант обучения с учителем предполагает, что известна только критическая оценка правильности выхода нейронной сети, но не сами правильные значения выхода. Обучение без учителя не требует знания правильных ответов на каждый пример обучающей выборки. В этом случае раскрывается внутренняя структура данных или корреляции между образцами в системе данных, что позволяет распределить образцы по категориям. При смешанном обучении часть весов определяется посредством обучения с учителем, в то время как остальная получается с помощью самообучения.

Теория обучения рассматривает три фундаментальных свойства, связанных с обучением по примерам: емкость, сложность образцов и вычислительная сложность. Под емкостью понимается, сколько образцов может запомнить сеть, и какие функции и границы принятия решений могут быть на ней сформированы. Сложность образцов определяет число обучающих примеров, необходимых для достижения способности сети к обобщению. Слишком малое число примеров может вызвать "переобученность" сети, когда она хорошо функционирует на примерах обучающей выборки, но плохо - на тестовых примерах, подчиненных тому же статистическому распределению. Известны 4 основных типа правил обучения: коррекция по ошибке, машина Больцмана, правило Хебба и обучение методом соревнования.

### **5.1.2. Обобщенная аппроксимационная теорема и вычислительные возможности искусственных нейронных сетей**

Исследуются вычислительные возможности искусственных нейронных сетей. В связи с этим происходит возврат к классическому вопросу о представлении функций многих переменных с помощью суперпозиций и сумм функций одного переменного и новая редакция этого вопроса (ограничение одной произвольно выбранной нелинейной функцией одного переменного).

Показано, что можно получить сколь угодно точное приближение любой непрерывной функции многих переменных, используя операции сложения и умножения на число, суперпозицию функций, линейные функции а также одну произвольную непрерывную нелинейную функцию одной переменной.

Аппроксимационная теорема Стоуна переносится с колец функций на любые их алгебры, замкнутые относительно произвольной нелинейной операции, а также относительно сложения и умножения на число.

Для нейронных сетей полученные результаты означают: от функции активации нейрона требуется только нелинейность - и более ничего. Какой бы она ни была, можно так построить сеть связей и подобрать коэффициенты линейных связей между нейронами, чтобы нейронная сеть сколь угодно точно вычисляла любую непрерывную функцию от своих входов.

### **5.1.3. Теоретический подход к возможности ускоренного обучения искусственных нейронных сетей за счет адаптивного упрощения обучающей выборки**

С увеличением размерности задач, возлагаемых на интеллектуальные системы, повышение скорости и качества обучения ИНС приобретает все большее значение. Подтверждением этому является значительное количество исследований, направленных на ускорение обучения ИНС. Во многих предложенных методах увеличение скорости достигается за счет модификации метода градиентного спуска, основанной на предположении о том или ином характере исходных данных [6]. Также часто используется динамическое изменение шага обучения [7], дополнительные методы вывода ИНС из локального минимума. Однако перечисленные методы не учитывают то обстоятельство, что дополнительное повышение скорости и качества обучения может быть достигнуто и за счет обработки данных, на которых это обучение происходит.

Обратившись к естественным обучающимся системам, можно заметить, что чаще всего обучение происходит не сразу на всем обучающем множестве (которым для естественных систем являются объекты реального мира), а на его упрощенной модели, отражающей лишь некоторые примеры и закономерности. По мере усвоения более простого материала модель становится все более подробной и адекватной. Т.е. обучение происходит как бы «от простого к сложному». Исследование применимости такого подхода для повышения качества и скорости обучения ИНС и является целью данной работы.

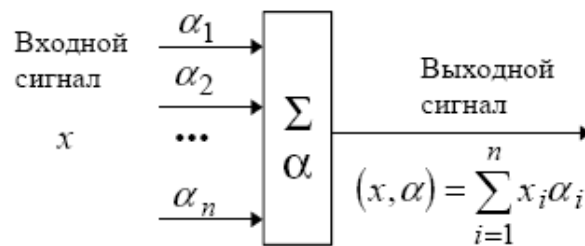
Для описания алгоритмов и устройств в нейроинформатике выработана специальная "схемотехника", в которой элементарные устройства - сумматоры, синапсы, нейроны и т.п. – объединяются в сети, предназначенные для решения задач.



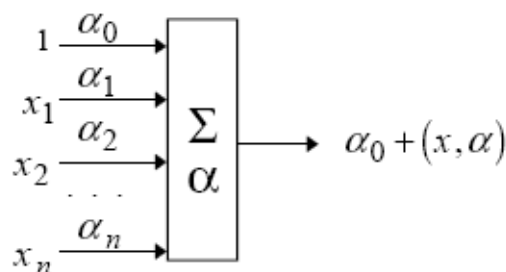
Интересен статус этой схмотехники - ни в аппаратной реализации нейронных сетей, ни в профессиональном программном обеспечении все эти элементы вовсе не обязательно реализуются как отдельные части или блоки. Используемая в нейроинформатике идеальная схмотехника представляет собой особый язык для представления нейронных сетей. При программной и аппаратной реализации описания, выполненные на этом языке, переводятся на языки другого уровня, более пригодные для реализации.

Важнейший элемент нейросистем - это *адаптивный сумматор*. Он вычисляет скалярное произведение вектора входного сигнала  $x$  на вектор параметров  $a$ . На схемах будем обозначать его так, как показано на рис. 5.2. Адаптивным называем его из-за наличия вектора настраиваемых параметров  $a$ .

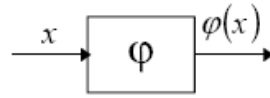
Для многих задач полезно иметь линейную неоднородную функцию выходных сигналов. Ее вычисление также можно представить с помощью адаптивного сумматора, имеющего  $n+1$  вход и получающего на 0-й вход постоянный единичный сигнал (рис. 5.3).



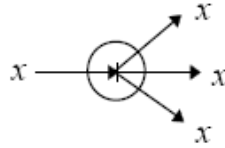
**Рис. 5.2.** Адаптивный сумматор



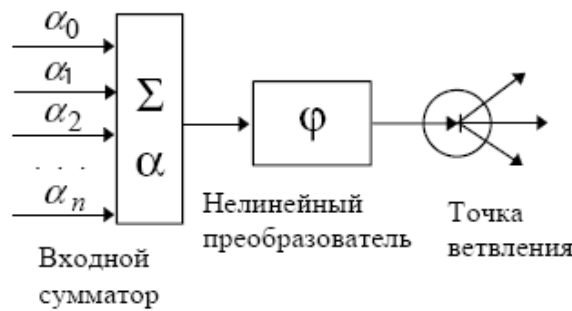
**Рис. 5.3.** Неоднородный адаптивный сумматор



**Рис.5.4.** Нелинейный преобразователь сигнала



**Рис.5.5.** Точка ветвления



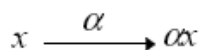
**Рис. 5.6.** Формальный нейрон

Нелинейный преобразователь сигнала изображен на рис. 5.4. Он получает скалярный входной сигнал  $x$  и переводит его в  $\phi(x)$ .

Точка ветвления служит для рассылки одного сигнала по нескольким адресам (рис. 5.5). Она получает скалярный входной сигнал  $x$  и передает его всем своим выходам.

Стандартный формальный нейрон составлен из входного сумматора, нелинейного преобразователя и точки ветвления на выходе (рис. 5.6).

Линейная связь - синапс - отдельно от сумматоров не встречается, однако для некоторых рассуждений бывает удобно выделить этот элемент (рис. 5.7). Он умножает входной сигнал  $x$  на «вес синапса»  $a$ .



**Рис. 5.7.** Синапс

Также бывает полезно «присоединить» связи не к входному сумматору, а к точке ветвления. В результате получаем элемент, двойственный адаптивному сумматору и называемый «выходная звезда». Его выходные связи производят умножение сигнала на свои веса.

Роль взвешивающей функции может выполнять функция от расстояния между входными векторами, удовлетворяющая следующим условиям:

1. Существовать и быть неотрицательной на всем множестве возможных значений расстояния.
2. Убывать с увеличением расстояния.
3. В зависимости от некоторого параметра  $\alpha$  изменять скорость убывания.

Таким образом, параметр  $\alpha$  определяет «крутизну» взвешивающей функции и задает степень упрощения исходной выборки. Одной из наиболее известных и широко применяемых функций, удовлетворяющих перечисленным условиям, является функция Гаусса, которую и предлагается использовать в качестве взвешивающей в данной работе.

$$c_{jk} = e^{-\left(\frac{\|X_k - X_j\|}{\alpha}\right)^2}.$$

Ниже приведен пример упрощения функции одной переменной при различных значениях параметра  $\alpha$  (рис. 5.8).

Для количественной оценки упрощения ОВ в процессе обучения ИНС рассмотрим следующие величины:  $\delta(\text{ОВ}; \text{ОВ})$  – отклонение упрощенной выборки от исходной;

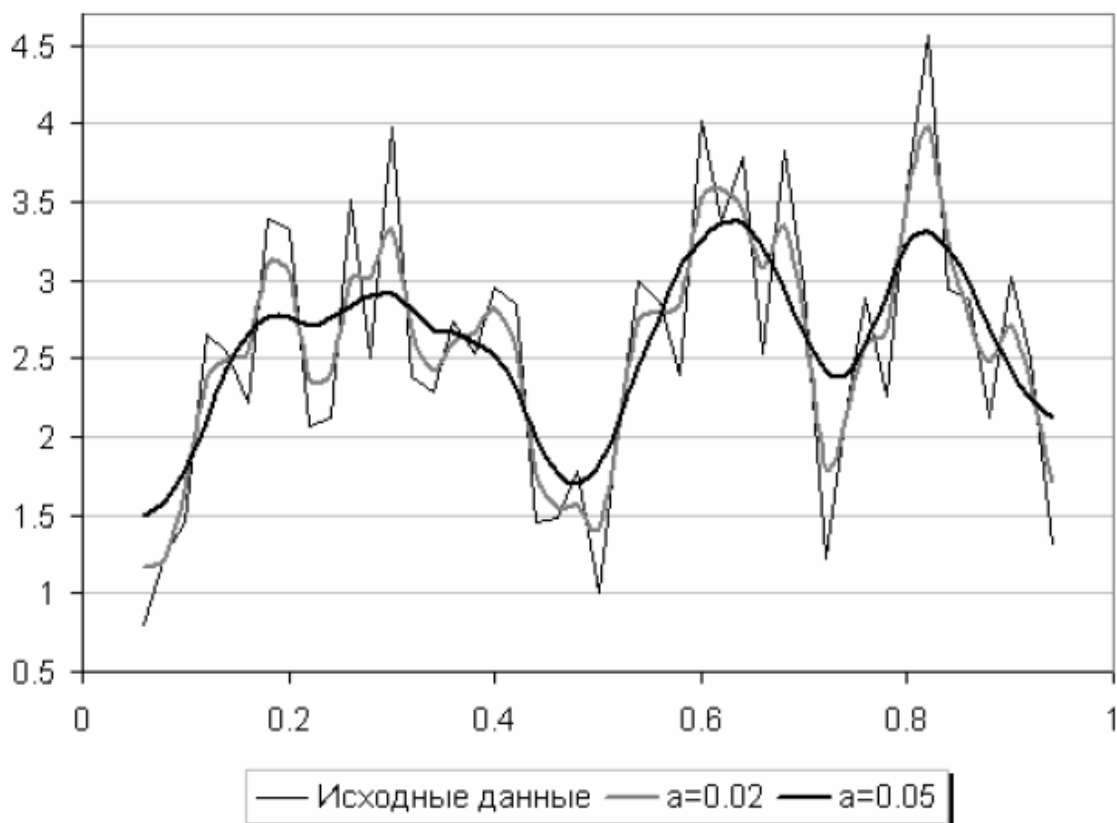
$\delta(\text{НС}; \text{ОВ}')$  – ошибка НС на упрощенной выборке ОВ';

$\delta(\text{НС}; \text{ОВ})$  – ошибка НС, обученной на упрощенной выборке, рассчитанная для исходной выборки ОВ. Пусть эти величины определены как среднее расстояние между выходными векторами в выбранной метрике. Тогда имеет место неравенство:

$$\delta(\text{ОВ}'; \text{ОВ}) + \delta(\text{НС}; \text{ОВ}') \geq \delta(\text{НС}; \text{ОВ}). \quad (1)$$

Это позволит воспользоваться левой частью неравенства в качестве критерия остановки обучения, а не тратить время на дополнительный расчет  $\delta(\text{НС}; \text{ОВ})$ . Вместе с тем нет необходимости обучаться на ОВ' с точностью, большей точности самой ОВ'. Следовательно, должно выполняться соотношение:

$$\delta(\text{ОВ}'; \text{ОВ}) \geq \delta(\text{НС}; \text{ОВ}'). \quad (2)$$



**Рис. 5.8.** Пример упрощения исходной функции

Учитывая (1) и (2), можно предложить следующий алгоритм обучения ИНС:

- 1. Задается начальное значение параметра упрощения.**
- 2. Формируется упрощенная выборка  $OB'$  и рассчитывается  $\delta(OB'; OB)$ .**
- 3. Производится обучение ИНС выборке  $OB'$  до тех пор, пока не выполнится одно из условий:**

**a)  $\delta(OB'; OB) + \delta(НС; OB') \geq \delta_{доп}$ , где  $\delta_{доп}$  - допустимая ошибка, определяемая требуемой точностью решения задачи. Выполнение условия означает окончание обучения.**

**b)  $\delta(OB'; OB) \geq \delta(НС; OB')$ . При выполнении – уменьшение параметра  $\alpha$  и переход на шаг 2.**

Данный алгоритм позволяет изменить процесс обучения так, что в начале ИНС будет обучаться основным тенденциям и закономерностям, несколько теряя в точности, но зато не повторяя возможно присутствующий в исходной выборке шум. По мере усложнения выборка  $OB'$  будет приближаться к исходной и, в конечном итоге, либо повторит ее, либо обеспечит достаточную точность решения задачи, что для ИНС будет означать финальный этап обучения.

## Практические результаты обучающей выборки

В ходе предварительных экспериментов производилось сравнение ошибок обучения ИНС с применением предложенного подхода и без него. Исходными данными служила функция 2-х переменных с шумовой составляющей, взятая из набора функций, сформированных для сравнения обучающих алгоритмов [21]. Обучающая выборка содержала 225 наборов, а контрольная - 10000. Сеть обучалась в течение 5000 эпох, затем рассчитывалась ошибка на контрольной выборке (КВ).

За время обучения было произведено 4 итерации упрощения ОВ. К моменту окончания обучения среднеквадратичное отклонение упрощенной выборки от исходной составляло 0,13, среднеквадратичная ошибка ИНС на упрощенной выборке – 0,14, на исходной – 0,17. Ошибка на КВ составила 0,48.

В случае с неизменяющейся ОВ, к окончанию обучения ошибка на ОВ составила 0,18; на КВ – 0,56.

В результате для случая с упрощением ОВ было отмечено снижение ошибок как на ОВ (6%), так и на КВ (15%).

Таким образом, использование адаптивного упрощения ОВ позволяет снизить время и, что более важно, повысить качество обучения ИНС. Это достигается в основном за счет снижения избыточной подробности обучающего множества на ранних этапах обучения, что вполне характерно для естественных обучающихся систем.

Используемые в подходе преобразования относятся только к исходным данным и не затрагивают алгоритма настройки весовых коэффициентов ИНС. Это делает подход совместимым со многими известными методами ускоренного обучения ИНС, тем самым давая дополнительный выигрыш во времени и качестве обучения.

В отношении дальнейшего развития подхода можно отметить исследование неравномерного упрощения ОВ, когда коэффициент упрощения различен для каждого набора и определяется с учетом ошибки ИНС на данном наборе, а не на всей выборке в среднем.

### 5.1. Модели ускоренного обучения искусственных нейронных сетей

В этом разделе поставленная задача рассматривается с точки зрения переобучения нейронной сети. В большинстве случаев нейронную сеть обучают, пока ее ошибка не станет равной нулю. Это приводит порой к неоправданным затратам драгоценных ресурсов времени, хотя для решения большей

части задач *достаточно*, чтобы эта ошибка не превышала определенного значения.

Иногда степень *достаточности* определяется исходя из условий задачи и искомого результата. Однако в большинстве случаев этот процесс протекает на интуитивном уровне и руководствующий принцип не фиксируется сознанием в достаточной мере. На самом деле этот момент является одним из самых важных в решении задач подобного типа, и оптимальное значение варьируемого параметра может зависеть от многих исходных величин и ограничений, накладываемых на решение задачи. Таким образом, появляется необходимость в формализации данного принципа; в дальнейшем — *Принципа Достаточности (ПД)*.

Идеально обученной считается такая сеть, которая полностью повторяет ОВ [5], т.е. ее глобальная ошибка равна нулю. Но обучение нейронной сети до такой степени представляет собой очень трудоемкую задачу, а нередко и вовсе неразрешимую. Эти трудности обычно связаны с тем, что разные классы имеют похожие объекты, и чем таких объектов больше и чем больше они похожи, тем труднее будет обучить нейронную сеть.

Суть ПД заключается в отказе от обязательного стремления к идеалу при поиске решения конкретной задачи. Рассматривая эту проблему с точки зрения ПД в рамках глобальной и локальной ошибки, можно сказать, что далеко не всегда необходима 100%-я точность распознавания. Иногда для того чтобы отнести исследуемый объект к заданному классу, достаточно, чтобы ошибка сети на данном наборе не превышала некоторого  $\delta$ .

Минимальное значение  $\delta$  зависит от характера обучающей выборки. В качестве параметров, характеризующих ОВ, рассмотрим ее полноту, равномерность и противоречивость. Полнота выборки характеризуется обеспеченностью классов обучающими наборами; количество обучающих наборов для класса должно в 3-5 раз больше, чем используемое в наборе число признаков класса. Пусть величина, характеризующая полноту выборки, вычисляется следующим образом:

$$F_{OB} = \frac{N_F}{N} \cdot 100 \%,$$

где  $N_F$  — число классов, удовлетворяющих указанному условию,  $N$  — общее число классов.

### 5.1.1. Модель нейросетевой структуры для оптимизации функционирования

Предлагается алгоритм, основанный на процессе обратного воспроизводства, который динамически развивает структуру наращиваемых многослойных нейронных сетей и демонстрирует их потенциал в плане применения для управления. Данный алгоритм содержит процесс "генерирования и испытания" и оценивает действие используемой структуры, а также изменяет ее в соответствии с используемыми альтернативами и отбирает наиболее перспективную. Алгоритм изменяет структуру нейронных сетей путем добавления или удаления нейронов или слоев. Эффективность алгоритма продемонстрирована при испытании в нескольких опытах с многообещающими результатами.

Нейронная сеть, представленная здесь, является многослойной наращиваемой нейронной сетью, основанной на алгоритме обратного распространения (Румельхарт, Хинтон, Уильямс, 1986 г.). Функция активации нейрона  $j$  может быть выражена так:

$$f(x) = \frac{2}{1 + e^{-x}} - 1, \quad (3)$$

и

$$x = \sum_j w_{ji} o_i + \theta_j, \quad (4)$$

где  $O_i$  выход части  $i$ ,  $W_{ji}$  вес между частью  $i$  и частью  $j$  и  $\theta_j$  пороговая величина части  $j$ .

Производная нейронной сети в формуле:

$$O_{NN} = N(W, I). \quad (5)$$

Формулы (3), (4) и (5) отображают  $n$ -слойную нейронную сеть с сигмоидальной функцией.

$$N(W, I) = F(W_n Z_{n-1});$$

$$Z_{n-1} = F(W_{n-1} Z_{n-2}); \quad (6)$$

$$Z_1 = F(W_1 I), \quad (7)$$

где  $I$  - входной набор;  $W_1$  - матрица весовых коэффициентов между входным и выходным слоями;  $W_{n-1}$  - матрица весовых коэффициентов между  $n-2$  и  $n-1$  слоями и  $W_{n-n-1}$  и выходным слоем.

Обобщенный алгоритм правила дельта (Румельхарт, 1986 г.) используется для обновления веса нейронной сети для того, чтобы минимизировать функцию стоимости (см. формулы (6) и (7)).

$$E = \frac{1}{2} \sum_k (t_{pk} - o_{pk})^2 ,$$

где  $t_{pk}$  и  $o_{pk}$  желаемая и полученная величины, соответственно, выхода  $k$

$$E = \frac{1}{2} \sum_k (t_k - o_k)^2 .$$

Полученная конвергенция (сходимость) выражается следующими формулами .

$$w_{ji}(t+1) = w_{ji}(t) + \nabla w_{ji}(t+1) ,$$

$$\nabla w_{ji}(t+1) = \eta \delta_j O_i ,$$

$$\delta_j = (t_k - o_k) O_j (1 - O_j)$$

если  $j$  – выходной слой, и

$$\delta_j = O_j (1 - O_j) \sum_k \delta_k w_{kj}$$

если  $j$  – внутренний слой, где  $k$  – преуспевающий уровень.

Вышеупомянутые уравнения показывают, что ВР перемещает веса в направлении максимального уменьшения, пока  $E_{\min}$  не достигнут, когда значение  $\|\nabla_w E\| = 0$  ,

где  $\|\nabla_w E\| = \sqrt{\sum_i (\Delta w_i)^2}$  .

Индекс  $i$  охватывает все веса в сети.

### 5.1.2. Исследование искусственных нейронных сетей на полупроводниковых нанобъектах – квантовых точках и ямах

В качестве элемента искусственной нейронной сети – нелинейного преобразователя выступает изомерная квантовая точка. При описании модели физического явления удобно использовать оболочечную модель ядра. В оболочечной модели ядра принимается, что энергетическая структура (уровни энергии нуклонов) ядра подобна энергетической структуре электронной оболочки атома. Сильное взаимодействие нуклонов в ядре и малый радиус этого взаимодействия позволяет рассматривать нуклоны движущимися независимо друг от друга в поле, обладающем сферически симметричным потенциалом. При этом нуклоны могут находиться в различных энергетических состояниях. Основному состоянию ядра должно соответствовать заполнение всех нижних уровней. Потеря нуклоном энергии при межнуклонных столкновениях не может перевести его в более низкое состояние, ибо все они заняты в соответствии с принципом Паули. Это приводит к то-



му, что длина свободного пробега нуклона в невозбужденном ядре становится больше радиуса ядра. Это означает возможность рассматривать нуклоны в рамках данной модели невзаимодействующими и несталкивающимися. Движение невзаимодействующих нуклонов в поле сферического потенциала, где орбитальный момент импульса является интегралом движения, характеризуется тем, что всем  $2l+1$  возможным ориентациям вектора  $l$  соответствует одинаковый энергетический уровень. На этом уровне размещаются  $2(2l+1)$  нуклонов данного типа. Таким образом, в оболочечной модели нуклоны располагаются в определенном количестве на энергетических *нуклонных оболочках*. Каждый нуклон характеризуется индивидуальной волновой функцией и индивидуальными квантовыми числами  $n$  и  $l$ . Существуют две системы нуклонных состояний - одна для протонов, другая для нейтронов; обе системы уровней заполняются нуклонами независимо друг от друга. Ядра, имеющие только заполненные *нуклонные оболочки*, должны обладать повышенной устойчивостью (проявляющейся, например, в их большей распространенности в природе), а также должны иметь сферически симметричное распределение заряда.

Порядок заполнения нуклонных оболочек с ростом  $A$  сходен с порядком заполнения электронных оболочек с ростом  $Z$ . Ввиду сильной спин-орбитальной связи все уровни с  $l \neq 0$  расщепляются на два подуровня с  $j = l \pm \frac{1}{2}$ , заполняющихся независимо.

Предсказания оболочечной модели, в общем, соответствуют действительности. Наиболее устойчивым по сравнению с соседними ядрами являются ядра со значениями  $N$  или  $Z$ , равными **2, 8, 20, 28, 50, 82, 126 и 152**. Эти числа называются *магическими*. Распространенность в природе таких ядер наиболее велика, а квадрупольные моменты их близки к нулю. Ядра, у которых магическими числами являются и  $N$  и  $Z$ , называются *дважды магическими*. Эти ядра ( ${}^4_2\text{He}$ ,  ${}^{16}_8\text{O}$ ,  ${}^{40}_{20}\text{Ca}$ ,  ${}^{208}_{82}\text{Pb}$ ) обладают особой устойчивостью, проявляющейся, в частности, в том, что они являются наиболее распространенными в природе изотопами этих элементов.

*Гамма-излучением* называется жесткое электромагнитное излучение, энергия которого высвобождается при переходах ядер из возбужденного в основное или в менее возбужденное состояние, а также при ядерных реакциях.

В первом случае энергия гамма-квантов равна разности энергий конечного и начального уровней ядра. В каждом акте перехода ядро излучает один гамма-квант. В связи с дискретностью энергетических уров-

ней ядра гамма-излучение имеет линейчатый спектр. Частоты гамма-квантов связаны с разностью энергий условием частот Бора.

При испускании ядром гамма-кванта само ядро, вследствие закона сохранения импульса, приобретает противоположно направленный импульс (*отдача*). Если ядра, испускающие гамма-кванты, находятся в твердом теле, то спектр гамма-лучей состоит из двух компонент:

а) компоненты с естественной шириной гамма-линии  $\Gamma$ , определяемой временем жизни ядер в данном возбужденном состоянии с энергией  $E$ ;

б) компоненты с шириной линии  $\Gamma_R \sim E \frac{\bar{u}}{c} \gg \Gamma$ , где  $\bar{u}$  - средняя квадратичная скорость теплового движения гамма-радиоактивных ядер в твердом теле; эта компонента имеет энергию, смещенную относительно значения  $E$  на величину энергии отдачи  $R = E^2 / 2M_0 c^2$ , где  $M_0$  - масса излучающего ядра (если считать его свободным и движущимся со скоростью  $\bar{u} \ll c$ ).

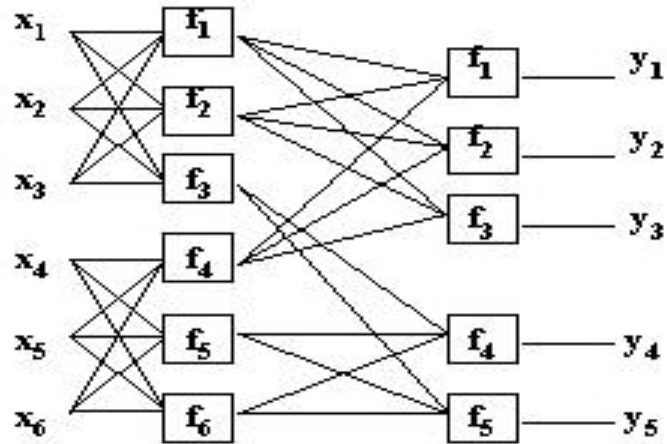
В результате линии гамма-излучения и поглощения (той же линии) сильно размываются и, кроме того, сдвинуты по энергии друг относительно друга на величину  $\sim 2R$ . Ввиду того, что для гамма-излучения  $R$  в общем не мало по сравнению с  $E$ , явление *резонансного поглощения гамма-лучей* ( $E_{изл} = E_{погл}$  или  $\nu_{изл} = \nu_{погл}$ ) обычно практически не наблюдается.

При определенных условиях удается добиться того, что излучаемый гамма-квант передает импульс не одному излучающему ядру, а всему кристаллу в целом. В результате излучаемой линии соответствует энергия отдачи  $R \approx 0$  ( $M$  – велико) и  $\Gamma_R \approx \Gamma$ , т.е. ширина линий приближается к естественной, а сдвиг по энергии практически исчезает. Одним из условий четкого проявления эффекта Мессбауэра, является условие  $R \leq 2k\Theta_D$ , где  $\Theta_D$  - дебаевская температура кристалла,  $k$  – постоянная Больцмана. При  $R \ll 2k\Theta_D$  гамма-переходы «без отдачи» можно наблюдать уже при комнатной температуре.

## 5.2. Варианты технических устройств для получения элементов искусственных нейронных сетей

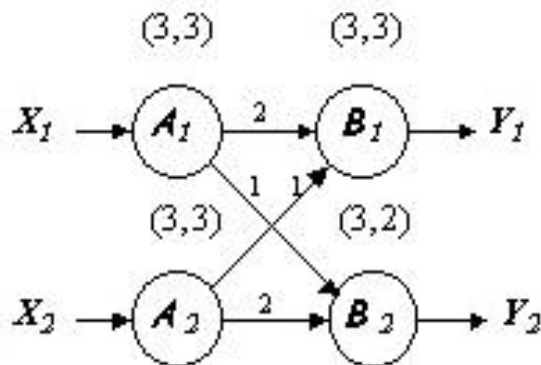
Рассмотрим несколько вариантов установок элементов искусственных нейронных сетей.

### 5.2.1. Искусственная нейронная сеть



*Рис. 5.9. Двухслойная искусственная нейронная сеть*

Для искусственной нейронной сети достаточным структурным описанием может служить ориентированный граф, в котором каждая вершина соответствует одному нейронному ядру, а дуги отвечают операторам межъядерных связей (рис. 5.9). Такой граф называется структурной моделью искусственной нейронной сети. На рис.5.10 показан пример структурной модели для топологической реализации, представленной на рис.5.9. На графе структурной модели показаны веса вершин и веса дуг. Вес вершины  $i$  определяется структурной характеристикой ядра  $(A_i, B_i)$ , а вес дуги  $(r_{ij})$  равен рангу проектирующего оператора межъядерной связи.



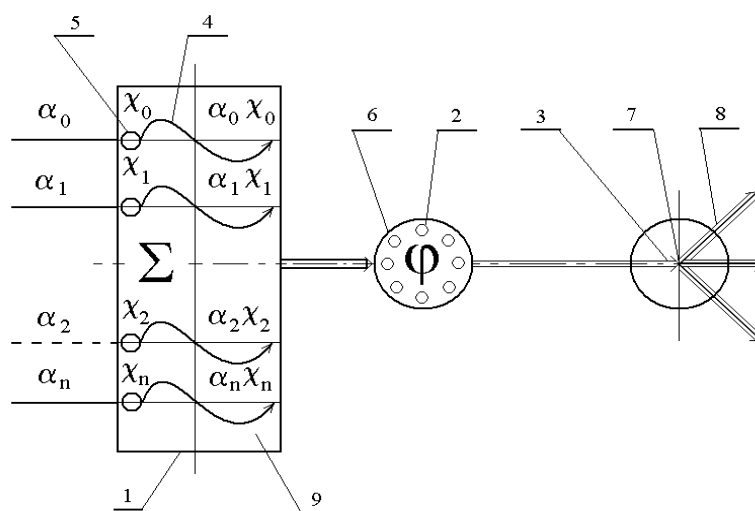
*Рис.5.10. Структурная модель*

Обратный переход от структурной модели к топологической реализации неоднозначен, что можно трактовать как свойство топологической пластичности ядерной нейронной сети. Степень топологической пластичности будем оценивать мощностью множества топологических реализаций, отвечающих данной структурной модели.

Искусственная нейронная сеть содержит входной сумматор, последовательно связанный с нелинейным преобразователем сигналов и точкой ветвления для рассылки одного сигнала по нескольким адресам, а также линейную связь – синапс, для умножения входного сигнала на вес синапса, и в качестве входного сумматора используются изомерные ядра со временем жизни в возбужденном состоянии несколько лет, нелинейным преобразователем, в свою очередь, являются квантовые точки. Точкой ветвления является узел, выполненный из углеродных нанотрубок, а синапс представляет собой отдельно взятый переход одного изомерного ядра с низших на один из высших энергетических уровней под действием фотона большой энергии.

Искусственная нейронная сеть работает следующим образом (Рис. 5.11). Адаптивный сумматор 1, имеющий  $n+1$  вход и получающий на 0 вход постоянный сигнал, вычисляет скалярное произведение вектора входного сигнала  $x$  5 на вектор параметров  $\alpha$ . Нелинейный преобразователь сигнала 2 получает скалярный входной сигнал  $x$  5 и переводит его в  $\varphi(x)$ . Точка ветвления 3 служит для рассылки одного сигнала по нескольким адресам, она получает скалярный входной сигнал  $x$  5, и передает его всем своим выходам. Стандартный формальный нейрон составлен из входного сумматора 1, нелинейного преобразователя 2 и точки ветвления 3 на выходе. Линейная связь – синапс 4 умножает входной сигнал на вес синапса  $\alpha$ . Нелинейным преобразователем 2 являются квантовые точки 6, точкой ветвления 3 является узел 7 выполненный из углеродных нанотрубок 8, а синапс представляет собой отдельно взятый переход одного изомерного ядра 5 с низших на один из высших энергетических уровней под действием фотона 9 большой энергии.

Применение предложенной модели технического решения искусственной нейронной сети позволяет снизить физические габариты при реализации нанoeлектронных схем и компонентов, что связано с повышением производительности системы.

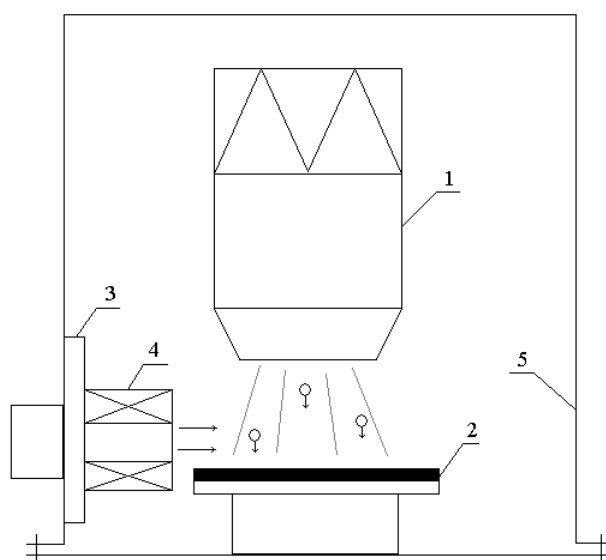


**Рис. 5.11.** Искусственная нейронная сеть

### 5.3.2. Устройство для получения углеродных пленок

В основу разработки положена задача снизить толщину при реализации получаемых углеродных пленок на наноуровне.

Устройство для получения углеродных пленок (рис. 5.12) содержит молекулярный источник 1, установленный с возможностью воздействия на подложку 2, выполненную из полупроводникового материала; между молекулярным источником 1 и подложкой установлена система подачи инертного газа 3, выполненная в виде тепловой трубы 4, связанной с криопанелью 5.



**Рис. 5.12.** Устройство для получения углеродных пленок

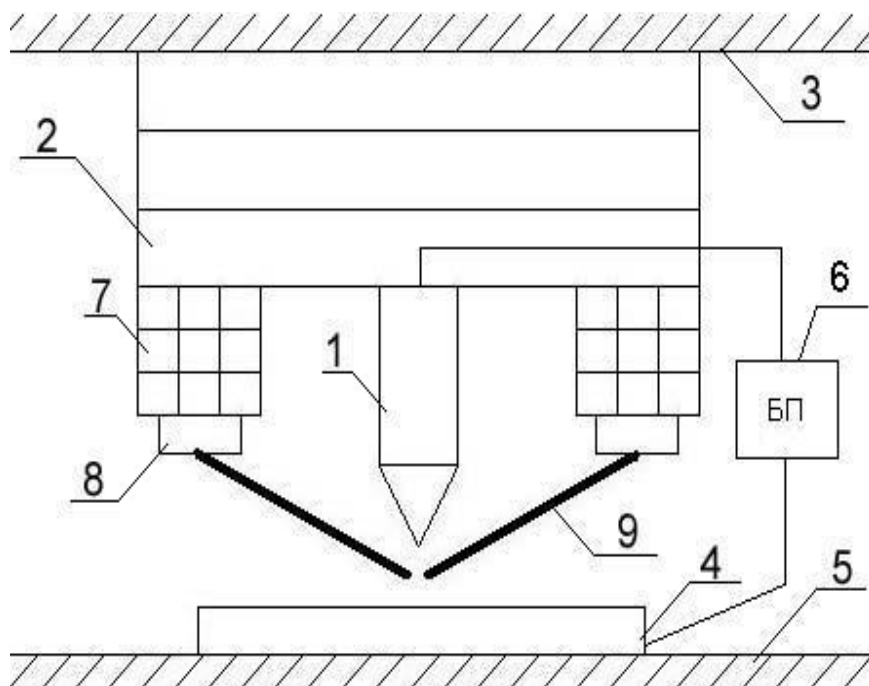
Устройство для получения углеродных пленок работает следующим образом .

При подаче напряжения на молекулярный источник 1, инертный газ поданный системой подачи 3 ионизируется и на подложке 2 выполненной из полупроводникового материала осаждается на подложке 2 в виде пленки наноразмерной толщины, технология осаждения изолирована от воздействия внешней среды криопанелью 5, а рабочая среда создается с помощью тепловой трубы 4.

Применение устройства для получения углеродных пленок позволяет получать углеродные пленки наноразмерной толщины.

### 5.3.3. Устройство для получения нанодорожек

В основу разработки положена задача обеспечить возможность оперативного управления процессом получения различных нанодорожек.



*Рис. 5.13. Устройство для получения нанодорожек*

Устройство для получения нанодорожек (Рис. 5.13) содержит зонд 1, жестко закреплённый на основном пьезоприводе 2, который установлен на

платформе 3, подложку 4, установленную на основании 5; и блок питания 6, электрически связанный с зондом 1 и подложкой 4; два дополнительных, двухкоординатных пьезопривода 7, жёстко закреплённых на торце основного пьезопривода 2, ёмкости 8, с рабочим веществом, установленные на торцах двухкоординатных пьезоприводов 7; а также капиллярные трубки 9, свободные торцы которых расположены вблизи острия зонда.

Устройство для получения нанодорожек работает следующим образом.

С помощью основного пьезопривода устанавливается туннельный зазор между зондом и подложкой. С помощью дополнительных, двухкоординатных пьезоприводов устанавливаются в исходное положение сопла с рабочим веществом. В туннельный зазор подается рабочее вещество из сопел 9, количество сопел, а, следовательно, и рабочих веществ не ограничено. При подаче разности потенциалов между зондом и подложкой, атомы рабочего вещества проникают в глубь подложки. В результате чего появляется возможность наносить мозаичную структуру на поверхность подложки.

Применение устройства для получения нанодорожек позволяет наносить мозаичную структуру на поверхность подложки.

### 5.3.4. Устройство наноперемещений

В основу разработки положена задача повысить быстродействие за счет нитевидного зонда и линий напряженности между обкладками конденсатора, выполняющих роль направляющих.

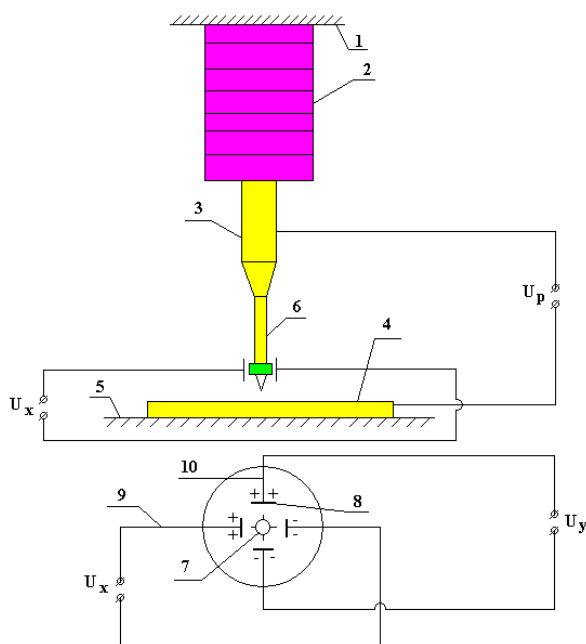


Рис. 5.14. Устройство наноперемещений

Устройство наноперемещений (рис. 5.14) содержит неподвижное основание 1, установленный на нем пьезопривод 2, на торце которого закреплен зонд 3 с возможностью электрического взаимодействия с подложкой 4, установленный на платформу 5, а также периферийную часть зонда 6 выполненного нитевидным, диаметром  $(0,05 \div 0,1)$  мм, с крестообразными закрепленными на ней электропроводящими пластинами 7 таким образом, что острие зонда 3 расположено ниже электропроводящих пластин 7 на  $(0,5 \div 1,0)$  мм, которые, в свою очередь установлены между обкладками 8 двух конденсаторов 9, 10, расположенными взаимноперпендикулярно, с возможностью независимой подачи электрического напряжения на их обкладки 8.

Устройство наноперемещений работает следующим образом .

С помощью пьезопривода периферийная часть зонда выполненная нитевидной, диаметром  $(0,05 \div 0,1)$  мм, устанавливается в зазор между электропроводящими пластинами, до возникновения туннельного тока, таким образом, что острие зонда расположено ниже электропроводящих пластин 7 на  $(0,5 \div 1,0)$  мм, пластины же свою очередь, установлены между обкладками 8 двух конденсаторов 9, 10, расположенными взаимноперпендикулярно, с возможностью независимой подачи электрического напряжения на их обкладки, при подаче разности потенциалов между зондом и подложкой, атомы рабочего вещества проникают в глубь подложки, а зонд, в свою очередь имеющий нитевидную форму, диаметром  $(0,05 \div 0,1)$  мм, позволяет существенно повысить быстродействие за счет своих малых габаритов.

Применение устройства наноперемещений в электронике позволяет повысить быстродействие за счет нитевидного зонда и линий напряженности между обкладками конденсатора, выполняющих роль направляющих, при изготовлении наносхем.

### **Заключение**

Развитие ИНС вызвало немало энтузиазма и критики. Некоторые сравнительные исследования оказались оптимистичными, другие - пессимистичными. Для многих задач, таких как распознавание образов, пока не создано доминирующих подходов. Выбор лучшей технологии должен диктоваться природой задачи. Нужно пытаться понять возможности, предпосылки и область применения различных подходов и максимально использовать их дополнительные преимущества для дальнейшего развития интеллектуальных систем. Подобные усилия могут привести к синергетическому подхо-



ду, который объединяет ИНС с другими технологиями для существенного прорыва в решении актуальных проблем. Индивидуальные модули важны, но мы также нуждаемся в методологии интеграции. Ясно, что взаимодействие и совместные работы исследователей в области ИНС и других дисциплин позволят не только избежать повторений, но и (что более важно) стимулируют и придают новые качества развитию отдельных направлений.

### Литература

1. DARPA Neural Network Study, AFCEAInt'l Press, Fairfax, Va., 1988.
2. J. Hertz, A. Krogh, and R.G. Palmer, Introduction to the Theory of Neural Computation, Addison-Wesley, Reading, Mass., 1991.
3. S. Haykin, Neural Networks: A Comprehensive Foundation, MacMillan College Publishing Co., New York, 1994.
4. W.S. McCulloch and W. Pitts, "A logical Calculus of Ideas Immanent in Nervous Activity", Bull. Mathematical Biophysics, Vol. 5, 1943, pp. 115-133.
5. Rosenblatt, "Principles of Neurodynamics", Spartan Books, New York, 1962.
6. M. Mimmsky and S. Papert, "Perceptrons: An Introduction to Computational Geometry", MIT Press, Cambridge, Mass., 1969.
7. J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", in Proc. National Academy of Sciences, USA 79, 1982, pp. 2554-2558.
8. P. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences", Phd Thesis, Dept. of Applied Mathematic
9. K.Alchalabi, D.Zimin, G.Kostorz, and H.Zogg. Phys.Rev.Lett., v.90, 026104 (2003).
10. Matthew Pelton, Charles Santori, Jelena Vuckovic et al. Phys.Rev.Lett., v.89, 233602 (2002).
11. Е.Н. Ивашов и др. Устройство для получения нанодорожек. Патент РФ на ПМ № 42696.- Опубликована 10.12.2004 Б. №34.
12. Производство тонкопленочных структур в электронном машиностроении: Учебник для вузов в 2-х томах. Т.1/ Н.В.Василенко, Е.Н.Ивашов, С.В.Степанчиков и др. – М.: Машиностроение, 2006 – 462с.



## **Глава 6. Квантовые наноразмерные структуры для систем кодирования и криптографии**

### **Введение**

Проблема безопасности информационных технологий возникла на пересечении двух активно развивающихся и, наверное, самых передовых в плане использования технических достижений направлений - безопасности технологий и информатизации. Сама проблема безопасности, конечно, не является новой, ведь обеспечение собственной безопасности - задача первостепенной важности для любой системы независимо от ее сложности и назначения будь то социальное образование, биологический организм или система обработки информации.

Жизнь современного общества немыслима без современных информационных технологий. Компьютеры обслуживают банковские системы, контролируют работу атомных реакторов, распределяют энергию, следят за расписанием поездов, управляют самолетами, космическими кораблями. Компьютерные сети и телекоммуникации определяют надежность и мощность систем обороны и безопасности страны. Компьютеры обеспечивают хранение информации, ее обработку и предоставление потребителям, реализуя таким образом информационные технологии.

Однако именно высокая степень автоматизации порождает риск снижения безопасности (личной, информационной, государственной, и т.п.). Доступность и широкое распространение информационных технологий, ЭВМ делает их чрезвычайно уязвимыми по отношению к деструктивным воздействиям в том числе и информационным. Тому есть множество примеров. Информационная безопасность, а главное надежность и достоверность информации становятся сегодня все более и более определяющими параметрами при взаимодействии технологических систем друг с другом.

Таким образом, чтобы быть защищенной, система должна успешно противостоять многочисленным и разнообразным угрозам безопасности, действующим в пространстве современных информационных технологий, и главным образом тем, которые носят целенаправленный характер.

На каждом этапе развития существуют инструментальные и экономические ограничения выражающиеся, в конце концов, в уровне и качестве продукции на данном этапе развития цивилизации. Прогресс в познании строения вещества неукоснительно приводит к созданию новых научных открытий и порождаемых ими новых научных технологиях.

С появлением нанотехнологий появилась техническая возможность сдвинуть ограничения на пространственное разрешение измерительных и исполнительных инструментов в нанометровую и субнанометровую область размеров, что и создало предпосылки развитию в направлениях

нанотехнологии, молекулярной нанотехнологии, наноэлектроники, базирующихся на возможности оперировать с веществом на уровне молекул, молекулярных кластеров и отдельных атомов.

Прогресс в познании строения вещества неразрывно связан с возможностью визуализации описывающих его параметров с максимально осуществимым пространственным и временным разрешением. Следующим шагом познания является попытка использования полученных знаний для построения новых функциональных структур максимально возможной информационной мощности, улучшения качества производимых продуктов, создания новых технологий.

Одним из наиболее мощных средств для исследования и проектирования технических систем является моделирование. Использование моделирования, начиная с ранних стадий, и постепенное накопление информации за счет уточнения и детализации модели позволяет говорить о расширяемой адаптивной модели всего цикла проектирования. Соответственно, при анализе различных свойств объекта проектирования (ОП) модельное представление должно формироваться наиболее подходящим для этой цели образом, независимо от конкретного процесса или этапа проектирования, и сохранять все требуемые свойства проектируемого объекта.

Для развития субмикронной технологии и нанотехнологии, в отличие от традиционной технологии, характерен "индивидуальный" подход, при котором внешнее "управление" достигает отдельных атомов и молекул, что позволяет создавать из них как "бездефектные" материалы с принципиально новыми физико-химическими свойствами, так и новые классы устройств с характерными нанометровыми размерами – наноразмерные структуры.

Одним из направлений решения этой проблемы является создание и развитие автоматизированных систем проектирования различных нанотехнологических процессов, в том числе формирование элементов наноразмерных структур на основе квантовых точек для систем кодирования и криптографии.

Инструментальный базис нанотехнологий, позволяющий ученым и исследователям не только визуализировать атомные структуры, но и манипулировать отдельными атомами и строить новые молекулы, основан на использовании так называемого эффекта туннелирования электронов. Его применение на вершинах зондов специальных конструкций позволило достичь высокой пространственной разрешающей способности управления атомно-молекулярными реакциями в отличие от известных групповых технологий осаждения материалов, методов оптической литографии, эпитаксии, а также электронной литографии, где высокая энергия фокусируемых электронов приводит к значительному разрушению используемых материалов.

Поэтому разработка элементов автоматизированной системы проектирования процесса формирования наноразмерных структур для систем

кодирования и криптографии в туннельно-зондовой нанотехнологии является задачей актуальной и своевременной.

## **6.1. Основные концепции построения «защищенных» систем**

### **6.1.1. Понятие «защищенная система» - определение и свойства**

Многие полагают, что защищенная технологическая система - это система обработки информации, в состав которой включен тот или иной набор средств защиты. С нашей точки зрения это неправильный подход, т. к. наличие средств защиты является лишь необходимым условием и не может рассматриваться в качестве критерия защищенности системы от реальных угроз, поскольку безопасность не является абсолютной характеристикой и может рассматриваться только относительно некоторой среды, в которой действуют определенные угрозы. Поэтому мы считаем, что:

*защищенная система обработки информации для определенных условий эксплуатации обеспечивает безопасность (конфиденциальность и целостность) обрабатываемой информации и поддерживает свою работоспособность в условиях воздействия на нее заданного множества угроз [1].*

Взяв за основу предложенное определение, рассмотрим свойства, которыми должна обладать защищенная система.

Как и все автоматизированные (компьютерные) системы обработки информации, защищенные системы решают задачу автоматизации некоторого процесса обработки информации. Под процессом обработки информации мы понимаем действия, связанные с ее хранением, преобразованием и передачей. В дополнение к этому кроме традиционных свойств, которыми обладают автоматизированные системы - надежности, эффективности, удобства использования и т. д., защищенная система обработки информации должна обладать еще одним - свойством безопасности, которое является для нее самым главным. Наконец, поскольку проблема безопасности компьютерных систем изучается и прорабатывается уже достаточно давно, защищенная система должна соответствовать сложившимся требованиям и представлениям. Кроме того, необходимо обеспечить возможность сопоставления параметров и характеристик защищенных систем для того, чтобы их можно было сравнивать между собой.

Таким образом, под защищенной системой обработки информации предлагается понимать систему, которая обладает следующими тремя свойствами:

- осуществляет автоматизацию некоторого процесса обработки конфиденциальной информации, включая все аспекты этого процесса связанные с обеспечением безопасности обрабатываемой информации;
- успешно противостоит угрозам безопасности, действующим в определенной среде;

- соответствует требованиям и критериям стандартов информационной безопасности.

Предложенный подход к определению понятия "защищенная система" отличается от существующих в первую очередь тем, что рассматривает проблему обеспечения безопасности систем как лежащую на стыке двух направлений: автоматизации обработки информации и общей безопасности. Это дает возможность объединить задачи автоматизации обработки конфиденциальной информации и разработки средств защиты в одну проблему создания защищенных информационных систем и процессе ее решения применять методы и технологии, разработанные как в той, так и в другой области.

### **6.1.2. Стандарты безопасности «защищенных систем»**

Безопасность является качественной характеристикой системы, ее нельзя измерить в каких-либо единицах, более того, нельзя даже с однозначным результатом сравнивать безопасность двух систем - одна будет обладать лучшей защитой в одном случае, другая - в другом. Кроме того, у каждой группы специалистов, занимающихся проблемами безопасности информационных технологий, имеется свой взгляд на безопасность и средства ее достижения, а, следовательно, и свое представление о том, что должна представлять собой защищенная система. Разумеется, любая точка зрения имеет право на существование и развитие, но для того, чтобы объединить усилия всех специалистов в направлении конструктивной работы над созданием защищенных систем все таки необходимо определить, что является целью исследований, что мы хотим получить в результате и чего в состоянии достичь?

Для ответа на эти вопросы и согласования всех точек зрения на проблему создания защищенных систем разработаны и продолжают разрабатываться стандарты информационной безопасности. Это документы, регламентирующие основные понятия и концепции информационной безопасности на государственном или межгосударственном уровне. Достигнут существенный прогресс, закрепленный в новом поколении документов разработки 90-годов.

Наиболее значимыми стандартами информационной безопасности являются (в хронологическом порядке):

- "Критерии безопасности компьютерных систем министерства обороны США" [2];
- Руководящие документы Гостехкомиссии России [3,4,5,6,7] (только для нашей страны);
- "Европейские критерии безопасности информационных технологий" [8];

- "Федеральные критерии безопасности информационных технологий США"[9];
- "Канадские критерии безопасности компьютерных систем" [10];
- "Единые критерии безопасности информационных технологий" [11].

### 6.1.3. Анализ существующих стандартов информационной безопасности

Главная задача стандартов информационной безопасности - согласовать позиции и цели производителей, потребителей и аналитиков-квалификаторов в процессе создания и эксплуатации продуктов информационных технологий. Каждая из перечисленных категорий специалистов оценивает стандарты и содержащихся в них требования и критерии по своим собственным параметрам.

Проведем сравнительный анализ существующих стандартов безопасности. Несмотря на то, что практически каждый из стандартов представляет оригинальный подход к определению понятия безопасной системы обработки информации, существует ряд понятий и концепций используемых всеми стандартами.

В качестве обобщенных показателей, характеризующих стандарты информационной безопасности и имеющих значение для всех трех сторон, предлагается использовать универсальность, гибкость, гарантированность, реализуемость и актуальность.

Классификация рассмотренных стандартов информационной безопасности по предложенным показателям приведена ниже:

#### Сопоставление стандартов информационной безопасности

Стандарты безопасности	Показатели сопоставления стандартов информационной безопасности				
	Универсальность	Гибкость	Гарантированность	Реализуемость	Актуальность
Оранжевая книга (1983 г.)	ограниченная	ограниченная	ограниченная	высокая (за исключением класса А)	Умеренная
Европейские критерии (1986 г.)	умеренная	умеренная	умеренная	высокая	Умеренная
Документы ГТК (1992 г.)	ограниченная	ограниченная	отсутствует	высокая	Ограниченная
Федеральные критерии (1992 г.)	высокая	отличная	достаточная	высокая	Высокая
Канадские критерии (1993 г.)	умеренная	достаточная	достаточная	достаточная	Достаточная
Единые критерии (1999 г.)	превосходная	превосходная	превосходная	превосходная	Превосходная

Степень соответствия стандартов предложенным показателям определяется по следующей качественной шкале:

- ограниченное соответствие - недостаточное соответствие, при применении стандарта возникают существенные трудности;
- умеренное соответствие - минимальное соответствие, при применении

- стандарта в большинстве случаев существенных трудностей не возникает;
- достаточное соответствие - удовлетворительное соответствие, при применении стандарта в большинстве случаев не возникает никаких трудностей, однако эффективность предлагаемых решений не гарантируется;
  - высокое соответствие - стандарт предлагает специальные механизмы и процедуры, направленные на улучшение данного показателя, применение которых позволяет получать достаточно эффективные решения;
  - превосходное соответствие - улучшение данного показателя рассматривалось авторами стандарта в качестве одной из основных целей его разработки, что обеспечивает эффективность применения предложенных решений.

Представленный анализ стандартов информационной безопасности и основных тенденций их развития позволяет сделать следующие выводы:

- 1) Развитие стандартов привело к отказу от единой шкалы ранжирования требований и критериев, замене их множеством независимых частных показателей и введению частично упорядоченных шкал.
- 2) Неуклонное возрастание роли требований адекватности реализации защиты и политики безопасности свидетельствует о тенденции преобладания "качества" обеспечения защиты над ее "количеством".
- 3) Определение ролей производителей, потребителей и экспертов по квалификации ИТ-продуктов и разделение их функций в процессе создания защищенных систем обработки информации свидетельствует о полноправной интеграции стандартов обеспечения безопасности в сферу информационных технологий.
- 4) Сложившееся на основе современных стандартов разделение ролей участников процесса создания и эксплуатации защищенных систем, применение соответствующих механизмов и технологий привело к сбалансированному распределению ответственности между всеми участниками процесса.
- 5) Современные тенденции интеграции информационных технологий и стремление к созданию безопасного всемирного информационного пространства привели к необходимости интернационализации стандартов информационной безопасности.



## **6.2. Моделирование «защищенных» систем**

### **6.2.1. Формальные модели безопасности**

Под *политикой безопасности* понимается совокупность норм и правил, регламентирующих процесс обработки информации, выполнение которых обеспечивает защиту от определенного множества угроз и составляет необходимое (а иногда и достаточное) условие безопасности системы. Формальное выражение политики безопасности называют *формальной моделью* политики безопасности. Формальные модели необходимы и используются достаточно широко, потому что только с их помощью можно доказать безопасность системы опираясь при этом на объективные и неопровержимые постулаты математической теории, а так же позволяют обосновать жизнеспособность системы и определяют базовые принципы ее архитектуры и используемые при ее построении технологические решения.

Основная цель создания политики безопасности информационной системы и описания ее в виде формальной модели - это определение условий, которым должно подчиняться поведение системы, выработка критерия безопасности и проведение формального доказательства соответствия системы этому критерию при соблюдении установленных правил и ограничений.

Кроме того, формальные модели безопасности позволяют решить еще целый ряд задач, возникающих в ходе проектирования, разработки и сертификации защищенных систем, поэтому их используют не только теоретики информационной безопасности, но и другие категории специалистов, участвующих в процессе создания и эксплуатации защищенных информационных систем (производители, потребители, эксперты-квалификаторы).

Среди моделей политик безопасности можно выделить два основных класса: дискреционные (произвольные) и мандатные (нормативные). Мы рассмотрим наиболее распространенные политики произвольного управления доступом, в основе которых лежат модель Харрисона-Руззо-Ульмана, модель типизированной матрицы доступа, фундаментальная нормативная модель безопасности Белла-ЛаПадулы.

### **6.2.2. Дискреционная модель Харрисона-Руззо-Ульмана**

Модель безопасности Харрисона-Руззо-Ульмана [12], являющаяся классической дискреционной моделью, реализует произвольное управление доступом субъектов к объектам и контроль за распространением прав доступа.

В рамках этой модели система обработки информации представляется в виде совокупности активных сущностей - субъектов (множество  $S$ ), которые осуществляют доступ к информации, пассивных сущностей - объектов

(множество  $O$ ), содержащих защищаемую информацию, и конечного множества прав доступа  $R = \{r_1, \dots, r_n\}$ , означающих полномочия на выполнение соответствующих действий (например, *чтение, запись, выполнение*).

Причем для того, чтобы включить в область действия модели и отношения между субъектами, принято считать, что все субъекты одновременно являются и объектами -  $S \subset O$ . Поведение системы моделируется с помощью понятия состояния. Пространство состояний системы образуется декартовым произведением множеств составляющих ее объектов, субъектов и прав —  $O \times S \times R$ . Текущее состояние системы  $Q$  в этом пространстве определяется тройкой, состоящей из множества субъектов, множества объектов и матрицы прав доступа  $M$ , описывающей текущие права доступа субъектов к объектам, -  $Q=(S,O,M)$ . Строки матрицы соответствуют субъектам, а столбцы - объектам, поскольку множество объектов включает в себя множество субъектов, матрица имеет вид прямоугольника. Любая ячейка матрицы  $M[s,o]$  содержит набор прав субъекта  $S$  к объекту  $o$ , принадлежащих множеству прав доступа  $R$ . Поведение системы во времени моделируется переходами между различными состояниями. Переход осуществляется путем внесения изменений в матрицу  $M$  с помощью команд следующего вида:

```

command  $\alpha(x_1, \dots, x_k)$ 
  if  $r_1$  in  $M[x_{s1}, x_{o1}]$  and (условия выполнения команды)
     $r_2$  in  $M[x_{s2}, x_{o2}]$  and
    .
    .
     $r_m$  in  $M[x_{sm}, x_{om}]$  and
  then
     $op_1, op_2 \dots op_n$  (операции, составляющие команду)

```

Здесь  $\alpha$  - имя команды;  $x_i$  - параметры команды, являющиеся идентификаторами субъектов и объектов,  $S_i$  и  $O_i$  - индексы субъектов и объектов в диапазоне от 1 до  $k$ ;  $op_i$  - элементарные операции. Элементарные операции, составляющие команду, выполняются только в том случае, если все условия, означающие присутствие указанных прав доступа в ячейках матрицы  $M$ , являются истинными. В классической модели допустимы только следующие элементарные операции:

```

enter  $r$  into  $M[s,o]$  (добавление субъекта  $s$  права  $r$  для объекта  $o$ )
delete  $r$  from  $M[s,o]$  (удаление субъекта  $s$  права  $r$  для объекта  $o$ )
create subject  $s$  (создание нового субъекта  $s$ )
create object  $o$  (создание нового объекта  $o$ )
destroy subject  $s$  (удаление существующего субъекта  $s$ )
destroy object  $o$  (удаление существующего объекта  $o$ )

```

Применение любой элементарной операции  $op$  в системе, находящейся в состоянии  $Q=(S,O,M)$  влечет за собой переход в другое

состояние  $Q'=(S',O',M')$ , которое отличается от предыдущего состояния  $Q$  по крайней мере одним компонентом. Выполнение базовых операций приводит к следующим изменениям в состоянии системы:

**enter**  $r$  **into**  $M[s,o]$  (где  $s \in S, o \in O$ )

$$O' = O$$

$$S' = S$$

$$M'[x_s, x_o] = M[x_s, x_o], \text{ если } (x_s, x_o) \neq (s, o)$$

$$M'[s, o] = M[s, o] \cup \{r\}$$

**delete**  $r$  **from**  $M[s,o]$  (где  $s \in S, o \in O$ )

$$O' = O$$

$$S' = S$$

$$M'[x_s, x_o] = M[x_s, x_o], \text{ если } (x_s, x_o) \neq (s, o)$$

$$M'[s, o] = M[s, o] \setminus \{r\}$$

**create subject**  $s$  (где  $s \notin S$ )

$$O' = O \cup \{s\}$$

$$S' = S \cup \{s\}$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S \times O$$

$$M'[s, x_o] = \theta \text{ для всех } x_o \in O'$$

$$M'[s, x_s] = \theta \text{ для всех } x_s \in S'$$

**destroy subject**  $s$  (где  $s \in S$ )

$$O' = O \setminus \{s\}$$

$$S' = S \setminus \{s\}$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S' \times O'$$

**create object**  $o$  (где  $o \in O$ )

$$O' = O \cup \{o\}$$

$$S' = S$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S \times O$$

$$M'[x_s, o] = \theta \text{ для всех } x_s \in S'$$

**destroy object**  $o$  (где  $o \in O \setminus S$ )

$$O' = O \setminus \{o\}$$

$$S' = S$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S' \times O'$$

Операция **enter** вводит право  $r$  в существующую ячейку матрицы доступа. Содержимое каждой ячейки рассматривается как множество, т.е. если это право уже имеется, то ячейка не изменяется. Операция называется **enter** *монотонной*, поскольку она только добавляет права в матрицу доступа и ничего не удаляет. Действие операции **delete** противоположно действию операции **enter**. Она удаляет право из ячейки матрицы доступа, если оно там присутствует. Поскольку содержимое каждой ячейки рассматривается как

множество, **delete** не делает ничего, если удаляемое право отсутствует в указанной ячейке. Поскольку **delete** удаляет информацию из матрицы доступа, она называется *немонотонной* операцией. Операции **create subject** и **destroy subject** представляют собой аналогичную пару монотонной и немонотонной операции.

Заметим, что для каждой операции существует еще и предусловие ее выполнения: для того чтобы изменить ячейку матрицы доступа с помощью операций **enter** или **delete** необходимо, чтобы эта ячейка существовала, т. е. чтобы существовали соответствующие субъект и объект. Предусловиями операций создания **create subject/object**, является отсутствие создаваемого субъекта/объекта, операций удаления **destroy subject/object** - наличие субъекта/объекта. Если предусловие любой операции не выполнено, то ее выполнение безрезультатно.

Формальное описание системы  $Z,(G,R,C)$  состоит из следующих элементов:

1. конечный набор прав доступа  $R = \{r_1, \dots, r_n\}$ ;
2. конечные наборы исходных субъектов  $S_0 = \{s_1, \dots, s\}$  и объектов  $O_0 = \{o_1, \dots, o_m\}$ , где  $S_0 \subseteq O_0$ ;
3. исходная матрица доступа, содержащая права доступа субъектов к объектам –  $M_0$ ;
4. конечный набор команд  $C = \{\alpha_i(x_1, \dots, x_k)\}$ , каждая из которых состоит из условий выполнения и интерпретации в терминах перечисленных элементарных операций.

Поведение системы во времени моделируется с помощью последовательности состояний  $\{Q_i\}$ , в которой каждое последующее состояние является результатом применения некоторой команды из множества  $C$  к предыдущему  $Q_{n+1} = C_n(Q_n)$ . Таким образом для заданного начального состояния только от условий команд из  $C$  и составляющих их операций зависит, сможет ли система попасть в то или иное состояние, или нет. Каждое состояние определяет отношения доступа, которые существуют между сущностями системы в виде множества субъектов, объектов и матрицы прав. Поскольку для обеспечения безопасности необходимо наложить запрет на некоторые отношения доступа, для заданного начального состояния системы должна существовать возможность определить множество состояний, в которые она сможет из него попасть. Это позволит задавать такие начальные условия (интерпретацию команд  $C$ , множества объектов  $O_0$ , субъектов  $S_0$  и матрицу доступа  $M_0$ ), при которых система никогда не сможет попасть в состояния, нежелательные с точки зрения безопасности. Следовательно, для построения системы с предсказуемым поведением необходимо для заданных начальных условий получить ответ на вопрос: сможет ли некоторый субъект  $S$  когда-либо приобрести право доступа  $r$  для некоторого объекта  $O$ ?

Поэтому критерий безопасности модели Харрисона-Руззо-Ульмана формулируется следующим образом:

*Для заданной системы начальное состояние  $Q_0=(S_0, O_0, M_0)$  является безопасным относительно права  $r$ , если не существует применимой к  $Q_0$  последовательности команд, в результате которой право  $r$  будет занесено в ячейку матрицы  $M$ , в которой оно отсутствовало в состоянии  $Q_0$*

Смысл данного критерия состоит в том, что для безопасной конфигурации системы субъект никогда получит право  $r$  доступа к объекту, если он не имел его изначально. На первый взгляд такая формулировка кажется довольно странной, поскольку невозможность получения права  $r$  вроде бы влечет за собой отказ от использования команд, в которых присутствует операция **enter into**  $M[s,o]$ , однако это не так. Дело в том, что удаление субъекта или объекта приводит к уничтожению всех прав в соответствующей строке или в столбце матрицы, но не влечет за собой уничтожение самого столбца или строки и сокращение размеров матрицы. Следовательно, если в какой-то ячейке в начальном состоянии существовало право  $r$ , и после удаления субъекта или объекта, к которым относилось это право, ячейка будет очищена, но впоследствии в результате создания субъекта или объекта появится вновь, и в эту ячейку с помощью соответствующей команды **enter** снова будет занесено право  $r$ , то это не будет означать нарушения безопасности.

Из критерия безопасности следует, что для данной модели ключевую роль играет выбор значений прав доступа и их использование в условиях команд. Хотя модель не налагает никаких ограничений на смысл прав и считает их равнозначными, те из них, которые участвуют в условиях выполнения команд, фактически представляют собой не права доступа к объектам (как, например, чтение и запись), а права управления доступом, или права на осуществление модификации ячеек матрицы доступа. Таким образом, по сути дела данная модель описывает не только доступ субъектов к объектам, а распространение прав доступа от субъекта к субъекту, поскольку именно изменение содержания ячеек матрицы доступа определяет возможность выполнения команд, в том числе команд, модифицирующих саму матрицу доступа, которые потенциально могут привести к нарушению критерия безопасности.

Необходимо отметить, что с точки зрения практики построения защищенных систем модель Харрисона-Руззо-Ульмана является наиболее простой в реализации и эффективной в управлении, поскольку не требует никаких сложных алгоритмов, и позволяет управлять полномочиями пользователей с точностью до операции над объектом, чем и объясняется ее распространенность среди современных систем. Кроме того, предложенный в данной модели критерий безопасности является весьма сильным в практическом плане, поскольку позволяет гарантировать недоступность определенного объекта для субъекта, которому изначально не выданы соответствующие полномочия.

Однако, Харрисон, Руззо и Ульман доказали, что в общем случае не существует алгоритма, который может для произвольной системы, ее

начального состояния  $Q_0=(S_0, O_0, M_0)$  и общего права  $r$  решить, является ли данная конфигурация безопасной. Доказательство опирается на свойства машины Тьюринга, с помощью которой моделируется последовательность переходов системы из состояния в состояние.

Для того, чтобы можно было доказать указанный критерий, модель должна быть дополнена рядом ограничений [14]. Мы не будем утомлять читателя многостраничными математическими выкладками, а только отметим, что указанная задача является разрешимой в любом из следующих случаев:

- команды  $\alpha_i(x_1, \dots, x_k)$  являются монооперационными, т. е. состоят не более чем из одной элементарной операции;
- команды  $\alpha_i(x_1, \dots, x_k)$  являются одноусловными и монотонными, т. е. содержат не более одного условия и не содержат операций **destroy** и **delete**;
- команды  $\alpha_i(x_1, \dots, x_k)$  не содержат операций **create**.

Эти условия существенно ограничивают сферу применения модели, поскольку трудно представить себе реальную систему, в которой не будет происходить создания или удаления сущностей.

Таким образом дискреционная модель Харрисона-Руззо-Ульмана в своей общей постановке не дает гарантий безопасности системы, однако именно она послужила основой для целого класса моделей политик безопасности, которые используются для управления доступом и контроля за распространением прав во всех современных системах.

### 6.2.3. Типизованная матрица доступа

Другая дискреционная модель, получившая название "Типизованная матрица доступа" (Type Access Matrix — далее ТАМ) [13], представляет собой развитие модели Харрисона-Руззо-Ульмана, дополненной концепцией типов, что позволяет несколько смягчить те условия, для которых возможно доказательство безопасности системы.

Формальное описание модели ТАМ включает следующие элементы:

1. конечный набор прав доступа  $R = \{r_1, \dots, r_i\}$ ;
2. конечный набор типов  $T = \{t_1, \dots, t_g\}$ ;
3. конечные наборы исходных субъектов  $S_0 = \{s_1, \dots, s\}$  и объектов  $O_0 = \{o_1, \dots, o_m\}$ , где  $S_0 \subseteq O_0$ ;
4. матрица  $M$ , содержащая права доступа субъектов к объектам, и ее начальное состояние  $M_0$ ;
5. конечный набор команд  $C = \{\alpha_i(x_1, \dots, x_k)\}$ , включающий условия выполнения команд и их интерпретацию в терминах элементарных операций.

Тогда состояние системы описывается четверкой  $Q=(S,O,t,M)$ , где  $S$ ,  $O$ , и  $M$  обозначают соответственно множество субъектов, объектов и матрицу

доступа, а  $t: O \rightarrow T$  - функция, ставящая в соответствие каждому объекту некоторый тип.

Состояние системы изменяется с помощью команд из множества  $C$ . Команды ТАМ имеют тот же формат, что и в модели Харрисона-Руззо-Ульмана, но всем параметрам приписывается определенный тип:

```

command  $\alpha(x_1:t_1, \dots, x_k:t_k)$ 
  if  $r_1$  in  $M[x_{s1}, x_{o1}]$  and (условия выполнения команды)
     $r_2$  in  $M[x_{s2}, x_{o2}]$  and
    .
    .
     $r_m$  in  $M[x_{sm}, x_{om}]$  and
  then
     $op_1, op_2, \dots, op_n$  (операции, составляющие команду)
  
```

Перед выполнением команды происходит проверка типов фактических параметров, и, если они не совпадают с указанными в определении, команда не выполняется. Фактически, введение контроля типов для параметров команд приводит к неявному введению дополнительных условий, т. к. команды могут быть выполнены только при совпадении типов параметров. В модели используются следующие шесть элементарных операций, отличающихся от аналогичных операций модели Харрисона-Руззо-Ульмана только использованием типизованных параметров:

Монотонные операции	Немонотонные операции
<b>enter</b> $r$ <b>into</b> $M[s,o]$	<b>delete</b> $r$ <b>from</b> $M[s,o]$
<b>create subject</b> $s$ <b>of type</b> $t$	<b>destroy subject</b> $s$
<b>create object</b> $o$ <b>of type</b> $t$	<b>destroy object</b> $o$

Смысл элементарных операций совпадает со смыслом аналогичных операций из классической модели Харрисона-Руззо-Ульмана с точностью до использования типов:

```

enter  $r$  into  $M[s,o]$  (где  $s \in S, o \in O$ )
   $O' = O$ 
   $S' = S$ 
   $t'(o) = t(o)$  для всех  $o \in O$ 
   $M'[x_s, x_o] = M[x_s, x_o]$ , если  $(x_s, x_o) \neq (s, o)$ 
   $M'[s, o] = M[s, o] \cup \{r\}$ 
delete  $r$  from  $M[s,o]$  (где  $s \in S, o \in O$ )
   $O' = O$ 
   $S' = S$ 
  
```

$$t'(o) = t(o) \text{ для всех } o \in O$$

$$M'[x_s, x_o] = M[x_s, x_o], \text{ если } (x_s, x_o) \neq (s, o)$$

$$M'[s, o] = M[s, o] \setminus \{r\}$$

**create subject s of type  $t_s$**  (где  $s \notin S$ )

$$O' = O \cup \{s\}$$

$$S' = S \cup \{s\}$$

$$t'(o) = t(o) \text{ для всех } o \in O$$

$$t'(s) = t_s$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S \times O$$

$$M'[s, x_o] = \theta \text{ для всех } x_o \in O'$$

$$M'[s, x_s] = \theta \text{ для всех } x_s \in S'$$

**destroy subject s** (где  $s \in S$ )

$$O' = O \setminus \{s\}$$

$$S' = S \setminus \{s\}$$

$$t'(o) = t(o) \text{ для всех } o \in O'$$

$$t'(s) = \text{не определено}$$

$$M'[x_s, x_o]' = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S' \times O'$$

**create object o of type  $t_o$**  (где  $o \in O$ )

$$O' = O \cup \{o\}$$

$$S' = S$$

$$t'(O) = t(O) \text{ для всех } O \in O$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S \times O$$

$$M'[x_s, o] = \theta \text{ для всех } x_s \in S'$$

**destroy object o** (где  $o \in O \setminus S$ )

$$O' = O \setminus \{o\}$$

$$S' = S$$

$$t'(x_o) = t(x_o) \text{ для всех } x_o \in O'$$

$$t'(o) = \text{не определено}$$

$$M'[x_s, x_o] = M[x_s, x_o] \text{ для всех } (x_s, x_o) \in S' \times O'$$

Таким образом ТАМ является обобщением модели Харрисона-Руззо-Ульмана, которую можно рассматривать как частный случай ТАМ с одним единственным типом, к которому относятся все объекты и субъекты. Появление в каждой команде дополнительных неявных условий, ограничивающих область применения команды только сущностями соответствующих типов, позволяет несколько смягчить жесткие условия классической модели, при которых критерий безопасности является разрешимым.



В работе [14] Харрисон, Руццо и Ульман показали, что критерий безопасности дискреционной модели может быть доказан для систем, в которых все команды  $\alpha_i(x_1, \dots, x_k)$  являются одноусловными и монотонными. Строгий контроль соответствия типов позволяет смягчить требование одноусловности, заменив его ограничением на типы параметров команд, при выполнении которых происходит создание новых сущностей.

Для того чтобы сформулировать это ограничение определим отношения между типами. Пусть  $\alpha(x_1:t_1, x_2:c, \dots, x_k:t_k)$  - некоторая команда ТАМ. Будем говорить, что  $t_i$  является *дочерним типом* в  $\alpha$ , если в теле  $\alpha$  имеет место одна из следующих элементарных операций: **create subject**  $x_i$  **of type**  $t_i$  или **create object**  $x_i$  **of type**  $t_i$ . В противном случае будем говорить, что  $t_i$  является *родительским типом* в  $\alpha$ .

Заметим, что в одной команде тип может быть одновременно и родительским, и дочерним, например:

```

command      foo(s1:u, s2:u s3:w,o:b)
                create subject s2 of type u;
                create subject s3 of type v;
end

```

Здесь  $u$  является родительским типом относительно  $s_1$  и дочерним типом относительно  $s_2$ . Кроме того,  $w$  и  $b$  являются родительскими типами, а  $v$  - дочерним типом. Тогда можно описать взаимосвязи связи между различными типами с помощью графа, определяющего отношение "наследственности" между типами, устанавливаемые через операции порождения сущностей (объектов и субъектов). Такой граф называется *графом создания* и представляет собой направленный граф с множеством вершин  $T$ , в котором ребро от  $u$  к  $v$  существует тогда и только тогда, когда в системе имеется создающая команда, в которой  $u$  является родительским типом, а  $v$  - дочерним типом. Этот граф для каждого типа позволяет определить:

1. сущности каких типов должны существовать в системе, чтобы в ней мог появиться объект или субъект заданного типа;
2. сущности каких типов могут быть порождены при участии сущностей заданного типа.

Модель монотонной типизированной матрицы доступа (МТАМ) идентична ТАМ за исключением того, что в ней отсутствуют немонотонные элементарные операции **delete**, **destroy subject** и **destroy object**.

Реализация МТАМ, состоящая из множеств объектов, субъектов, типов, матрицы прав доступа и множества команд, называется *ациклической* тогда и только тогда, когда ее граф создания не содержит циклов, в противном случае говорят, что реализация *циклическая*. Например, граф создания для приведенной выше команды `foo`, содержит следующие ребра:  $\{(u, u), (u, v), (w, u), (w, v), (b, u), (b, v)\}$ . Реализация МТАМ, содержащая эту

команду, будет циклической, поскольку тип  $u$  является для нее одновременно и родительским и дочерним, что приводит к появлению на графе цикла ( $u, u$ ).

Доказано, что критерий безопасности, предложенный Харрисоном, Руццо и Ульманом, разрешим для ациклических реализаций МТАМ, и что требование одноусловности команд можно заменить требованием ациклическости графа создания [13]. Смысл этой замены состоит в том, что последовательность состояний системы должна следовать некоторому маршруту на графе создания, поскольку невозможно появление сущностей дочерних типов, если в системе отсутствуют сущности родительских типов, которые должны участвовать в их создании. Отсутствие циклов на графе создания позволяет избежать заикливания при доказательстве критерия безопасности, т. к. количество путей на графе без циклов является ограниченным.

Это означает, что поведение системы становится предсказуемым, поскольку в любом состоянии можно определить сущности каких типов могут появиться в системе, а каких - нет. Но, к сожалению, доказано, что в общем случае сложность проверки критерия безопасности для МТАМ является NP-трудной задачей, т. е. с ростом размерности задачи (количества объектов и субъектов) время на ее решение растет в степенной зависимости от ее размерности. Этот недостаток может быть преодолен с помощью *тернарной* ТАМ, в которой команды могут иметь не более трех параметров. Тернарная МТАМ является монотонной версией тернарной ТАМ. Для тернарной МТАМ доказательство безопасности радикально упрощается, поскольку при проверке условной части команды всегда используется только небольшой фрагмент матрицы доступа. Тернарная МТАМ по своим выразительным способностям эквивалентна МТАМ, несмотря на это, доказано, что безопасность ее ациклической реализации разрешима за время, зависящее от размера начальной матрицы доступа; полиномиально.

Следовательно введение строгого контроля типов в дискреционную модель Харрисона-Руццо-Ульмана позволило доказать критерий безопасности систем для более приемлемых ограничений, что существенно расширило область ее применения.

#### **6.2.4. Мандатная модель Белла-ЛаПадулы**

Система в модели безопасности Белла-ЛаПадулы [15], как и в модели Харрисона-Руццо-Ульмана, представляется в виде множеств субъектов  $S$ , объектов  $O$  (множество объектов включает множество субъектов,  $S \subset O$ ) и прав доступа  $read$  (чтение) и  $write$  (запись). В мандатной модели рассматриваются только эти два вида доступа, и, хотя она может быть расширена введением дополнительных прав (например, правом на добавление информации, выполнение программ и т.д.), все они будут

отображаться в базовые (чтение и запись). Использование столь жесткого подхода, не позволяющего осуществлять гибкое управление доступом, объясняется тем, что в мандатной модели контролируются не операции, осуществляемые субъектом над объектом, а потоки информации, которые могут быть только двух видов; либо от субъекта к объекту (запись), либо от объекта к субъекту (чтение).

Уровни безопасности субъектов и объектов задаются с помощью функции уровня безопасности  $F: S \cup O \rightarrow L$ , которая ставит в соответствие каждому объекту и субъекту уровень безопасности, принадлежащий множеству уровней безопасности  $L$ , на котором определена решетка  $\Lambda$  [16].

### Решетка уровней безопасности.

Решетка уровней безопасности  $\Lambda$  - это формальная алгебра  $(L, \leq, \bullet, \otimes)$ , где  $L$  - базовое множество уровней безопасности, а оператор  $\leq$  определяет частичное нестрогое отношение порядка для элементов этого множества, т.е. оператор  $\leq$  - антисимметричен, транзитивен и рефлексивен. Отношение  $\leq$  на  $L$ :

- 1) рефлексивно, если  $\forall a \in L: a \leq a$ ;
- 2) антисимметрично, если  $\forall a_1, a_2 \in L: (a_1 \leq a_2 \wedge a_2 \leq a_1) \Rightarrow a_1 = a_2$ ;
- 3) транзитивно, если  $\forall a_1, a_2, a_3 \in L: (a_1 \leq a_2 \wedge a_2 \leq a_3) \Rightarrow a_1 \leq a_3$ .

Другое свойство решетки состоит в том, что для каждой пары  $a_1$  и  $a_2$  элементов множества  $L$  можно указать единственный элемент *наименьшей верхней границы* и единственный элемент *наибольшей нижней границы*. Эти элементы также принадлежат  $L$  и обозначаются с помощью операторов  $\bullet$  и  $\otimes$  соответственно:

$$a_1 \bullet a_2 = a \Leftrightarrow a_1, a_2 \leq a \wedge \forall a' \in L: (a' \leq a) \Rightarrow (a' \leq a_1 \vee a' \leq a_2)$$

$$a_1 \otimes a_2 = a \Leftrightarrow a \leq a_1, a_2 \wedge \forall a' \in L: (a' \leq a_1 \wedge a' \leq a_2) \Rightarrow (a' \leq a)$$

Смысл этих определений заключается в том, что для каждой пары элементов (или множества элементов, поскольку операторы  $\bullet$  и  $\otimes$  транзитивны) всегда можно указать единственный элемент, ограничивающий ее сверху или снизу таким образом, что между ними и этим элементом не будет других элементов.

Функция уровня безопасности  $F$  назначает каждому субъекту и объекту некоторый уровень безопасности из  $L$ , разбивая множество сущностей системы на классы, в пределах которых их свойства с точки зрения модели безопасности являются эквивалентными. Тогда оператор  $<$  определяет направление потоков информации, то есть, если  $F(A) \leq F(B)$ , то информация может передаваться от элементов класса  $A$  элементам класса  $B$ .

Покажем, почему в модели Белла-ЛаПадулы для описания отношения доминирования на множестве уровней безопасности используется решетка.

Если информация может передаваться от сущностей класса А к сущностям класса В, а также от сущностей класса В к сущностям класса А, то классы А и В содержат одноуровневую информацию и с точки зрения безопасности эквивалентны одному классу (АВ). Поэтому для удаления избыточных классов необходимо, чтобы отношение  $\leq$  было антисимметричным.

Если информация может передаваться от сущностей класса А сущностям класса В, а также от сущностей класса В к сущностям класса С, то очевидно, что она будет также передаваться от сущностей класса А к сущностям класса С. Таким образом, отношение  $\leq$  должно быть транзитивным.

Так как класс сущности определяет уровень безопасности содержащейся в ней информации, то все сущности одного и того же класса содержат с точки зрения безопасности одинаковую информацию. Следовательно, нет смысла запрещать потоки информации между сущностями одного и того же класса. Более того, из чисто практических соображений нужно предусмотреть возможность для сущности передавать информацию самой себе. Следовательно, отношение  $\leq$  должно быть рефлексивным.

Покажем, что для любого множества сущностей должны существовать единственная наименьшая верхняя и наибольшая нижняя границы Множества соответствующих им уровней безопасности. Для пары сущностей  $x$  и  $y$ , обладающих уровнями безопасности  $a$  и  $b$  соответственно, обозначим наибольший уровень безопасности их комбинации как  $(a \bullet b)$ , при этом  $a \leq (a \bullet b)$  и  $b \leq (a \bullet b)$ . Тогда, если существует некоторый уровень  $c$  такой, что  $a \leq c$  и  $b \leq c$ , то должно иметь место отношение  $(a \bullet b) \leq c$ , поскольку  $(a \bullet b)$  - это минимальный уровень субъекта, для которого доступна информация как из  $x$ , так и из  $y$ . Следовательно,  $(a \bullet b)$  должен быть наименьшей верхней границей  $a$  и  $b$ . Аналогично обозначим наименьший уровень безопасности комбинации сущностей  $x$  и  $y$  как  $(a \otimes b)$ , при этом  $(a \otimes b) \leq a$  и  $(a \otimes b) \leq b$ . Тогда, если существует некоторый уровень  $c$  такой, что  $c \leq a$  и  $c \leq b$ , то должно иметь место отношение  $c \leq (a \otimes b)$ , поскольку  $(a \otimes b)$  - это максимальный уровень субъекта, для которого разрешена передача информации как в  $x$ , так и в  $y$ . Следовательно,  $(a \otimes b)$  должен быть наибольшей нижней границей  $a$  и  $b$ .

Использование решетки для описания отношений между уровнями безопасности позволяет использовать в качестве атрибутов безопасности (элементов множества  $L$ ) не только целые числа, для которых определено отношение "меньше или равно", но и более сложные составные элементы. Например, в государственных организациях достаточно часто в качестве атрибутов безопасности используется комбинация, состоящая из уровня безопасности, представляющего собой целое число, и набора категорий из некоторого множества. Такие атрибуты невозможно сравнивать с помощью арифметических операций, поэтому отношение доминирования  $\leq$

определяется как композиция отношения "меньше или равно" для уровней безопасности и отношения включения множеств  $\subseteq$  для наборов категорий. Причем это никак не сказывается на свойствах модели, поскольку отношения "меньше или равно" и "включение множеств" обладают свойствами антисимметричности, транзитивности и рефлексивности, и, следовательно, их композиция также будет обладать этими свойствами, образуя над множеством атрибутов безопасности решетку. Точно также можно использовать любые виды атрибутов и любое отношение частичного порядка, лишь бы их совокупность представляла собой решетку.

### Основная теорема безопасности Белла-ЛаПадулы

Система  $\Sigma (v_0, R, T)$  безопасна тогда и только тогда, когда:  
 начальное состояние  $v_0$  безопасно и для любого состояния  $v$ , достижимого из  $v_0$  путем применения конечной последовательности запросов из  $R$  таких, что  $T\{v, r\} = v^*$ ,  $v = (F, M)$  и  $v^* = (F^*, M^*)$  для каждого  $s \in S$  и  $o \in O$  выполняются следующие условия:

- если  $read \in M^*[s, o]$  и  $read \notin M[s, o]$ , то  $F^*(s) \geq F^*(o)$ ;
- если  $read \in M[s, o]$  и  $F^*(s) < F^*(o)$ , то  $read \notin M^*[s, o]$ ;
- если  $write \in M^*[s, o]$  и  $write \notin M[s, o]$ , то  $F^*(o) \geq F^*(s)$ ;
- если  $write \in M[s, o]$  и  $F^*(o) < F^*(s)$ , то  $write \notin M^*[s, o]$ .

Доказательство:

Необходимость. Если система безопасна, то состояние  $v_0$  безопасно по определению. Допустим, существует некоторое состояние  $v^*$ , достижимое из  $v_0$  путем применения конечного числа запросов из  $R$  и полученное путем перехода из безопасного состояния  $V$ :  $T(v, r) = v^*$ . Тогда, если при этом переходе нарушено хотя бы одно из первых двух ограничений, накладываемых теоремой на функцию  $T$ , то состояние  $v^*$  не будет безопасным по чтению, а если функция  $T$  нарушает хотя бы одно из последних двух условий теоремы, то состояние  $v^*$  не будет безопасным по записи. В любом случае при нарушении условий теоремы система небезопасна.

Достаточность. Проведем доказательство от противного. Предположим, что система небезопасна. В этом случае, либо  $v_0$  небезопасно, что явно противоречит условиям теоремы, либо должно существовать небезопасное состояние  $v^*$ , достижимое из безопасного  $v_0$  путем применения конечного числа запросов из  $R$ . В этом случае обязательно будет, иметь место переход  $T(v, r) = v^*$  при котором состояние  $v$  - безопасно, а состояние  $v^*$  - нет, однако четыре условия теоремы делают такой переход невозможным.

Таким образом, теорема утверждает, что система с безопасным начальным состоянием является безопасной тогда и только тогда, когда при любом переходе системы из одного состояния в другое не возникает никаких новых и не сохраняется никаких старых отношений доступа, которые будут

небезопасны по отношению к функции уровня безопасности нового состояния. Формально эта теорема определяет все необходимые и достаточные условия, которые должны быть выполнены для того, чтобы система, начав свою работу в безопасном состоянии, никогда не достигла небезопасного состояния.

### Безопасная функция перехода

Недостаток основной теоремы безопасности Белла-ЛаПадулы состоит в том, что ограничения, накладываемые теоремой на функцию перехода, совпадают с критериями безопасности состояния, поэтому данная теорема является избыточной по отношению к определению безопасного состояния. Кроме того, из теоремы следует только то, что все состояния, достижимые из безопасного состояния при определенных ограничениях, будут в некотором смысле безопасны, но при этом не гарантируется, что они будут достигнуты без потери свойства безопасности в процессе осуществления перехода. Поскольку у нас нет никаких определенных ограничений на вид функции перехода, кроме указанных в условиях теоремы, и допускается, что уровни безопасности субъектов и объектов могут изменяться, то можно представить такую гипотетическую систему (она получила название *Z*-системы [17]), в которой при попытке низкоуровневого субъекта прочитать информацию из высокоуровневого объекта будет происходить понижение уровня объекта до уровня субъекта и осуществляться чтение. Функция перехода *Z*-системы удовлетворяет ограничениям основной теоремы безопасности, и все состояния такой системы также являются безопасными в смысле критерия Белла-ЛаПадулы, но вместе с тем в этой системе любой пользователь сможет прочитать любой файл, что, очевидно, несовместимо с безопасностью в обычном понимании.

Следовательно, необходимо сформулировать теорему, которая бы не только констатировала, безопасность всех достижимых состояний для системы, соответствующей определенным условиям, но и гарантировала бы безопасность в процессе осуществления переходов между состояниями. Для этого необходимо регламентировать изменения уровней безопасности при переходе от состояния к состоянию с помощью дополнительных правил.

Такую интерпретацию мандатной модели осуществил Мак-Лин [18], предложивший свою формулировку основной теоремы безопасности, основанную не на понятии безопасного состояния, а на понятии безопасного перехода. При таком подходе функция уровня безопасности представляется с помощью двух функций, определенных на множестве субъектов и объектов:  $F_s: S \rightarrow L$  и  $F_o: O \rightarrow L$ .

Функция перехода  $T$  считается безопасной по чтению, если для любого перехода  $T(r,v)=v^*$ , выполняются следующие три условия:

если **read**  $\in M^*[s,o]$  и

**read**  $\notin M[s,o]$  то:

$$F_s^*(s) \geq F_o(o) \text{ и } F=F^*;$$

если  $F_s \neq F_s^*$  то:

$$M=M^*,$$

$$F_o=F_o^*,$$

для  $\forall s$  и  $o$ , для которых  $F_s^*(s) < F_o^*(o)$ , **read**  $\notin M[s,o]$ ;

если  $F_o \neq F_o^*$  то:

$$M=M^*,$$

$$F_s=F_s^*,$$

для  $\forall s$  и  $o$ , для которых  $F_s^*(s) < F_o^*(o)$ , **read**  $\notin M[s,o]$ ;

Функция перехода  $T$  считается безопасной по записи, если для любого перехода  $T(r,v) = v^*$  выполняются следующие три условия:

если **write**  $\in M^*[s,o]$  и

**write**  $\notin M[s,o]$  то:

$$F_o^*(o) \geq F_s(s) \text{ и } F=F^*;$$

если  $F_s \neq F_s^*$  то:

$$M=M^*,$$

$$F_o=F_o^*,$$

для  $\forall s$  и  $o$ , для которых  $F_s^*(s) > F_o^*(o)$ , **write**  $\notin M[s,o]$ ;

если  $F_o \neq F_o^*$  то:

$$M=M^*,$$

$$F_s=F_s^*,$$

для  $\forall s$  и  $o$ , для которых  $F_s^*(s) > F_o^*(o)$ , **write**  $\notin M[s,o]$ ;

Функция перехода является безопасной тогда и только тогда, когда она одновременно безопасна и по чтению и по записи.

Смысл введения перечисленных ограничений и их принципиальное отличие от условий теоремы Белла-ЛаПадулы состоит в следующем: нельзя изменять одновременно более одного компонента состояния системы - в процессе перехода либо возникает новое отношение доступа, либо изменяется уровень объекта, либо изменяется уровень субъекта.

Следовательно, *функция перехода является безопасной тогда и только тогда, когда она изменяет только один из компонентов состояния и изменения не приводят к нарушению безопасности системы.*

Поскольку безопасный переход из состояния  $v$  в состояние  $v^*$  позволяет изменяться только одному элементу из  $v$  и так как этот элемент может быть изменен только способами, сохраняющими безопасность состояния, была доказана следующая теорема о свойствах безопасной системы [17]:

**Теорема безопасности Мак-Лина.** Система безопасна в любом состоянии и в процессе переходов между ними, если ее начальное состояние является безопасным, а ее функция перехода удовлетворяет критерию Мак-Лина.

Обратное утверждение неверно. Система может быть безопасной по определению Белла-ЛаПадулы, но не иметь безопасной функции перехода, о чем свидетельствует рассмотренный пример Z-системы.

Такая формулировка основной теоремы безопасности предоставляет в распоряжение разработчиков защищенных систем базовый принцип их построения, в соответствии с которым для того, чтобы обеспечить безопасность системы как в любом состоянии, так и в процессе перехода между ними, необходимо реализовать для нее такую функцию перехода, которая соответствует указанным условиям.

### **6.2.5. Моделирование квантовых наноразмерных структур для систем кодирования и криптографии**

В качестве элемента системы кодирования и криптографии выступает изомерная квантовая точка. При описании модели физического явления удобно использовать оболочечную модель ядра. В оболочечной модели ядра принимается, что энергетическая структура (уровни энергии нуклонов) ядра подобна энергетической структуре электронной оболочки атома. Сильное взаимодействие нуклонов в ядре и малый радиус этого взаимодействия позволяет рассматривать нуклоны движущимися независимо друг от друга в поле, обладающем сферически симметричным потенциалом. При этом нуклоны могут находиться в различных энергетических состояниях. Основному состоянию ядра должно соответствовать заполнение всех нижних уровней. Потеря нуклоном энергии при межнуклонных столкновениях не может перевести его в более низкое состояние, ибо все они заняты в соответствии с принципом Паули. Это приводит к тому, что длина свободного пробега нуклона в невозбужденном ядре становится больше радиуса ядра. Это означает возможность рассматривать нуклоны в рамках данной модели невзаимодействующими и несталкивающимися. Движение невзаимодействующих нуклонов в поле сферического потенциала, где орбитальный момент импульса является интегралом движения, характеризуется тем, что всем  $2l+1$  возможным ориентациям вектора  $l$  соответствует одинаковый энергетический уровень. На этом уровне размещаются  $2(2l+1)$  нуклонов данного типа. Таким образом, в оболочечной модели нуклоны располагаются в определенном количестве на энергетических *нуклонных оболочках*. Каждый нуклон характеризуется индивидуальной волновой функцией и индивидуальными квантовыми числами  $n$  и  $l$ . Существуют две системы нуклонных состояний - одна для протонов, другая для нейтронов; обе системы уровней заполняются нуклонами независимо друг от друга. Ядра, имеющие только заполненные *нуклонные оболочки*, должны обладать повышенной устойчивостью (проявляющейся, например, в их большей распространенности в природе), а также должны иметь сферически симметричное распределение заряда.



Порядок заполнения нуклонных оболочек с ростом  $A$  сходен с порядком заполнения электронных оболочек с ростом  $Z$ . Ввиду сильной спин-орбитальной связи все уровни с  $l \neq 0$  расщепляются на два подуровня с  $j = l \pm 1/2$ , заполняющихся независимо.

Предсказания оболочечной модели, в общем, соответствуют действительности. Наиболее устойчивым по сравнению с соседними ядрами являются ядра со значениями  $N$  или  $Z$ , равными **2, 8, 20, 28, 50, 82, 126** и **152**. Эти числа называются *магическими*. Распространенность в природе таких ядер наиболее велика, а квадрупольные моменты их близки к нулю. Ядра, у которых магическими числами являются и  $N$  и  $Z$ , называются *дважды магическими*. Эти ядра ( ${}^4_2\text{He}$ ,  ${}^{16}_8\text{O}$ ,  ${}^{40}_{20}\text{Ca}$ ,  ${}^{208}_{82}\text{Pb}$ ) особой устойчивостью, проявляющейся, в частности, в том, что они являются наиболее распространенными в природе изотопами этих элементов.

*Гамма-излучением* называется жесткое электромагнитное излучение, энергия которого высвобождается при переходах ядер из возбужденного в основное или в менее возбужденное состояние, а также при ядерных реакциях.

В первом случае энергия гамма-квантов равна разности энергий конечного и начального уровней ядра. В каждом акте перехода ядро излучает один гамма-квант. В связи с дискретностью энергетических уровней ядра гамма-излучение имеет линейчатый спектр. Частоты гамма-квантов связаны с разностью энергий условием частот Бора.

При испускании ядром гамма-кванта само ядро, вследствие закона сохранения импульса, приобретает противоположно направленный импульс (*отдача*). Если ядра, испускающие гамма-кванты, находятся в твердом теле, то спектр гамма-лучей состоит из двух компонент:

а) компоненты с естественной шириной гамма-линии  $\Gamma$ , определяемой временем жизни ядер в данном возбужденном состоянии, с энергией  $E$ ;

б) компоненты с шириной линии  $\Gamma_R \sim E \frac{\bar{u}}{c} \gg \Gamma$ , где  $\bar{u}$  - средняя квадратичная скорость теплового движения гамма-радиоактивных ядер в твердом теле; эта компонента имеет энергию, смещенную относительно значения  $E$  на величину энергии отдачи  $R = \frac{E^2}{2M_0 c^2}$ , где  $M_0$  - масса излучающего ядра (если считать его свободным и движущимся со скоростью  $u \ll c$ ).

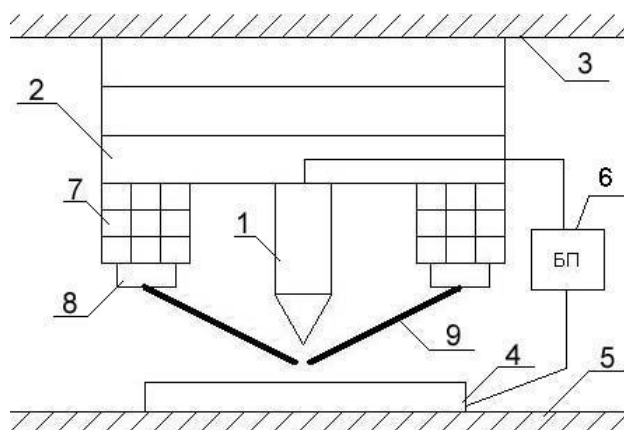
В результате линии гамма-излучения и поглощения (той же линии) сильно размываются и, кроме того, сдвинуты по энергии друг относительно друга на величину  $\sim 2R$ . Ввиду того, что для гамма-излучения  $R$  в общем не мало по сравнению с  $E$ , явление *резонансного поглощения гамма-лучей* ( $E_{\text{изл}} = E_{\text{погл}}$  или  $\nu_{\text{изл}} = \nu_{\text{погл}}$ ) обычно практически не наблюдается.

При определенных условиях удается добиться того, что излучаемый гамма-квант передает импульс не одному излучающему ядру, а всему кристаллу в целом. В результате излучаемой линии соответствует энергия отдачи  $R \approx 0$  ( $M$  – велико) и  $\Gamma_R \approx \Gamma$ , т.е. ширина линий приближается к естественной, а сдвиг по энергии практически исчезает. Одним из условий четкого проявления эффекта Мессбауэра, является условие  $R \leq 2k\Theta_D$ , где  $\Theta_D$  – дебаевская температура кристалла,  $k$  – постоянная Больцмана. При  $R \ll 2k\Theta_D$  гамма-переходы «без отдачи» можно наблюдать уже при комнатной температуре;

### 6.3. Варианты технических устройств для получения элементов систем кодирования и криптографии

#### 6.3.1. Устройство для получения нанодорожек

В основу разработки положена задача обеспечить возможность оперативного управления процессом получения различных нанодорожек.



*Рис. 6.1. Устройство для получения нанодорожек*

Устройство для получения нанодорожек (Рис. 6.1.) содержит зонд 1, жестко закреплённый на основном пьезоприводе 2, который установлен на платформе 3, подложку 4, установленную на основании 5 и блок питания 6, электрически связанный с зондом 1 и подложкой 4, два дополнительных, двухкоординатных пьезопривода 7, жёстко закреплённых на торце основного пьезопривода 2, ёмкости 8, с рабочим веществом, установленные на торцах двухкоординатных пьезоприводов 7, а так же капиллярные трубки 9, свободные торцы которых расположены вблизи острия зонда.

Устройство для получения нанодорожек работает следующим образом.

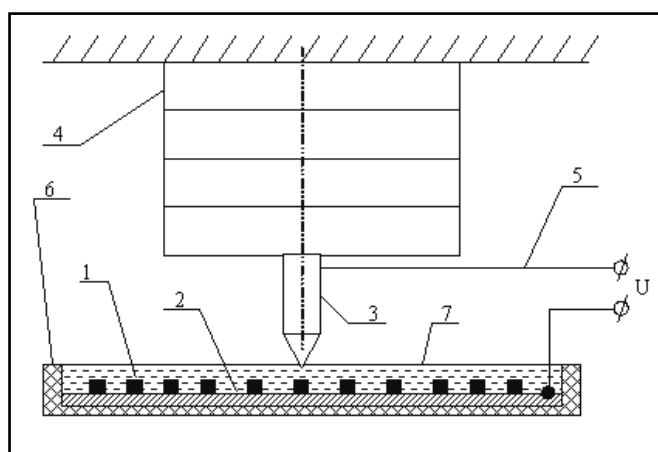
С помощью основного пьезопривода устанавливается туннельный зазор между зондом и подложкой. С помощью дополнительных,

двухкоординатных пьезоприводов устанавливаются в исходное положение сопла с рабочим веществом. В туннельный зазор подается рабочее вещество из сопла 9, количество сопел, а, следовательно, и рабочих веществ не ограничено. При подаче разности потенциалов между зондом и подложкой, атомы рабочего вещества проникают в глубь подложки. В результате чего появляется возможность наносить мозаичную структуру на поверхность подложки.

Применение устройства для получения нанодорожек позволяет наносить мозаичную структуру на поверхность подложки.

### 6.3.2. Устройство формирования изомерных квантовых точек

В основу разработки положена задача создания наногетероструктур, способных сохранять возбужденное состояние сроком до нескольких лет.



*Рис. 6.2. Устройство формирования изомерных квантовых точек*

Устройство формирования изомерных квантовых точек 1 для систем (рис. 6.2) долговременной памяти на подложке 2, содержит зонд 3, закрепленный на пьезоприводе 4, подложку, электрически связанную 5 с зондом, установленную в ванне 6 с соляным раствором вещества 7, ядра которого обладают ядерной изомерией.

Устройство формирования изомерных квантовых точек работает следующим образом.

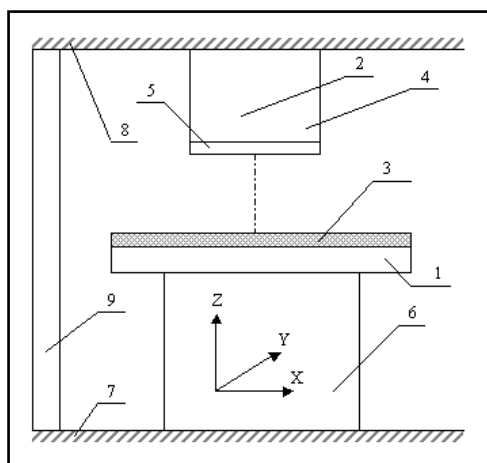
При подаче рабочего напряжения в соляном растворе начинается протекать процесс электролиза, обеспечивающего формирование квантовых точек на подложке.

Применение устройства формирования изомерных квантовых точек позволяет создать наногетероструктуры, способные сохранять возбужденные состояния в течение нескольких лет.

### 6.3.3. Устройство долговременной памяти

В основу разработки положена задача повысить плотность записи информации на записывающей матрице.

Устройство долговременной памяти (рис. 6.3) содержит записывающую матрицу 1, излучатель электромагнитных волн 2 вещество – приемник 3, расположенное на матрице 1. Вещество – приемник 3 выполнено из радиоактивного материала с изомерными ядрами. Излучатель электромагнитных волн 2 выполнен в виде источника 4 гамма–лучей – фотонов большой энергии. Причем в качестве излучателя гамма–лучей – фотонов большой энергии использован лазерный генератор 5, с возможностью перевода изомерных ядер вещества – приемника 3 в возбужденное состояние на период до нескольких лет, а записывающая матрица 1 связана с трехкоординатным приводом 6, установленным на неподвижном основании 7.



*Рис. 6.3. Устройство долговременной памяти*

Лазерный генератор 5 также закреплен на втором неподвижном основании 8, причем первое и второе основания 7,8 связаны между собой жесткой планкой 9.

Устройство долговременной памяти работает следующим образом.

Источник 4 гамма–лучей – фотонов большой энергии – лазерный генератор 5 воздействует на изомерные ядра вещества – приемника 3, переводя их в возбужденное состояние, в котором они могут находиться несколько лет. Причем одно изомерное возбужденное ядро соответствует одному биту информации. Таким образом плотность записи информации резко повышается.

Применение устройства долговременной памяти позволяет повысить плотность записи информации на три порядка, при котором размер одного бита информации составляет не более 1нм.

## Заключение

Научно-техническая революция в последнее время приняла грандиозные масштабы в области информатизации общества на базе современных средств вычислительной техники, связи, а также современных автоматизированных систем. Применение этих средств и методов приняло всеобщий характер. В современном мире информационный ресурс стал одним из наиболее мощных рычагов экономического развития. Владение информацией необходимого качества в нужное время и в нужном месте является залогом успеха в любом виде хозяйственной деятельности. Монопольное обладание определенной информацией оказывается зачастую решающим преимуществом в конкурентной борьбе и предопределяет, тем самым, высокую цену "информационного фактора". Развитие нанотехнологий привели к возможности создания защищенных информационных технологических систем элементами которых выступали бы наноструктуры.

## Литература

1. Зегжда Д.П., Ивашко А.М. Основы безопасности информационных систем. –М.: Горячая линия – Телеком, 2000.
2. Trusted Computer System Evaluation Criteria. Us Department of Defense 5200.28-STD, 1993.
3. Гостехкомиссия России. Руководящий документ. Концепция защиты средств вычислительной техники от несанкционированного доступа к информации. Москва, 1992.
4. Гостехкомиссия России. Руководящий документ. Руководящий документ. Средства вычислительной техники. Защита от несанкционированного доступа к информации. Показатели защищенности от несанкционированного доступа к информации. Москва, 1992.
5. Гостехкомиссия России. Руководящий документ. Автоматизированные системы. Защита от несанкционированного доступа к информации. Классификация автоматизированных систем и требования по защите информации. Москва, 1992.
6. Гостехкомиссия России. Руководящий документ. Временное положение по организации разработки, изготовления и эксплуатации программных и технических средств защиты информации от несанкционированного доступа в автоматизированных системах и средствах вычислительной техники. Москва, 1992.
7. Гостехкомиссия России. Руководящий документ. Защита от несанкционированного доступа к информации. Термины и определения. Москва, 1992.

8. Information Technology Security Evaluation Criteria. Harmonized Criteria of France-Germany-Netherlands-United Kingdom. – Department of Trade and Industry, London, 1991.
9. Federal Criteria for Information Technology Security. National institute of Standards and Technology & National Security Agency. Version 1.0, December 1992.
10. Canadian Trusted Computer Product Evaluation Criteria. Canadian System Security Centre Communication Security Establishment, Government of Canada. Version 3.0e. January 1993.
11. Common Criteria for Information Technology Security Evaluation. National Institute of Standards and Technology & National Security Agency (USA), Communication Security Establishment (Canada), Communication-Electronics Security Group (United Kingdom), Bundesamt fur Sicherheit in der Informationstechnik (Germany), Service Central de la Securite des Systemes d'Information (France), National Communications Security Agency (Netherlands). Version 2.1, August 1999.
12. M. Harrison, W. Ruzzo, J. Uhlman “Protection of operating systems”, Communications of the ACM, 1976.
13. Ravi S. Sandhu “The Typed Access Matrix Model” Proceedings of IEEE Symposium on Security and Privacy, Oakland, California, May 4-6, 1992, pages 122-136.
14. M Harrison, W. Ruzzo, “Monotonic protection systems”, Foundation of secure computation, 1978.
15. Leonard J. LaPadula and D. Elliot Bell “Secure Computer Systems: A Mathematical Model”, MITRE Corporation Technical Report 2547, Volume II, 31 May 1973.
16. Ciaran Bryce “Lattice-Based Enforcement of Access Control Policies”, Arbeitspapiere der GMD (Research Report), Nummar 1020, August 1996.
17. John McLean “The Specification and Modeling of Computer Security”, Computer, 23(1):9-16, January 1990.
18. John McLean “Security Models”, Encyclopedia of software engineering, 1994.
19. А.В.Панов и др. Устройство для получения нанодорожек. Патент РФ на ПМ № 42696.- Опубликована 10.12.2004 Б. №34.
20. Е.Н.Ивашов, Б.Г.Львов, С.В.Степанчиков. Способ получения нанотрубок. Патент РФ на изобретение №2225655. – Опубл. 10.03.2004. Бюл. №7.



## Глава 7. Формирование рельефа подложки в туннельно-зондовой нанотехнологии

### Введение

Многokратное увеличение быстродействия, уровня интеграции и расширение функциональных возможностей в электронике, оптике, робототехнике, материаловедении, биологии, информатике и других областях науки и техники возможно на основе нанотехнологии - науки об управлении процессами на уровне атомов и молекул (1 нанометр= $10^{-3}$  мкм). Практически не остается ни одного раздела науки и техники, которые не претерпят кардинальных изменений с развитием нанотехнологии. Поэтому развитие нанотехнологии за рубежом называется нанотехнологической революцией.

Предлагается дать следующее рабочее определение для нанотехнологии.

*Нанотехнологией* называется междисциплинарная область науки, в которой изучаются закономерности физико-химических процессов в пространственных областях нанометровых размеров с целью управления отдельными атомами, молекулами, молекулярными системами при создании новых молекул, наноструктур, наноустройств и материалов со специальными физическими, химическими и биологическими свойствами.

Задача уменьшения линейных размеров используемых элементов микросхем - одна из основных в микроэлектронике. Создание элементов схем с характерными размерами порядка единиц и десятков нанометров качественно меняет электронику и переводит ее в новую область - наноэлектронику. Изменяется и физика работы элементов - они будут работать преимущественно на принципах квантовой механики.

Создание интегральных наноэлектронных квантовых схем, по существу, конечная цель нанотехнологии. Нанотехнологию также можно определить как совокупность способов и приемов создания функциональных элементов нанометровых размеров на поверхности подложек, в том числе из отдельных молекул и атомов, с возможностью одновременной их визуализации и контроля.

### 7.1. Основные направления в нанотехнологии

Традиционный метод, включающий создание масок на поверхности полупроводниковой пластины с последующим применением микролитографий все более высокого разрешения, в том числе рентгено-, электроно- и ионной литографий, а также синхротронного излучения, возможно, приведет к созданию дорожек с нанометровыми поперечными размерами. Однако создание элементов на основе отдельных молекул и атомов традиционными путями невозможно.

Изобретение в 1982 г. Г. Биннигом и Г. Рорером, сотрудниками швейцарского филиала фирмы IBM, сканирующего туннельного микроскопа



(СТМ) открыло прежде всего новый путь для неразрушающего контроля и анализа металлических и полупроводниковых подложек с разрешением до 0,01 нм. С помощью СТМ получено изображение поверхности с атомным разрешением различных проводящих монокристаллических и поликристаллических материалов, разработаны новые методы исследования поверхностей твердых тел.

Принцип работы СТМ прост: сканирующий туннельный зонд представляет собой металлический игольчатый электрод, закрепленный в трехкоординатном пьезоприводе и располагающийся перпендикулярно исследуемой поверхности. С помощью пьезопривода зонд подводится к поверхности образца до возникновения туннельного тока, который определяется прежде всего величиной зазора между зондом и поверхностью, а также напряжением между электродами. Если туннельный ток и напряжение поддерживать постоянными, то при сканировании с помощью зонда можно получать непосредственную информацию о рельефе поверхности.

СТМ является уникальным инструментом для исследования физики поверхности на атомном уровне. Туннельная микроскопия поверхности подложек позволила исследовать различные процессы, в том числе изменение структуры поверхностей при различных процессах химического или ионного травления, а также позволила визуализировать разнообразные процессы при осаждении пленок.

Создатели туннельного микроскопа, по-видимому, первыми предложили его технологическое использование для получения проводящих дорожек с помощью испарения материала с туннельного зонда. Последующие исследования показали, что на базе СТМ возможна разработка новой технологии - зондовой нанотехнологии (существенно отличающейся от традиционных подходов), в основе которой лежит применение туннельного зонда как для визуализации объектов на подложке, так и для формирования их в нанометровых областях.

Идея применения зонда для сканирования поверхности подложки оказалась чрезвычайно плодотворной. Были созданы сканирующие атомно-силовые микроскопы (АСМ), в которых зонд крепится на весьма чувствительной балочке (балочку с зондом и её держателем называют кантилевером), позволяющие исследовать рельеф поверхности диэлектрических подложек с атомным разрешением, лазерно-силовые микроскопы и микроскопы с термическими зондами.

Туннельно-зондовая нанотехнология (ТЗН) начала развиваться по двум направлениям: сверхвысоковакуумная нанотехнология и нанотехнология в газах и жидкостях при атмосферном давлении, поскольку были созданы СТМ, работающие как в высоковакуумных, так и в атмосферных условиях.



*Рис. 7.1. Развитие нанотехнологии в электронике*

Основное преимущество высоковакуумной ТЗН - возможность иметь исходно чистые подложки в чистом объеме, что позволяет манипулировать с отдельными молекулами и атомами. Однако массоперенос между игольчатым электродом и подложкой, накопление молекул и атомов на подложке, возможность их ухода из межэлектродного зазора и поступление примесей из вакуумного объема влияют на исходные условия на поверхности и в объеме. Концепция ТЗН в газах и жидкостях исходит из того, что при наличии соответствующим образом подобранных технологических носителей ультравысокой чистоты возможно получение результатов, не уступающих по многим параметрам нанотехнологии в глубоком вакууме. Объекты воздействия в этом случае уже не отдельные атомы и молекулы, а образования с размерами до 10-30 нм, например, кластеры. По сравнению с высоковакуумной ТЗН в газах и жидкостях может, по-видимому, иметь существенное преимущество в производительности и технологичности.

Традиционные литографии высокого уровня разрешения с использованием масок и шаблонов ведут от микротехнологии к субмикротехнологии, а в ближайшем будущем, возможно, к нанотехнологии. Зондовые микроскопы позволяют создать на их базе устройства для реализации туннельно-зондовой нанотехнологии. Такой путь продвижения к нанотехнологии выгодно отличается тем, что он позволяет одновременно и контролировать и визуализиро-

вать процессы ТЗН. Вакуумная и атмосферная нанотехнологии должны привести сначала к созданию дискретных устройств наноэлектроники в виде отдельных функциональных элементов, в том числе устройств памяти со сверхплотной записью информации, а в последующем интегральных квантовых схем с нанометровыми размерами элементов (рис. 7.1). При этом в полной мере могут быть реализованы идеи молекулярной электроники, когда в качестве элементной базы предполагается использование и модификация отдельных молекул. Более простой, но более "грубой" элементной базой наноэлектроники, по-видимому, могут быть металлические и полупроводниковые кластеры, покрытые лигандами органических и неорганических соединений.

Эти образования с характерными размерами до 30 нм сами по себе обладают уникальными свойствами. Лигандное покрытие переводит их в устойчивые состояния. В свою очередь, лигандные кластеры могут образовывать, например, металлоцепи с металлической связью между остовами.

Элементная база на основе лигандных кластеров, по-видимому, сделает возможным создание интегральных схем, работающих, в частности, на основе эффекта дискретного одноэлектронного туннелирования, в том числе ОЗУ со сверхбольшой производительностью и емкостью на кристалле.

Особенность нанотехнологии заключается в том, что рассматриваемые процессы и совершаемые действия происходят в нанометровом диапазоне пространственных масштабов, а характерный размер атома составляет несколько десятых нанометра. В этом диапазоне размеров "сырьем" являются отдельные атомы, молекулы, молекулярные системы, а не привычные в традиционной технологии микронные или макроскопические объемы материала, содержащие, по крайней мере, миллиарды атомов и молекул. В отличие от традиционной технологии, для нанотехнологии характерен "индивидуальный" подход, при котором внешнее "управление" достигает отдельных атомов и молекул, что позволяет создавать из них как "бездефектные" материалы с принципиально новыми физико-химическими и биологическими свойствами, так и новые классы устройств с характерными нанометровыми размерами.

Анализ текущего состояния позволяет выделить в нанотехнологии ряд важнейших направлений.

**I. Молекулярный дизайн.** Препарирование молекул и синтез новых молекул в сильно неоднородных электромагнитных полях.

**II. Наноматериаловедение.** Создание "бездефектных" высокопрочных материалов, материалов с высокой проводимостью.

**III. Наноприборостроение.** Создание сканирующих туннельных микроскопов, атомно-силовых микроскопов, магнитных силовых микроскопов, многоострийных систем для молекулярного дизайна, миниатюрных сверхчувствительных датчиков, нанороботов.

**IV. Наноэлектроника.** Конструирование нанометровой элементной базы для ЭВМ следующего поколения, нанопроводов, полевых транзисторов, выпрямителей, дисплеев, акустических систем.

**V. Нанооптика.** Создание нанолазеров, синтез многоострийных систем с нанолазерами.

**VI. Нанокатализ.** Разработка катализаторов с наноструктурами для классов реакций селективного катализа.

**VII. Наномедицина.** Проектирование наноинструментария для уничтожения вирусов, локального "ремонта" органов, высокоточной доставки доз лекарств в определенные места живого организма.

**VIII. Нанотрибология.** Определение связи наноструктуры материалов и сил трения и использование этих знаний для изготовления перспективных пар трения.

**IX. Управляемые ядерные реакции.** Наноускорители частиц, не статистические ядерные реакции.

## **7.2. Физические основы процесса формирования нанорельефа подложки**

### **7.2.1. Основные нанотехнологические процессы**

Основой нанотехнологических процессов является проведение локальных атомно-молекулярных взаимодействий.

В настоящее время наиболее распространены групповые технологии создания объектов нанометровых размеров с помощью осаждения и литографии (рис. 7.2).

Групповые технологии осаждения характеризуются особенностями, существенно ограничивающими возможности создания структур нанометровых размеров.

Из-за одновременного осаждения на различные участки подложки возникают зерна, дислокации, поры и другие дефекты. Применение методов эпитаксии позволяет преодолеть данные недостатки, однако из-за высокой температуры эпитаксиальных процессов (необходимой для повышения поверхностной миграции) ликвидируется возможность локального осаждения. Локализация осаждаемого материала возможна в методе графоэпитаксии, однако его развитие сдерживается возможностями методов литографии.

Традиционно основным направлением развития методов литографии, обеспечивающим повышение разрешающей способности, считалось применение свободно распространяющихся в пространстве частиц с меньшей длиной волны. Поэтому проводились разработки в направлении укорочения длины волны используемого излучения - базирующиеся на применении ультрафиолетового или синхротронного излучения, а также высокоэнергетичной электронной или даже ионной литографии.

Методы оптической литографии пока ограничены техническими возможностями фокусирования света - традиционными линзовыми системами, осуществляющими передачу излучения через открытое пространство в размеры, соразмеримые с длиной волны излучения. Методы электронной и ион-

ной литографии позволяют осуществить фокусировку воздействующего электронного потока в малые размеры. Однако высокая энергия фокусируемых электронов приводит к значительному разрушению используемых материалов, что ограничивает пространственную разрешающую способность метода.

В то же время известен физический эффект, позволяющий получить пространственное ограничение потока излучения в размерах, меньших длины волны используемых частиц. Главная особенность эффекта заключается в наличии условий, запрещающих свободное распространение частиц через определенную область пространства. Этот эффект первоначально в ядерной физике получил название эффекта туннелирования. В оптике его называют эффектом нарушенного полного внутреннего отражения, а в СВЧ-технике - эффектом запертого волновода.

Применение данных эффектов на вершинах зондов специальных конструкций позволило достичь высокой пространственной разрешающей способности без применения высокоэнергетических частиц и создать новые методы техники сканирующей зондовой микроскопии на их основе. Например, эффективная ширина потока туннелирующих электронов при энергии в доли электрон-вольт не превышает 0,1-0,2 нм, а оптическое излучение металлооптическими волноводами можно локализовать в области в десятки раз меньшей длины волны используемого излучения.

На основе техники сканирующей зондовой микроскопии стали создаваться методы нанотехнологии, использующие частицы с величинами энергий, определяемыми не энергией, необходимой для их фокусировки, а оптимальной энергией стимуляции нанотехнологических процессов. Именно данное сочетание позволило достичь высокой разрешающей способности при формировании элементов нанометровых размеров. Рассмотрим основные фазы нанотехнологических процессов. При этом целью повышения эффективности разработки нанотехнологии целесообразно выделить и проанализировать две важнейшие фазы атомно-молекулярных процессов, заключающиеся в фиксации и активации атомов.

Наиболее распространенным типом фиксации является локализация нанообъектов (атомов, молекул, наночастиц) в неоднородном электрическом поле, создаваемом между вершиной зонда и поверхностью образца.

За счет дипольного момента и градиента электрического поля нанообъектам энергетически наиболее выгодно расположиться в области максимальной напряженности электрического поля, т.е. в промежутке зонд-подложка.

Для повышения эффективности удержания целесообразно увеличивать величину градиента электрического поля, что достигается применением более острых зондов и увеличением напряжения между зондом и образцом.

Для увеличения эффективности удержания целесообразно увеличивать величину градиента магнитного поля, что достигается применением более острых зондов с вершинами изготавливаемых из материалов с максимально

технически допустимой по условиям применения в эксперименте магнитной проницаемостью и увеличением напряженности магнитного поля.

При протекании электрического тока через нанообъект и одновременном действии магнитного поля (в том же направлении) возникает сила Лоренца, дополнительно локализующая движение нанообъекта. Возможно также локализовать атомы и молекулы с помощью стоящих оптических волн. Для повышения величины барьера и степени локализации целесообразно увеличивать мощность и частоту применяемого излучения. Одним из возможных вариантов использования данного эффекта является использование зонда в качестве элемента металло-оптического волновода.

Следует отметить, что при расположении нанообъектов между вершиной зонда и подложкой на них начинают влиять силы межатомного взаимодействия.

С целью увеличения величины межатомных взаимодействий, кроме подбора соответствующего материала целесообразно использовать зонды специальных конструкций, включая зонды из нанотрубок, увеличивающих размеры области взаимодействия,

Сравнение различных методов локализации показывает, что наиболее эффективным способом локализации является использование трехмерного потенциального барьера, создаваемого на конце нанотрубки. Поэтому сочетание возможности инъекции частиц по нанотрубке и их локализация на конце является одним из наиболее перспективных направлений дальнейших разработок нанотехнологии.

Основной фазой нанотехнологических процессов, позволяющей изменить структуру объектов, является фаза активации. Возможности активационных процессов зависят прежде всего от величины энергии активации. Величины энергий активационных процессов находятся в диапазоне от долей электрон-вольт до десятков электрон-вольт. Рассмотрим диапазон энергий активации, соизмеримых с величиной энергии тепловых колебаний ( $\sim 0.25$  эВ). При таких энергиях активация может привести к увеличению эффекта поверхностной миграции, что позволяет преодолеть барьеры, локализирующие атомы при физической сорбции на поверхности. Кроме того, в ряде ситуаций данная величина энергии достаточна для начала процессов диссоциации молекул. При проведении нанотехнологических процессов необходимо даже создавать дополнительные условия диссоциации молекул, размещая их в областях с повышенной плотностью электронов или повышенной напряженностью сил межатомного взаимодействия. Создание энергии активации в несколько эВ осуществляется приложением внешнего электрического поля, межатомного взаимодействия, теплового нагрева, рассеяния электронного потока, воздействия коллективных процессов. Диапазон энергий в десятые доли эВ соответствует дальнему инфракрасному диапазону.

Основными процессами в данном диапазоне являются перемещение частиц и их тепловая активация. Перемещение частиц позволяет осуществить

планарную сборку вещества. Однако создаваемые структуры вследствие невысоких энергий связи характеризуются невысокой стабильностью.

В диапазоне энергий активации от нескольких эВ до десятка эВ данный диапазон фактически является основным для большинства нанотехнологических процессов, здесь происходит разрыв и синтез органических молекул, ионизация и т.д. По величине используемых квантов данный диапазон энергий соответствует оптическому. В данном диапазоне энергий возможна трехмерная сборка ряда материалов с невысокой энергией связи. Создаваемые вещества характеризуются большей стабильностью (по сравнению с предыдущим методом).

В диапазон энергий более десятка эВ основными процессами являются ионизация практически всех материалов и испарение ионов. В данном диапазоне возможна 3-мерная сборка высокотемпературных материалов, характеризующихся наивысшей стабильностью. Важным вопросом является соотношение энергий воздействия и энергии, необходимой для проведения данного процесса. Для проведения нанотехнологического процесса, необходимо, чтобы энергия воздействия превышала пороговую энергию начала соответствующего процесса. Однако при проведении процесса с существенным превышением данных энергий разница энергий будет преобразовываться в движения, мешающие стабилизации результатов процесса. К ним относятся - разлет частиц, стимуляция многокаскадных процессов, включая коллективные, ухудшающие локальность воздействия.

Процессы, происходящие в результате данных методов активации, можно классифицировать по следующим основным типам.

### **Полевые процессы**

- Поляризация атомов и молекул, приводящая к направленной ориентации, повышающая структурную упорядоченность вещества и тем самым повышающая вероятность соответствующих атомно-молекулярных процессов. В результате поляризации может образовываться состояние технологической среды, похожее на жидкокристаллическое.
- Стимулированная миграция атомов, молекул и наноструктур под действием градиента электрического поля. В результате стимулированной миграции возможно перемещение адсорбированных частиц на поверхности с целью формирования заданных структур.
- Диссоциация молекул и образование новых связей в сильных электрических полях и в областях с повышенной концентрацией электронов. В сильных электрических полях и в присутствии электродов, повышающих пространственную плотность электронов, возможна стимуляция процессов диссоциации молекул и образования новых связей. Данный метод позволяет образовывать химические связи из адсорбированного вещества с подложкой, отличающиеся большей устойчивостью создаваемых наноструктур по сравнению со структурами, удерживаемыми за счет физической сорбции.

- Десорбция молекул и атомов с вершины зонда или соответствующего места подложки. Метод позволяет не только производить предварительную очистку вершины зонда и соответствующего места подложки, но и осаждать материалы с вершины зонда на подложку и наоборот.
- Ионизация атомов и молекул. Процесс значительно повышает скорость прохождения реакций.
- Полевая деформация поверхности зонда и подложки. Процесс является следствием ряда процессов – создания кулоновского взаимодействия, рассеяния электронов и повышения скорости миграции.
- Полевая эмиссия электронов и ионов. В сильных электрических полях на первом этапе возникает полевая эмиссия электронов, а при дальнейшем увеличении напряженности поля до нескольких вольт на ангстрем начинается полевое испарение материала.

### **Кинетические процессы**

- Рассеяние электронов. Процесс порождает возбуждение различных колебаний в молекуле, приводит к локальному разогреву, симулирует поверхностную миграцию. Не всегда эффективность процесса достаточна для стимуляции нанотехнологических процессов, потому что масса электрона в тысячи раз меньше массы молекулы, на которой он рассеивается, и поэтому процесс передачи энергии происходит недостаточно эффективно.
- Рассеяние нейтральных частиц. Рассеяние нейтральных частиц в большинстве процессов, стимулируемых вышеописанными методами, не используется или является вторичным эффектом. Однако в будущем, при специальном конструировании инжекторов нейтральных частиц, процесс мог бы быть более эффективным, чем процесс рассеяния электронов.
- Рассеяние заряженных частиц. Рассеяние заряженных частиц более эффективно по сравнению с рассеянием электронов, однако реализации данного режима противодействует обычно присутствующий поток электронов и достаточно малое время перезарядки.
- Тепловое воздействие. Является наиболее эффективным способом стимулировать физико-химические процессы. Однако несмотря на высокую скорость диссипации энергии, для обеспечения нанометровой пространственной разрешающей способности требуется осуществлять воздействия с пикосекундными длительностями.

### **Электромагнитные процессы**

- Нерезонансные. При вводе излучения с частотой, не соответствующей пику поглощения, возможно использование процессов локализации, нагрева или охлаждения объектов, располагаемых в зазоре зонд-поверхность образца.
- Резонансные в одном спектральном диапазоне. Применение излучения, соответствующего пику поглощения, позволяет повысить эффективность поглощения и соответствующего разделения молекул, а также за-



претить образование химических связей с энергией, равной энергии кванта вводимого излучения.

- Резонансные в нескольких спектральных диапазонах. Введение излучения на нескольких частотах позволяет повысить эффективность управления технологическим процессом, сужая диапазон возможных разрешенных результатов путем увеличения числа запрещенных.
- Спектрально синхронизованные. Прохождение во времени нанотехнологических процессов сопровождается сложным изменением во времени энергетических связей. При синхронизированной во времени накачке, амплитудной, частотной и фазовой модуляции возможно более эффективное управление ходом атомно-молекулярных процессов.
- Пространственно-когерентные. Прохождение нанотехнологического процесса сопровождается изменением пространственного положения атомов и молекул. Введение излучения не только по времени, но и в пространстве когерентного с пространственным преобразованием положения атомов и молекул позволит обеспечить самое оптимальное управление ходом нанотехнологических процессов.
- Межатомные взаимодействия. Потенциально данный вид взаимодействий является наиболее эффективным по величине энергий, по соотношению масс и т.д. Однако использование его требует создания специальных конструкций, изготовление которых в свою очередь также требует использования нанотехнологий. В дальнейшем возможно создание специализированных конструкций типа нанотрубок - транспортеров атомов, в которых будут происходить несколько процессов транспортировка, активация и т. д.

Методы активации нанотехнологических процессов определяются в основном техническими возможностями нанотехнологических установок. К основным из них относятся создание между вершиной зонда и образцом:

- разности электрических потенциалов;
- механического воздействия;
- теплового воздействия (создание градиента температуры);
- оптического воздействия;
- ультразвукового воздействия;
- электромагнитного воздействия.

В зависимости от технологического режима возможна также комбинация воздействий (в различных амплитудных и временных пропорциях).

При проведении конкретных процессов не всегда удается использовать только один из физических эффектов. Реально используется их совокупность.

Перечень классифицированных физических эффектов, используемых в нанотехнологических процессах, приведен на рис. 7.2.

ПОЛЕВЫЕ	КИНЕТИЧЕСКИЕ	ЭЛЕКТРОМАГ- НИТНЫЕ
Поляризация	Рассеяние электронов	Нерезонансные
Стимуляция миграции	Рассеяние нейтральных частиц	Резонансные в од- ном спектральном диа- пазоне
Диссоциация	Рассеяние заряженных частиц	Резонансные в не- скольких спектральных диа- пазонах
Десорбция	Тепловое воздействие	Спектрально- синхронизирован- ные
Полевая деформация		Пространственно- когерентные
Полевая эмиссия		Межатомные взаимодействия

*Рис. 7.2. Нанотехнологические процессы*

В настоящее время нанотехнология находится в начале развития и поэтому еще рано утверждать, что найдены самые оптимальные сочетания физических эффектов, используемые для проведения нанотехнологических процессов.

Однако некоторые сочетания уже сейчас можно анализировать и использовать.

### **7.2.2. Типы нанотехнологических воздействий**

В зависимости от уровня сложности и соответствующих технических возможностей нанотехнологических установок в настоящее время применяют, а также могут быть использованы следующие нанотехнологические воздействия.

## **1. Полевое воздействие**

Полевое воздействие является самым простым и поэтому самым распространенным. В результате полевого воздействия получены экспериментальные результаты по перемещению атомов и молекул, локальному электрохимическому травлению и осаждению из газовой и жидкой фаз. Основным недостатком метода являются технические проблемы активации нанотехнологических процессов при использовании изолирующих подложек.

Следует отметить, что для создания наноэлектронных элементов можно воспользоваться тем обстоятельством, что все элементы имеют электрическое подключение в конечном счете к внешним электродам. Поэтому формируя проводящие элементы от данных электродов, располагаемых на изолирующей подложке, возможно создание наноэлектронных функциональных элементов. Другой способ преодоления существующего недостатка заключается в использовании нескольких зондов. При этом возможно пространственное разделение областей активации и области осаждения продуктов реакции. Данный способ проведения нанотехнологических процессов открывает возможность осуществлять их на подложках с любым типом проводимости.

Дальнейшее развитие многозондового метода проведения нанотехнологических процессов заключается в реализации мажоритарного способа активации, заключающегося в инъекции носителей зарядов из нескольких зондов в один. При этом у данного зонда локально возрастает температура и повышается тем самым вероятность проведения локальной физико-химической реакции. Направленность процесса обеспечивается приложением напряжения между зондами соответствующей полярности.

## **2. Термополевое воздействие**

Термополевое воздействие является сочетанием создания разницы температур между вершиной зонда и подложкой. Модификацией данного метода является приложение дополнительного напряжения между вершиной зонда и подложкой (если подложка проводящая). Достоинством метода термополевой десорбции является возможность осаждения с атомарной разрешающей способностью материалов любых типов. Отличие метода заключается в необходимости достижения достаточно высоких температур и создания соответствующих градиентов.

## **3. Оптическое воздействие**

Достоинствами метода являются:

- возможность резонансного поглощения излучения, приводящего к диссоциации молекулы;
- высокая степень локализации атомов и молекул в узле напряженности электромагнитного поля;
- возможность работы с материалами подложки любых типов, включая диэлектрические;

- высокая скорость начала и окончания процесса, определяемая скоростью модуляции оптического излучения, достигающая пикосекундного диапазона и соизмеримая со скоростью прохождения атомно-молекулярных процессов.

В настоящее время данное направление является наиболее перспективным. В дальнейшем, при появлении возможности создания эффективных систем, генерирующих управляемое по частоте лазерное излучение, будет возможно генерировать излучение с перестраиваемой частотой во времени наиболее оптимальной для проведения соответствующего технологического процесса.

### 7.2.3. Физические эффекты в туннельно-зондовой нанотехнологии

В растровом туннельном микроскопе при напряжении между игольчатым электродом и подложкой 5 В и зазоре между ними 0.5 нм возникают электрические поля приблизительно  $10^8$  В/см, сравнимые с внутриатомными. Преимуществами таких полей являются их локальность, по крайней мере в областях на поверхности подложек с поперечником до 20 нм, и низкие приложенные напряжения, которые не могут вызывать ионизацию молекул и атомов в межэлектродном зазоре. При таких полях возможны плотности токов электронной эмиссии до  $10^8$  А/см<sup>2</sup>, которые можно оценить по формулам Фаулера - Нордгейма:

$$j = 1.55 \cdot 10^{-6} (E^2 / \Phi) \exp(-6.836 \cdot 10^7 \Phi^{3/2} \Theta / E),$$

$$\Theta = 0.935 - 10.79 \cdot 10^{-8} E / \Phi^2,$$

где  $E$  - напряженность электрического поля, В/см,  $\Phi$  - работа выхода электрона из электрода, к которому приложено электрическое поле, В. Сверхплотный пучок энергетических электронов из игольчатого электрода может вызвать на подложке локальный разогрев. Локальное повышение температуры по радиусу вдоль поверхности от оси пучка для изотропных подложек можно оценить по формуле:

$$\delta T(r) = (UI / 4\pi k l) \cdot (2l/r(1 - e^{-r/l}) - e^{-r/l}),$$

где  $U$  - напряжение, приложенное к электродам;  $I$  - туннельный ток;  $k$  - коэффициент термической проводимости подложки;  $l$  - длина неупругого рассеяния электронов в подложке.

Электростатическое поле порождает также нормальное к поверхности электродов механическое напряжение:

$$\sigma = \epsilon \epsilon_0 E^2 / 2,$$

где  $\epsilon$  - диэлектрическая проницаемость среды между электродами,  $\epsilon_0$  - диэлектрическая проницаемость вакуума. Это поле может быть достаточным для локальных упругих и пластических деформаций поверхности металличе-

ских электродов. Например, на воздухе при полях  $E > E_0$  ( $E_0$  - электростатическое поле порога пластической деформации), где

$$E_0 = 2.1 \cdot 10^3 \tau^{1/2}, \text{ В/см,}$$

( $\tau$  - механическое напряжение, при котором уже наблюдается пластическая деформация, Па), возможна локальная пластическая деформация металлической подложки в виде бугорков. При этом более жесткий игольчатый электрод может оставаться неизменным.

Для полупроводниковых подложек, когда внешнее электрическое поле проникает в объем полупроводника, ситуация усложняется, при этом возможна, в частности, локальная глубинная деформация подложек.

С помощью сильного электрического поля в межэлектродном зазоре возможна заметная поляризация молекул среды и их перестройка (например, геометрическая изомерия), а за счет диполь-дипольного взаимодействия молекул и допирующих примесей возможно образование проводящих молекулярных мостиков из адсорбата электродов, либо из жидкой диэлектрической фазы, находящейся в межэлектродном зазоре. При этом существует критическое электрическое поле  $E_m$  для образования молекулярных мостиков:

$$E_m = \frac{1}{\alpha} \left[ (\mu^2 + 2\alpha kT)^{1/2} - \mu \right],$$

где  $\mu$  - постоянный дипольный момент молекулы,  $\alpha$  - ее поляризуемость,  $T$  - абсолютная температура,  $k$  - постоянная Больцмана.

При  $E > E_m$  поляризованные молекулы будут связаны диполь-дипольным взаимодействием и ориентированы в направлении внешнего поля. При  $E < E_m$  тепловое движение молекул должно разрушать мостики. Наконец, при полях  $E > E_p$  возможно полевое испарение отдельных атомов и их комплексов с игольчатого электрода или подложки. Значение электрического поля  $E_f$  для массопереноса в виде положительных ионов можно вычислить по формуле:

$$E_p^+ = \min \left\{ (\beta e)^{-3} \left( \lambda + \sum_{k=1}^{\beta} I_k - \beta \phi \right)^2 \right\},$$

где  $e$  - заряд электрона,  $\beta$  - кратность ионизации испаряемого атома;  $I_k$  - потенциал  $k$ -кратной ионизации атома;  $\lambda$  - энергия испарения атома,  $\phi$  - работа выхода электрона. Выражение для плотности тока испаряемых ионов при комнатных температурах и ниже можно приближенно записать в виде:

$$j = j_0 \chi(E - E_p),$$

где  $j_0$  - плотность тока насыщения ионов, определяющаяся условиями на электроде, с которого происходит испарение:

$$\chi(E - E_p) = \begin{cases} 1, & E \geq E_p \\ 0, & E < E_p \end{cases}.$$

Для массопереноса в виде отрицательных ионов имеем:

$$E_p^- = \min \left\{ (\beta e)^{-3} \left( \lambda - \sum_{k=1}^{\beta} A_k + \beta \phi \right)^2 \right\},$$

здесь  $A_k$  - сродство к электрону в  $k$ -кратном зарядовом состоянии;  $\beta$  - кратность зарядового состояния атома,  $\phi$  - работа выхода электрона.

Для стационарного полевого испарения атомов, очевидно, должно выполняться условие:

$$E_0 > E_f.$$

В противном случае, вместо испарения полем будут возникать металлические перемычки между игольчатым электродом и подложкой. Поскольку поля  $E_f$  велики ( $\sim 10^8$  В/см), актуальной является проблема понижения величины этих полей.

При лазерном излучении в межэлектродном зазоре для образования на поверхности электродов возбужденных атомов вместо величины  $I_k$  в формуле (1) для  $E_f$  необходимо использовать  $I_k = I_k - h\nu$ , где  $\nu$  - частота резонансного возбуждения атомов,  $h$  - постоянная Планка.

Концентрация возбужденных атомов определяется соотношением:

$$N_1 = N_0 \rho(\nu) B_{10} / (\rho(\nu) B_{10} + A_{01}),$$

где  $N_0$  - концентрация атомов в основном состоянии;  $\rho(\nu)$  - плотность лазерного излучения на частоте  $\nu$ ,  $B_{10}$  - коэффициент вероятности перехода из основного состояния в возбужденное;  $A_{01}$  - вероятность спонтанного перехода из возбужденного состояния в основное. Используя метастабильные уровни, для которых  $A_{01}$  мала, можно существенно снизить плотность необходимого электромагнитного излучения. При более детальном рассмотрении необходимо еще учитывать вероятность "тушения" возбужденных атомов на поверхности электродов. Заметим, что с помощью перестраиваемого лазерного излучения можно управлять избирательностью процесса полевого массопереноса.

Электронные токи с плотностями до  $10^9$  А/см<sup>2</sup>, которые могут протекать в СТМ, вызывают не только локальный разогрев подложки. Может оказаться существенным влияние ponderomotorных объемных сил, распределенных с плотностью:

$$F=(1/c)[\vec{j}, \vec{B}],$$

где  $c$  - скорость света,  $B$  - индукция магнитного поля, которое создается втекающим в подложку током. Как показывает опыт, через туннельные зонды из вольфрама в защитной среде можно пропускать кратковременные токи до 100 мА. Такие токи создают небольшие магнитные поля, однако за счет сверхбольших плотностей токов значение электродинамической силы может быть существенным.

При удельных энергиях электронного пучка более  $10^6$  Вт/см<sup>2</sup> возрастают механические силы, действующие на облучаемый участок и обусловленные давлением на приповерхностный слой самого электронного пучка. Давление электронного пучка можно оценить по формуле:

$$F_e=jmv/e=1,06 \cdot 10^{-3}(P_e/\sqrt{W}),$$

где  $e$ ,  $m$ ,  $v$  - заряд, масса и скорость электронов соответственно;  $P_e$  - удельная мощность, Вт/см<sup>2</sup>;  $W$  - энергия электронов в ускоряющем поле, кэВ;  $F_e$  - давление, Па.

Таким образом, основными факторами, определяющими процессы нанотехнологии, являются локальные электрические поля, сравнимые с внутримолекулярными и атомными; сверхбольшие плотности токов до  $10^9$  А/см<sup>2</sup> и их электродинамическое воздействие; сверхплотные локальные потоки тепла, вызванные протекающими токами; возможны и внешние иницирующие воздействия.

При взаимодействии ускоренных электронов с твердым телом может происходить множество взаимосвязанных процессов, приводящих к необратимым изменениям в локальной приповерхностной области образца, которые не описываются в общем случае простыми соотношениями и должны рассматриваться в каждом конкретном случае.

### 7.3. Нанотехнологические установки и устройства

#### 7.3.1. Нанотехнологическая установка "Луч-2"

Основным инструментом для проведения нанотехнологических процессов и исследования физико-химических характеристик создаваемых нанобъектов являются нанотехнологические установки.

В процессе локального воздействия в зазоре зонд-подложка на поверхности направленным образом преобразуются межатомные и межмолекулярные связи вещества подложки и технологической среды. В результате обра-

зуются новые структурные композиции вещества, кластеры, локальные образования нового вещества, композиции органических веществ с неорганическими и формируются нужные объекты с нанометровыми размерами.

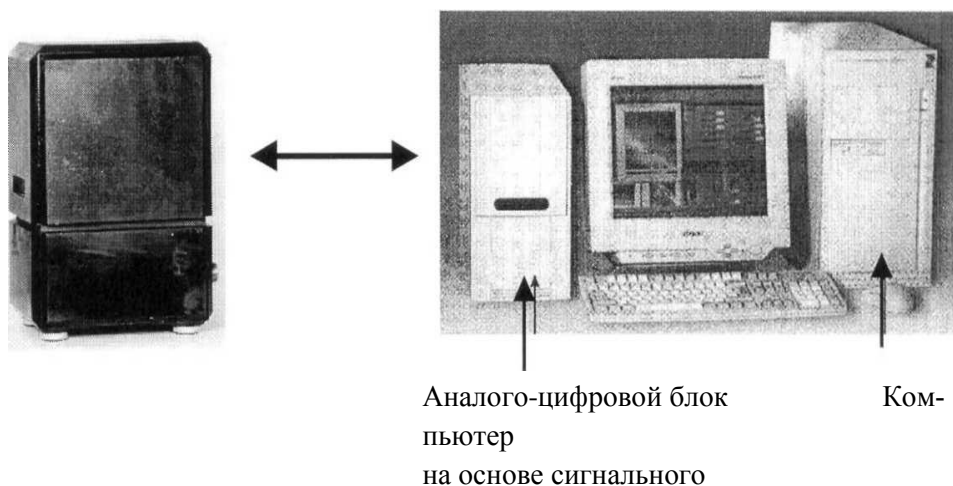
Проведение нанотехнологических процессов осуществляется на установках в зазоре зонд-подложка, при инъекции в данный объем технологических сред и комплексном воздействии различными методами на происходящие атомно-молекулярные процессы.

Исследование характеристик создаваемых нанообъектов и поверхности осуществляется методами сканирующей зондовой микроскопии и туннельной спектроскопии.

Нанотехнологические установки, разработанные Институтом Нанотехнологий Международного фонда конверсии (ИНАТ МФК) предназначены для проведения нанотехнологических процессов в газовых и жидких средах. Важнейшим отличием созданных установок является использование специального технологического модуля и соответствующей системы позиционирования, обеспечивающих возможность работы даже с агрессивными средами и стабилизацию параметров технологической среды.

Нанотехнологическая установка «Луч-2» - предназначена для использования в химическом лабораторном шкафу (с вытяжкой) рис.7.4. Используется внешняя система подготовки технологических сред.

Технологический модуль (устойчивый к агрессивным воздействиям)  
Электронная система управления



*Рис.7.3. Нанотехнологическая установка «Луч-2»*

Нанотехнологические установки  
- Обеспечивают проведение нанотехнологических реакций в самых разнообразных газовых и жидких рабочих средах, включая кислоты, щелочи, хлорорганику, металлоорганику, т.е. допускают применение не только демонстрационных но и более необходимых материалов.



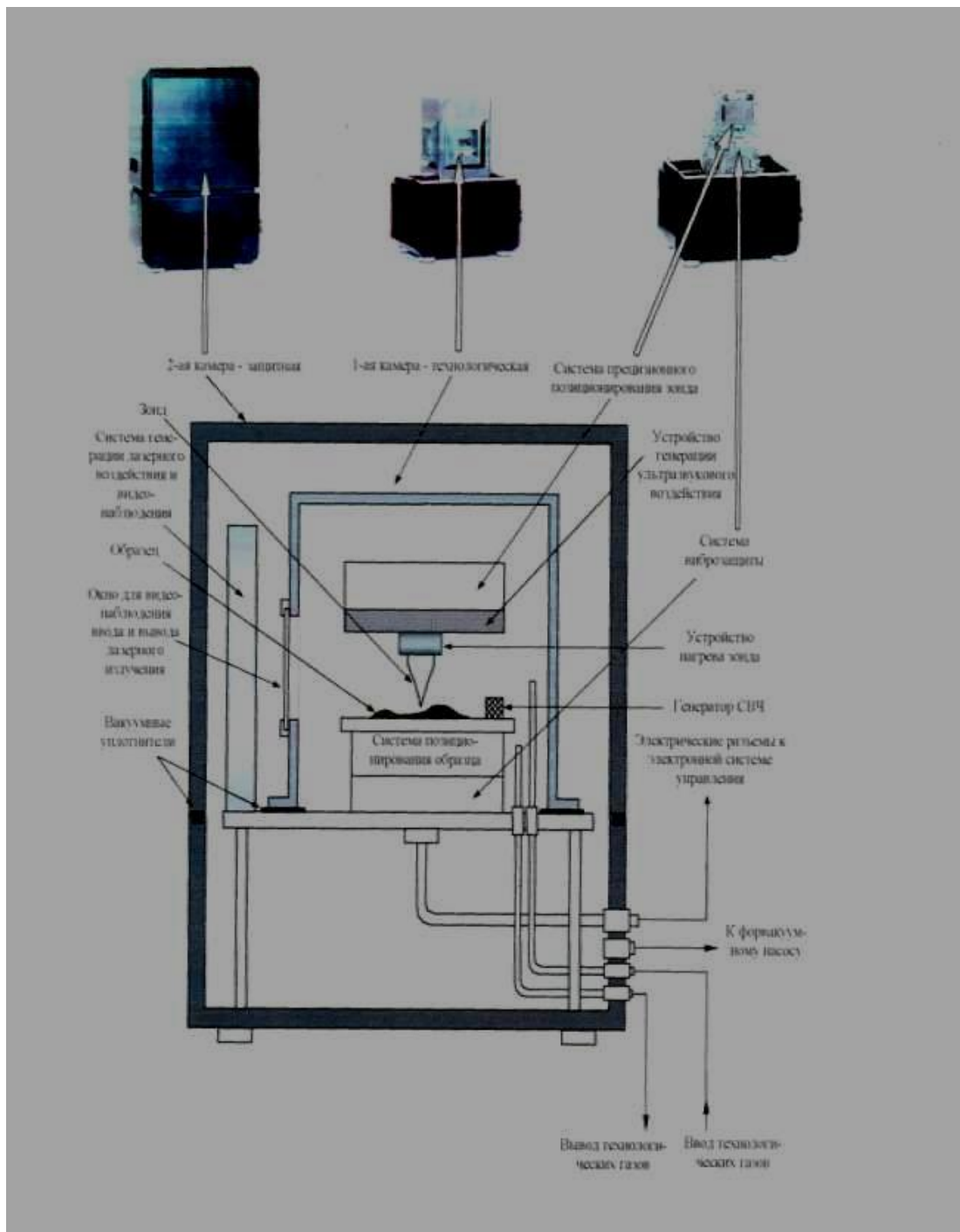
- Реализуют в зазоре зонд-образец различные по природе полевые, электромагнитные, кинетические воздействия на технологический процесс.
- Характеризуются высокой разрешающей способностью, вплоть до атомарной.
- Обладают повышенным быстродействием, благодаря специальной конструкции манипуляторов, электронной системы управления, программного обеспечения, что позволяет осуществить большой объем экспериментов.

Нанотехнологическая установка «Луч-2» (рис.7.3) состоит из:

- технологического модуля с системой прецизионного позиционирования, технологической и защитной камерами, устройствами гашения сейсмических и акустических колебаний, системой инъекции и вывода технологических сред, исполнительными элементами воздействий на процесс в зоне зонд-образец;
- электронной аналого-цифровой системы с высокочувствительными и высоковольтными электронными схемами, управляемой сигнальным процессором;
- математического обеспечения для управления нанотехнологическими процессами, проведения измерений нанообъектов и поверхности, тестирования узлов и систем установки.

В дополнение к базовой комплектации установки «Луч-2» выпускаются: Система видеонаблюдения - визуализирует на экране компьютера место подвода зонда к исследуемому образцу. Ускоряет вывод зонда на требуемую позицию. Использование афокальной оптики с зеркальными оптическими поверхностями, в которых отсутствует хроматическая аберрация, предоставляет возможность измерения объектов, расположенных под зондом, и введения излучения с длинами волн, отличными от длин волн, на которых работает система видеонаблюдения. Система характеризуется сочетанием большой глубины резкости (до нескольких мм) и высоким пространственным разрешением (до 1-2 мкм). Кроме того, применяемая оптическая система позволяет вводить лазерное излучение в зазор зонд образец и юстировать его с помощью управляемых от компьютера пьезодефлекторов.

Система расположена вне технологического модуля, и для компенсации смещения изображения применяется специальная программа стабилизации изображения на экране компьютера и оптимального наведения излучения в область проведения нанотехнологических процессов.



*Рис. 7.5. Блок-схема технологического модуля*

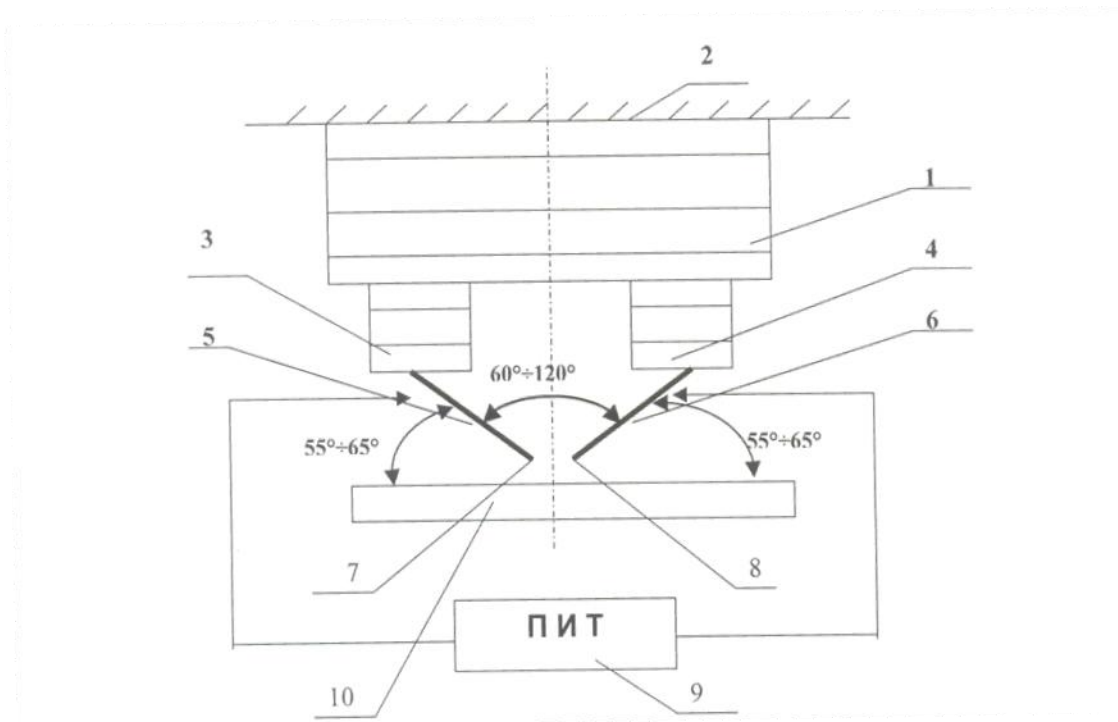
### 7.3.2. Устройство определения нанорельефа подложки

В основу разработки положена задача, обеспечить определения нанорельефа подложки, выполненного из диэлектрического или полупроводникового материала.

Устройство определения нанорельефа подложки (рис. 7.6) содержит основной пьезопривод 1, жестко связанный с основанием 2. На основном пьезоприводе 1 установлены дополнительные 3 и 4, к которым крепятся основной 5 и дополнительный 6 зонды с острями 7 и 8 соответственно.

В зондовой цепи 5-6 установлен прецизионный измеритель тока 9, изолированный от подложки 10. Зонды 5,6 установлены под углом  $55^{\circ}$ - $65^{\circ}$  к подложке и  $60^{\circ}$ - $120^{\circ}$  между собой.

Способ определения нанорельефа подложки реализуется следующим образом.



*Рис. 7.5. Устройство определения нанорельефа подложки*

Регулируют расстояние между острями 7,8 основного и дополнительного зондов 5,6 до получения туннельного зазора с соответствующим туннельным током, величину которого измеряют посредством прецизионного измерителя тока 9. Затем посредством основного пьезопривода 1 осуществляют сканирование над подложкой 10 системы из двух зондов 5,6, и измеря-

ют значение тока в зондовой цепи 5-6 посредством прецизионного измерителя тока 9. Величина тока в зондовой цепи 5-6 зависит не только от нанорельефа поверхности подложки 10, но и её материала, т.к. у различных материалов различные величины "туннельного облака" электронов вблизи поверхности.

Применение предложенного устройства позволяет определить нанорельеф подложки из любого материала, а также его стехиометрические характеристики.

### 7.3.3. Устройство определения отклонения острия зонда

В основу разработки положена задача, повысить чувствительность устройства регистрации отклонения острия зонда.

Устройство для регистрации отклонения острия зонда (рис. 7.7), содержит прецизионный измеритель перемещений 1, связанный с зондом 2 и острием 3. Прецизионный измеритель перемещений 1 выполнен в виде кварцевого резонатора 4, жестко связанного с зондом 2 и установленный соосно с ним. Устройство снабжено датчиком разности частот 5, электрически связанным с кварцевым резонатором 4. Зонд 2 с острием 3 установлен с возможностью бесконтактного взаимодействия с подложкой 6, закрепленной на основании 7.

Устройство регистрации отклонения острия зонда работает следующим образом.

Под действием сил электростатического и межатомного взаимодействия острие 3 зонда 2 перемещается в вертикальном направлении и воздействует на кварцевый резонатор 4. В результате этого взаимодействия изменяется частота колебаний кварцевого резонатора 4. Датчик разности частот 5 регистрирует разность частот колебаний кварцевого резонатора 4 до и после возникновения электростатических и межатомных взаимодействий между острием 3 зонда 2 и исследуемой поверхностью 6.

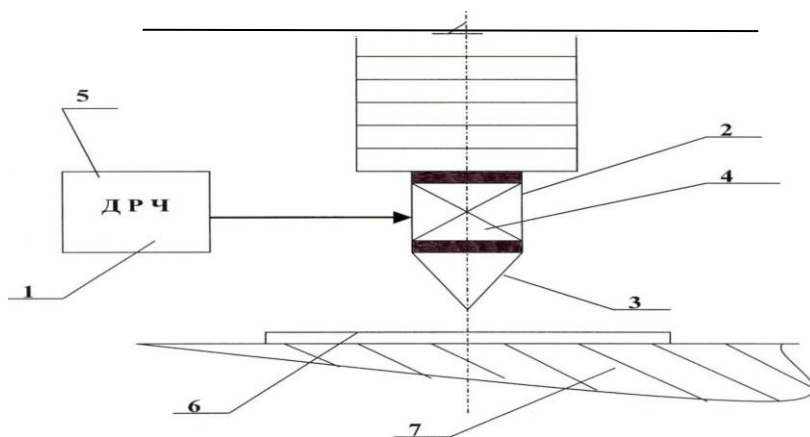
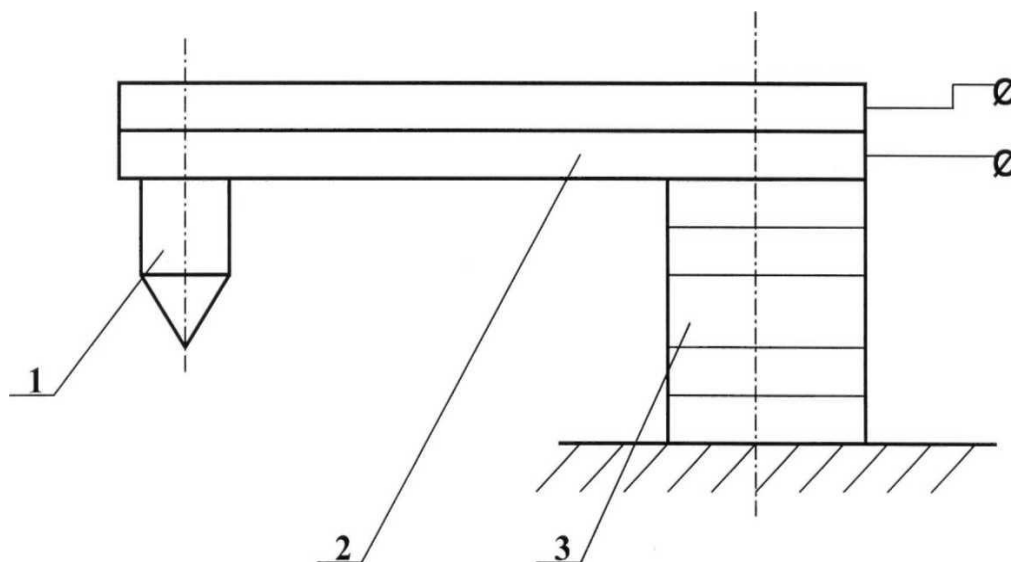


Рис. 7.6. Устройство для регистрации отклонения острия зонда

Применение предложенного устройства позволяет повысить чувствительность при отклонении острия зонда на  $1 \text{ пм} (10^{-12} \text{ м})$ .

#### 7.3.4. Устройство перемещения зонда для нанотехнологии

В основу разработки положена задача компенсировать тепловые деформации в пьезоприводе.



*Рис. 7.7. Устройство перемещения зонда для нанотехнологии*

Устройство перемещения зонда для нанотехнологии (рис. 7.8) содержит механический зонд 1, закрепленный посредством кронштейна 2 к пьезоприводе 3. Кронштейн 2 выполнен в виде двух полупроводниковых пластин, находящихся в жестком контакте, с возможностью подачи на эти пластины электрического потенциала разного знака, а зонд 1 и кронштейн 2 электрически изолированы между собой.

Устройство перемещения зонда для нанотехнологии работает следующим образом.

При подаче напряжения на пьезопривод 3 происходит перемещение кронштейна 2, а для компенсации тепловых деформаций в пьезоприводе на кронштейн 2, состоящих из двух полупроводниковых пластин, находящихся в жестком контакте, подают электрические потенциалы разного знака. При прохождении тока через полупроводниковые пластины, в соответствии с эффектом Пельтье одна пластина нагревается, а вторая охлаждается, в результате чего две

пластины начинают изгибаться, что приводит к изменению положения механического зонда 1 в пространстве и компенсации тепловых деформаций в пьезоприводе, зависящих от природы контактирующих полупроводниковых пластин и температур.

Таким образом, за счет подачи разного знака на кронштейн 2, состоящий из двух полупроводниковых пластин, находящихся в жестком контакте, происходит его изгибание и компенсация тепловых деформаций в пьезоприводе.

Применение предлагаемого технического решения устройства перемещения зонда для нанотехнологии позволяет компенсировать тепловые деформации в пьезоприводе.

### 7.3.5. Устройство кантилевера для нанотехнологии

В основу разработки положена задача, снизить рабочие электрические напряжения в промежутке зонд-подложка.

Устройство кантилевера (рис. 7.9) содержит механический зонд 1, направленный на исследуемую поверхность 2, закрепленный посредством кронштейна 3 к пьезоприводе 4. К кронштейну 3 жестко прикреплен лазер 5 и два зеркала 6,7 таким образом, что лазер 5, первое зеркало 6, исследуемая поверхность 2 и второе зеркало 7 связаны между собой посредством опорного луча, причем угол падения опорного луча на второе зеркало 7 равен нулю.

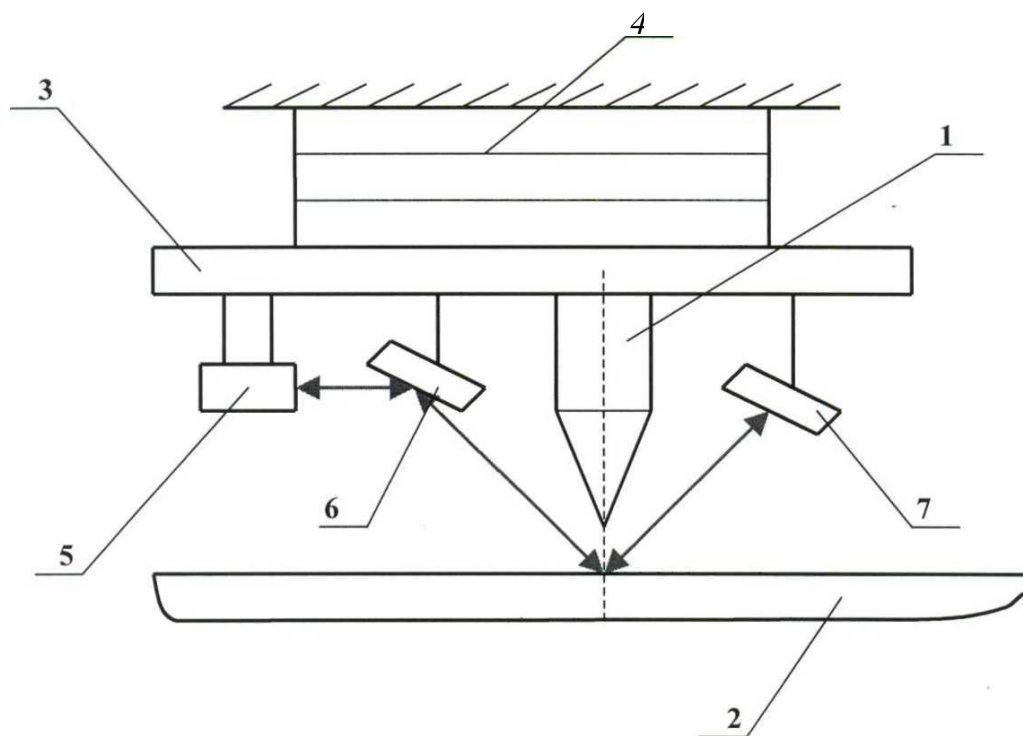


Рис. 7.8. Устройство кантилевера для нанотехнологии

Устройство кантилевера для нанотехнологии работает следующим образом.

При подачи напряжения на пьезопривод 4 происходит перемещение кронштейна 3 до возникновения туннельного тока. Для понижения рабочих электрических напряжений в промежутке зонд-подложка лазер 5 жестко прикрепленный к кронштейну 3 испускает опорный луч. При испускании лазером 5 луча на исследуемой поверхности 2 происходит переход части электронов в возбужденное состояние, что приводит к снижению работы выхода и увеличению туннельного тока. Увеличение туннельного тока позволяет снизить рабочие электрические напряжения и вероятность электрического пробоя в промежутке зонд-подложка.

Применение предлагаемого технического решения устройства кантилевера для нанотехнологии позволяет снизить рабочие электрические напряжения в промежутке зонд-подложка.

### **7.3.6. Основные параметры зондов**

Зонды предназначены для исследования и локальной модификации поверхностей в вакууме, газовых средах и непроводящих жидкостях в нанотехнологических установках. Зонды изготавливаются из вольфрамовой или платино-иридиевой проволоки с острием, сформированным методом электрохимического травления

Изготавливаются несколько типов зондов для различных целей. Базовый тип зондов - это зонд без покрытия. На его основе изготавливаются большинство других типов зондов. Базовый тип зондов предназначен для исследования и локальной модификации поверхностей в вакууме и газовой среде. В случае необходимости, для проведения некоторых нанотехнологических процессов локального осаждения, полимеризации и т.д. на вершину зонда, изготовленного из вольфрамовой проволоки, наносят покрытие в виде пленки золота толщиной 0.1 мкм. Базовый тип зондов может быть использованным в большинстве видов исследований.

Основные характеристики зондов:

- Материал зондов - вольфрам или платино-иридиевый сплав.
- Способ формирования острия - электрохимическое травление.
- Зонды изготавливаются по форме, близкой к оптимальной, характеризующейся резонансной частотой конструкции более 10 кГц.
- Радиус острия составляет не более 200 А.
- Диаметр основания зонда может составлять 0.2, 0.3 и 0.5 мм, а длина от 6 до 10 мм.
- Зонды могут иметь золотое покрытие толщиной 0.1 мкм.

Используемые в нанотехнологии зонды:

- из мономатериалов, например, вольфрама, с оптимальной формой для высокоскоростного сканирования и радиусом вершины менее 100 Å;
- с гидрофобным покрытием для проведения исследований на влажной поверхности биообъектов;
- с изолирующим на боковой поверхности покрытием для проведения электрохимических экспериментов;
- с термической активацией вершины для повышения эффективности локальных воздействий;
- с плазменной активацией вершины для стимулирования локальных химических реакций.

Наиболее часто используемыми зондами для проведения нанотехнологических процессов являются вольфрамовые зонды с радиусом вершины менее 10 нм.

После механической обработки или травления на острие появляются миниострия, через одно из которых и протекает туннельный ток, но длинные и тонкие острия или нитевидные кристаллы (вискеров) чувствительны к вибрациям и изменениям температуры, поэтому в случае наличия на острие вискеров туннельный промежуток становится неустойчивым.

В процессе туннельного зондирования, на зонд действуют межатомные силы, которые затупляют его острие. Затупление или увеличение радиуса вершины зонда снижает величину градиента электрического поля и приводит к понижению эффективности удерживания промежутка зонд-подложка. В результате чего качество изображения, получаемого в результате зондирования, существенно снижается. Затупление острия зонда является одной из проблем нанотехнологии.

В таблице приведены используемые типы зондов и их параметры.

Тип зонда	Материал	Покрытие	Диаметр, мм	Длина, мм
W-5.10	W	-	0.5	10
W-3.7	W	-	0.3	7
Pt-3.10	Pt-Ir	-	0.3	10
W/A-5.10	W	Au	0.5	10

### **Зонды с гидрофобным покрытием для биологических исследований**

Зонды с гидрофобным покрытием предназначены для проведения исследований поверхности биологических объектов, покрытых тонким слоем проводящей жидкости (например, физиологического раствора). При изготовлении зондов данного вида в качестве основы используется "базовый" тип зондов. Острие зондов формируется методом электрохимического травления. Покрытие наносят на вершину зонда и прилегающее основание. Гидрофобное покрытие изготавливается на основе углеродных материалов. Материал зондов - вольфрам, платиново-иридиевый сплав.



### **Зонды с изолирующим покрытием для электрохимии**

Зонды с изолирующим покрытием предназначены для проведения локальных электрохимических реакций. При изготовлении зондов данного вида в качестве основы используется "базовый" тип зондов. Материал покрытия - полиимид. Материал зондов - вольфрам, платиново-иридиевый сплав.

Тип зонда	Материал	Диаметр, мм	Длина, мм	Покрытие
W/P-5.10	W	0.5	10	полиимид

### **Зонды со сверхузкими вершинами**

Зонды со сверхузкими вершинами предназначены для проведения исследований поверхностей с резким перепадом I высот (кристаллографических ступеней, сколов, углублений) и применения в многозондовых установках с независимым позиционированием зондов. Рабочее острие зондов формируется на основе методов локального осаждения и ионного травления. Зонд со сверхузкой вершиной представляет собой осажденное острие на металлическом основании. В качестве основания используется металлическая проволока диаметром 0,2-0,5 мм. Разрабатывается несколько перспективных базовых типов: Зонды с металлическим острием, осажденным из газообразных металлоорганических соединений на металлическом основании вольфрамовой проволоке; Зонды с острием из проводящих оксидов, осажденным на металлическом основании - медной проволоке.

### **Литература**

1. Лускинович П.Н. и др. Нанотехнологические процессы и установки. - Математическое моделирование нанотехнологических процессов и наноструктур, Труды научного семинара, вып. 1.-Москва, 2001
2. Неволин В.К. Физические основы туннельно-зондовой нанотехнологии.- Учебное пособие / М.: МИЭТ, 2000 - 69 с.
3. Биннинг Г., Рорер Г. Сканирующая туннельная микроскопия - от рождения к юности. - Успехи физических наук, т.154,вып. 2.- М.: Наука, 1988-с. 261-278.
4. П.А.Ващенко, Е.Н.Ивашов, С.В.Степанчиков. Система виброзащиты. Патент РФ на полезную модель №43047. – Опубл. 27.12.2004. Бюл №36.

## **Глава 8. Приповерхностные свойства нанообъектов , гетероструктур и нанобластей с различным химическим и фазовым составом**

### **Введение**

Изучение наноструктур знаменует новый этап в развитии физики и биологии, поскольку по своему размеру они занимают промежуточное положение между молекулами и микроскопическими объектами. Они содержат поддающееся счету число атомов и, следовательно, подходят для решения технологических задач на атомном уровне.

Последние 30 лет микроструктура поверхности твердых тел интенсивно изучалась методами дифракции и рассеяния электронных и ионных пучков, а также электронной спектроскопии. Однако большинство этих методов первоначально разрабатывалось для исследования объемной структуры твердых тел, поэтому они не всегда годятся для получения информации о структуре поверхности, тем более на атомном уровне.

Изучение природы сил трения на атомарном уровне стало возможным после создания атомно-силовых микроскопов . За десятилетний отрезок времени интенсивное развитие этой принципиально новой техники физических исследований привело к становлению чрезвычайно богатой области физики — нанотрибологии, объединившей экспериментальное и теоретическое изучение адгезии, трения, износа и смазки, химической активности и трибоэлектромагнетизма на наноструктурном уровне.

Изучение трения и связанных с ним процессов на наноструктурном уровне представляет значительный интерес для широкого спектра технических применений, таких как технология производства и поверхностного покрытия жестких магнитных дисков для компьютеров, изготовление микро-сенсорных датчиков и т.д. Эти и традиционно важные инженерные применения трибологии в машиностроении требуют более глубокого понимания свойств материалов на атомарном уровне с целью оптимизации и прогнозирования трибологических характеристик трущихся поверхностей. Успешное решение этих задач требует сокращения существующего разрыва между макро- и наноструктурными уровнями знаний о свойствах материалов вообще, и о трении в частности.

Вместе с атомно-силовой и фрикционной микроскопией (АФМ) важное место в изучении нанотрибоэффектов занимают также исследования процессов наноиндентирования поверхностей.

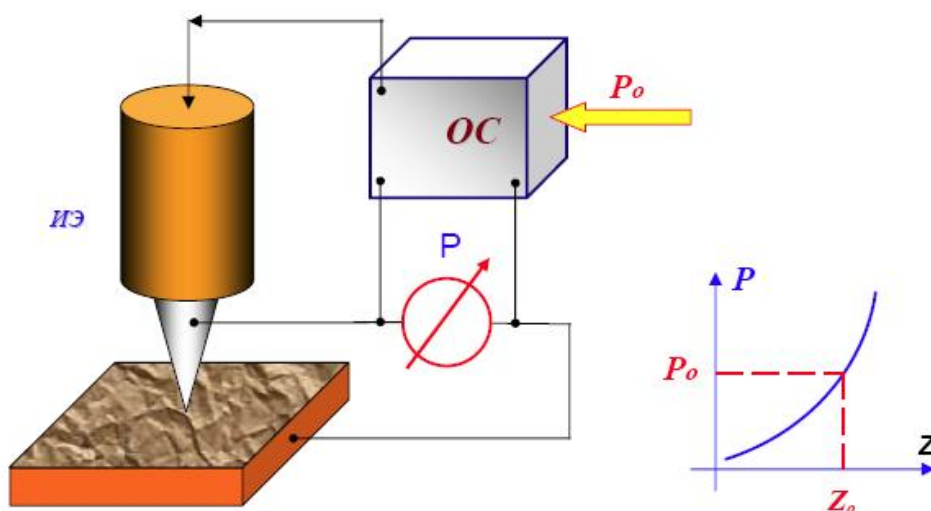
## 8.1. Основные принципы сканирующей зондовой микроскопии

В сканирующих зондовых микроскопах исследование микрорельефа поверхности и ее локальных свойств проводится с помощью специальным образом подготовленных зондов в виде игл. Рабочая часть таких зондов (острие) имеет размеры порядка десяти нанометров. Характерное расстояние между зондом и поверхностью образцов в зондовых микроскопах по порядку величины составляет 0,1-10нм.

Контакт зонда и образца подразумевает их взаимодействие. Чтобы осуществлять исследование с помощью конкретного прибора, из широкого спектра выбирается какое-либо одно рабочее взаимодействие. Природа этого выбранного взаимодействия и определяет принадлежность прибора к тому или иному типу в рамках семейства зондовых микроскопов. Информация о поверхности извлекается путем фиксации (при помощи системы обратной связи (рис. 8.1)) или детектирования взаимодействия зонда и образца.

В основе атомно-силовой микроскопии лежит взаимодействие зонда и образца с силами притяжения или отталкивания. Миниатюрные размеры зонда и высокая чувствительность детектирующей системы микроскопа позволяет достигать нано- и субнанометрового пространственного разрешения при детектировании поверхностных свойств (разрешающая способность прибора, как правило, тем выше, чем более короткодействующий характер имеет взаимодействие зонда и образца.)

Рассмотрим общие черты, присущие различным зондовым микроскопам. Пусть взаимодействие зонда с поверхностью характеризуется некоторым параметром  $P$ . Если существует достаточно резкая и взаимно однозначная зависимость параметра  $P$  от расстояния зонд-образец  $P=P(z)$ , то данный параметр может быть использован для организации системы обратной связи (ОС), контролирующей расстояние между зондом и образцом.



**Рис.8.1.** Схема организации системы обратной связи зондового микроскопа

Система обратной связи поддерживает значение параметра  $P$  постоянным, равным величине  $P_0$ , задаваемой оператором. Если расстояние зонд-поверхность изменяется (например, увеличивается), то происходит изменение (увеличение) параметра  $P$ . В системе ОС формируется разностный сигнал, пропорциональный величине  $\Delta P = P - P_0$ , который усиливается до нужной величины и подается на исполнительный элемент ИЭ. Исполнительный элемент обрабатывает данный разностный сигнал, приближая зонд к поверхности или отодвигая его до тех пор, пока разностный сигнал не станет равным нулю. Таким образом можно поддерживать расстояние зонд-образец с высокой точностью. В существующих зондовых микроскопах точность удержания расстояния зонд-поверхность достигает величины  $\sim 0,01 \text{ \AA}$ . При перемещении зонда вдоль поверхности образца происходит изменение параметра взаимодействия  $P$ , обусловленное рельефом поверхности. Система ОС обрабатывает эти изменения, так что при перемещении зонда в плоскости  $X, Y$  сигнал на исполнительном элементе оказывается пропорциональным рельефу поверхности.

Процесс сканирования осуществляется при помощи пьезокерамического манипулятора (или системы манипуляторов). Зонд движется последовательно, строка за строкой, вдоль поверхности (изменяются координаты  $X$  и  $Y$ ). Для оцифровки данных участок сканирования разбивается на  $N$  строк, а каждая строка на  $M$  точек, таким образом, положение иглы в плоскости  $X, Y$  описывается двумя координатами  $X_i, Y_j$  из множества  $\{X_i, Y_j\}$   $N \times M$  точек (обычно выбирают  $N=M$ ). Результатом работы сканирующего зондового микроскопа является установление соответствия между каждой парой координат из множества  $\{X_i, Y_j\}$  и некоторым числовым значением (или рядом значений), характеризующим анализируемый параметр поверхности (или ряд параметров).

По способу движения зонда над поверхностью можно провести следующую дифференциацию работы СЗМ.

- Если зонд движется над поверхностью при постоянной координате  $Z$ , то говорят, что сканирование осуществляется по способу *постоянной высоты*. В этом случае в каждой точке из множества  $\{X_i, Y_j\}$  измеряется *интенсивность* рабочего взаимодействия  $\Phi_{ij|Z=const}$ . Результатом исследования является массив (или несколько массивов, если измерялось несколько параметров)  $\{\Phi_{ij|Z=const} X_i, Y_j\}$ , описывающий зависимость функции двух переменных  $\Phi|_{Z=const}(X, Y)$ .

- Если же *система обратной связи* фиксирует в процессе сканирования на заданном уровне величину рабочего взаимодействия  $A(X, Y, Z)$  вариацией вертикальной  $Z$  координаты зонда, то говорят, что сканирование осуществляется по способу *постоянного взаимодействия*. Результатом работы СЗМ в этом режиме будет массив  $\{Z_{ij|A=const} X_i, Y_j\}$ , коррелирующий с топографией исследуемой поверхности. Помимо «топографического» массива, можно, проводя в каждой точке измерения какого-либо дополнительного параметра (или нескольких), получать зависимости вида  $\Phi_{ij|A=const}(X_i, Y_j)$ .

Таким образом, результатом СЗМ-исследования является получение функциональных зависимостей двух типов: по способу постоянной высоты:  $\Phi|_{Z=const}(X, Y)$  и по способу постоянного взаимодействия:  $Z|_{A=const}(X, Y)$  («топография»), плюс какая-либо дополнительная зависимость  $\Phi|_{A=const}(X, Y)$ . С помощью компьютерного программного обеспечения можно проводить анализ полученных зависимостей (анализ характерных латеральных и вертикальных размеров поверхностных особенностей, построение сечений, фурье-анализ, оценка шероховатости и т.п.), отображать полученные зависимости на экране монитора и выводить их на принтер.

Следует учитывать отличие «топографического» массива, полученного в режиме постоянного взаимодействия:  $\{Z_{ij}|_{A=const}, X_i, Y_j\}$  от реальной топографии поверхности. В случае неоднородного распределения поверхностных свойств, определяющих интенсивность взаимодействия зонда и образца, для извлечения точной информации о топографии объекта необходимо в каждой точке проводить дополнительный анализ взаимодействия зонда и образца.

## **8.2. Силовое взаимодействие зонда с поверхностью**

При сканировании зондом поверхности исследуемого образца на зонд действуют различные силы, возникающие в приповерхностных областях образца. Эти силы можно регистрировать в различных режимах работы зондового микроскопа. Но при исследовании каждого конкретного вида взаимодействия остальные исключить невозможно, поэтому при получении экспериментальных данных необходимо проводить их обработку, чтобы окончательные результаты исследования соответствовали действительности.

Для анализа получаемых результатов необходимо представление о видах силового взаимодействия зонда с исследуемой поверхностью.

### **8.2.1. Электростатическое взаимодействие**

Даже если к зонду и образцу не приложена разность потенциалов, различные участки поверхности могут содержать не скомпенсированные поверхностные заряды, находящиеся в материалах зонда и образца. Также друг с другом будут взаимодействовать диполи, находящиеся в материале зонда и образца. Эти взаимодействия носят силовой характер, и тем самым могут приводить к искажению других измеряемых величин в процессе сканирования зондом по поверхности образца.

#### Отдельные заряды

Сила взаимодействия двух электрических зарядов выражается известным законом Кулона:

$$F = \frac{kq_1q_2}{\epsilon r^2},$$

где  $q_1, q_2$  – электрические заряды,  $r$  – расстояние между ними,  $\varepsilon$  – диэлектрическая проницаемость,  $k = 9 \cdot 10^9$  м/Ф.

Свободные заряды индуцируют поверхностные заряды на ближайших поверхностях, известным и наглядным образом описания которых является метод электрических изображений. Его удобство состоит в том, что наведенное распределение зарядов можно заменить одним или несколькими точечными зарядами-изображениями и затем для расчета силы взаимодействия пользоваться законом Кулона. В случае заряда над плоскостью этот метод приводит к следующей формуле:

$$F = \frac{-kq^2}{4r^2} \left( \frac{\varepsilon_s - \varepsilon_m}{\varepsilon_s + \varepsilon_m} \right),$$

где  $\varepsilon_m, \varepsilon_s$  – диэлектрическая проницаемость среды и образца соответственно, а  $r$  – расстояние от заряда до плоскости. В случае металлов диэлектрическая проницаемость  $\varepsilon_s$  бесконечна, и поэтому выражение в скобках равно единице, и взаимодействие сводится к кулоновскому притяжению заряда к своему заряду-изображению другого знака, находящегося на расстоянии  $2r$ . Другой практический случай – погружение системы в жидкую среду с некоторым  $\varepsilon_m$ . Электростатическое взаимодействие может значительно уменьшиться и даже превратиться в отталкивание при  $\varepsilon_m > \varepsilon_s$ .

### Диполи

Помимо нескомпенсированных зарядов в толще кантилевера и образца для анализа электростатического взаимодействия необходимо принимать во внимание диполи.

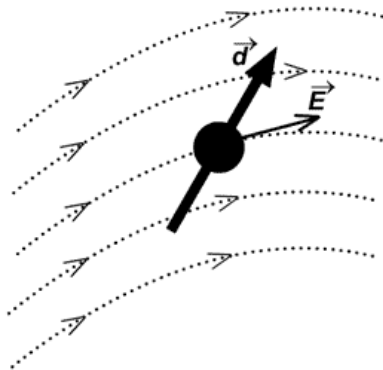
Энергия диполя  $d$ , помещенного в поле  $E$  (рис. 8.2),

$$W_D = -dE,$$

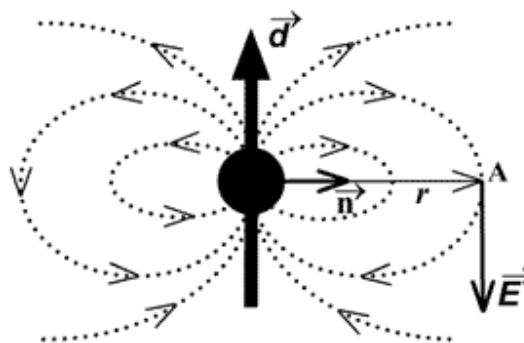
Электрическое поле, создаваемое диполем (рис. 8.3),

$$E = \frac{(3(\mathbf{n}d)\mathbf{n} - d)}{r^3}, \quad |E| \sim \frac{1}{r^3},$$

где  $\mathbf{n}$  – единичный вектор в направлении на диполь из точки, где ищется поле.



**Рис. 8.2.** Диполь, помещенный в электрическое поле



**Рис. 8.3.** Электрическое поле, созданное диполем

Сила взаимодействия двух диполей убывает с расстоянием гораздо быстрее кулоновской:

$$F = \frac{3d_1d_2 - (nd_1)(nd_2)}{r^4},$$

где  $d_1, d_2$  – векторы дипольных моментов,  $r$  – расстояние между ними, а  $n$  – вектор нормали вдоль линии, их соединяющей. Малость силы дипольного взаимодействия по сравнению с Кулоновской компенсируется большим количеством диполей в веществе, в роли которых выступают молекулы. Полярные молекулы, один атом или группа атомов которых более электроотрицательны, чем другие, обладают дипольным моментом.

### Контактные явления Работа выхода полупроводников

Работой выхода  $A$  (термодинамической работой выхода) электрона из полупроводника называют наименьшую энергию, которую нужно затратить для удаления электрона из полупроводника в вакуум, если начальная энергия электрона в полупроводнике равна электрохимическому потенциалу. Уровень энергии, равной электрохимическому потенциалу, обычно называют уровнем химического потенциала или уровнем Ферми полупроводника. Его положение по отношению к дну зоны проводимости определяется как и для металлов, значением химического потенциала  $\mu$  (обычно  $\mu < 0$ , т.е. уровень Ферми в полупроводниках, в отличие от металлов, располагается ниже дна зоны проводимости).

Внешней работой выхода  $A_{внеш}$  электрона из полупроводника называют наименьшую работу, которую нужно затратить, для удаления в вакуум электрона, начальная энергия которого в полупроводнике равна энергии, соответствующей дну зоны проводимости. Связь между  $A$  и  $A_{внеш}$  имеет вид  $A = A_{внеш} - \mu$ . Внешняя работа выхода определяется свойствами кристаллической решетки полупроводника и для различных полупроводников составляет от 1 до 6 эВ.

Для чистого (беспримесного) полупроводника

$$\mu = -\frac{\Delta W_0}{2} + \frac{kT}{2} \ln \frac{v_h}{v_e},$$

где  $\Delta W_0$  – энергия активации собственной проводимости,  $k$  – постоянная Больцмана,

$$v_e = \frac{2}{h^3} (2\pi m_e^* kT)^{3/2},$$

$$v_h = \frac{2}{h^3} (2\pi m_h^* kT)^{3/2},$$

$h$  – постоянная Планка,  $m_e^*, m_h^*$  – эффективные массы электрона и дырки. При  $T=0^\circ K$   $\mu = -\Delta W_0/2$ , т.е. уровень Ферми располагается посередине запрещенной зоны.

### Контакт металл-полупроводник

Если зонд выполнен из полупроводникового материала, а поверхность исследуемого образца металлическая, то при подводе зонда микроскопа к исследуемой поверхности электроны начинают переходить из тела с меньшей работой выхода в тело с большей работой выхода. При этом между металлом и полупроводником устанавливается контактная разность потенциалов, обусловленная двойным электрическим слоем, образующимся в области контакта и называемым контактным слоем. Вследствие малой концентрации носителей тока в полупроводнике (порядка  $(10^{14} \div 10^{16}) \text{ см}^{-3}$  вместо  $10^{22} \text{ см}^{-3}$  в металлах) толщина контактного слоя в полупроводнике достигает  $(10^{-5} \div 10^{-4}) \text{ см}$ , т.е. на несколько порядков больше толщины контактного слоя в металлах. В пределах этого слоя в полупроводнике распределен объемный заряд и существует контактное электрическое поле. Знак объемного заряда зависит от соотношения между работами выхода электрона из металла ( $A_1$ ) и из полупроводника ( $A_2$ ): если  $A_1 > A_2$ , то объемный заряд положительный, а если  $A_1 < A_2$ , - отрицательный.

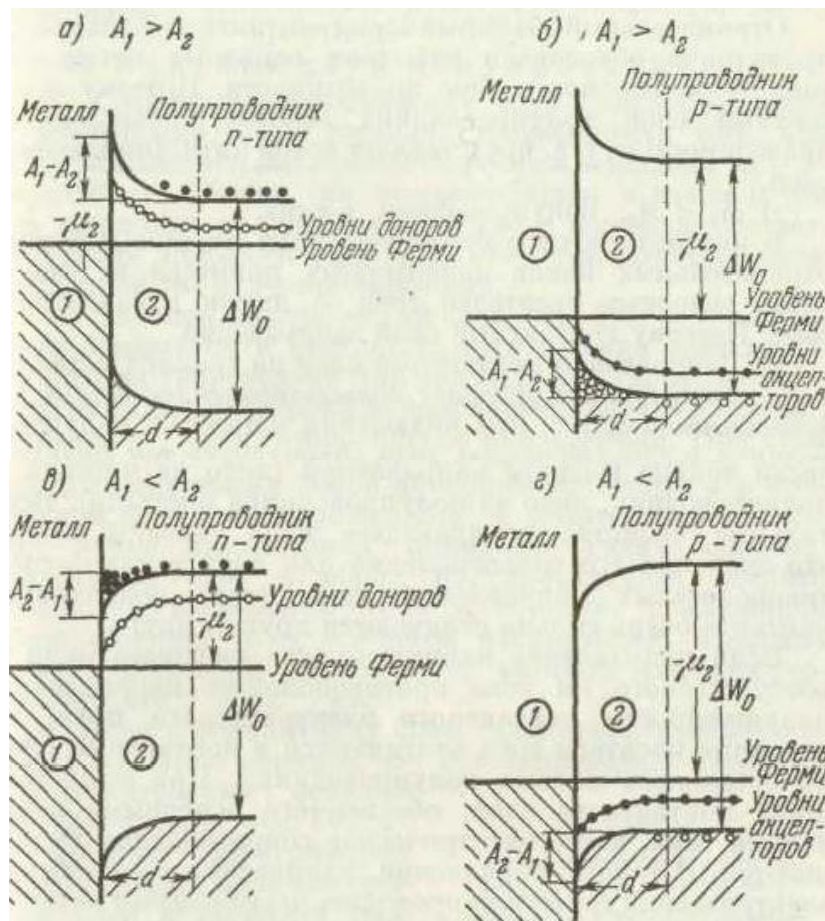
Вследствие существования электрического поля контактного слоя потенциал в точках этого слоя полупроводника при  $A_1 > A_2$  ниже, чем в остальном объеме полупроводника, а при  $A_1 < A_2$  - выше. Поэтому, если  $A_1 > A_2$ , то при прочих равных условиях энергия электронов в контактном слое полупроводника больше, чем в остальном объеме. Так как электрохимический потенциал во всех частях полупроводника одинаков, то в контактном слое дно зоны проводимости поднимается вверх, удаляясь от уровня Ферми, а верхний край валентной зоны также поднимается вверх, приближаясь к уровню Ферми. При этом ширина запрещенной зоны  $\Delta W_0$  между верхним краем валентной зоны и дном зоны проводимости остается такой же, как и в остальном объеме полупроводника.

Если  $A_1 < A_2$ , то направление вектора напряженности контактного электрического поля изменяется на противоположное, так что потенциал в точках контактного слоя полупроводника выше, чем в остальном объеме. Поэтому в контактном слое дно зоны проводимости опускается вниз, приближаясь к уровню Ферми, а верхний край валентной зоны также опускается вниз, удаляясь от уровня Ферми.

Возможны следующие четыре случая контакта металла с примесными полупроводниками (рис. 8.4).

**$A_1 > A_2$ , полупроводник  $n$ -типа.** Объемный положительный заряд обусловлен избытком в контактном слое полупроводника положительных ионов донорных примесей. Контактный слой полупроводника обеднен основными носителями тока - электронами в зоне проводимости. Поэтому удельное сопротивление контактного слоя во много раз больше удельного сопротивления остальной части полупроводника. Такой контактный слой называют запирающим.





**Рис. 8.4.** ● - электроны в зоне проводимости и отрицательные ионы акцепторных примесей; ○ - дырки в валентной зоне и положительные ионы донорных примесей;  $d$  - толщина контактного слоя

$A_1 > A_2$ , **полупроводник  $p$ -типа.** В контактном слое полупроводника имеется избыток основных носителей тока – дырок в валентной зоне. Поэтому контактный слой обладает повышенной проводимостью. Такой контактный слой называют антизапирающим.

$A_1 < A_2$ , **полупроводник  $n$ -типа.** Отрицательный объемный заряд контактного слоя полупроводника обусловлен избытком основных носителей тока – электронов в зоне проводимости. Поэтому контактный слой полупроводника обладает повышенной проводимостью, т.е. представляет собой антизапирающий слой.

$A_1 < A_2$ , **полупроводник  $p$ -типа.** В контактном слое полупроводника имеется избыток отрицательных ионов акцепторных примесей и недостаток основных носителей тока – дырок в валентной зоне. Поэтому контактный слой запирающий.

### Контакт полупроводник-полупроводник

Граница соприкосновения двух полупроводников с различными,  $n$ - и  $p$ -типами проводимости называется электронно-дырочным переходом. Двойной электрический слой  $p$ - $n$ -перехода образуется в результате перемещения электронов из  $n$ - в  $p$ -полупроводник (так как  $A_p > A_n$ , где  $A_p$  и  $A_n$  – работы выхода электрона из полупроводников  $p$ - и  $n$ -типа), а дырок – в противоположном направлении. При этом в контактном поле  $n$ -полупроводника оказывается избыток положительных ионов донорных примесей, а в контактном слое  $p$ -полупроводника – избыток отрицательных ионов акцепторных примесей. Таким образом, контактный слой в обоих полупроводниках обеднен основными носителями тока и обладает пониженной проводимостью, т.е. является запирающим. Толщина  $d$   $p$ - $n$ -перехода в практически важных случаях составляет  $(10^{-4} \div 10^{-5})$  см.

### Трибоэлектрический эффект

Также, разность потенциалов между проводниками может появиться в следствие возникновения электрических зарядов при трении – трибоэлектрический эффект. Этот эффект наблюдается при взаимном трении двух диэлектриков, полупроводников или металлов различного химического состава или одинакового состава, но разной плотности, при трении металлов о диэлектрики, при трении двух одинаковых диэлектриков, при трении жидких диэлектриков друг о друга или о поверхность твёрдых тел и др. При этом электризуются оба тела; их заряды одинаковы по величине и противоположны по знаку.

При трении двух химически одинаковых тел положительные заряды получает более плотное из них. При трении двух диэлектриков положительно заряжается диэлектрик с большей диэлектрической проницаемостью  $\epsilon$ . Тела можно расположить в трибоэлектрические ряды, в которых предыдущее тело электризуется положительно, а последующее — отрицательно (ряд Фарадея: (+) мех, фланель, слоновая кость, перья, горный хрусталь, флинтглас, бумажная ткань, шёлк, дерево, металлы, сера (-)). Для диэлектриков, расположенных в трибоэлектрический ряд, наблюдается убывание твёрдости (ряд Гезе-хуса: (+) алмаз (твёрдость 10), топаз (8), горный хрусталь (7), гладкое стекло (5), слюда (3), кальцит (3), сера (2), воск (1) (-)). Для металлов характерно возрастание твёрдости. У жидких диэлектриков положительный заряд приобретает вещество с большей  $\epsilon$  или поверхностным натяжением.

У твёрдых тел трибоэлектричество объясняется переходом носителей тока при трении от одного тела к другому. В случае двух металлов, двух полупроводников или металла и полупроводника оно обусловлено переходом электронов от вещества с меньшей работой выхода к веществу с большей. При контакте металла с диэлектриком трибоэлектричество возникает за счёт перехода электронов из металла в диэлектрик и перехода ионов того или

иною знака из диэлектрика на поверхность металла. При трении двух диэлектриков - диффузия электронов и ионов. Существенную роль может играть также разное нагревание тел при трении, что вызывает переход носителей тока с локальных неоднородностей более нагретой поверхности. При трении двух диэлектриков, отличающихся только плотностью, из более плотного вещества будет диффундировать больше электронов и оно зарядится положительно; при контакте двух разных диэлектриков из вещества с большей  $\varepsilon$  будет переходить больше электронов.

### 8.2.2. Сила Ван-дер-Ваальса

При дальнейшем сближении зонда АСМ и исследуемой поверхности, на зонд начинают действовать межмолекулярные силы Ван-дер-Ваальса. Эта сила имеет три составляющие  $W_{ориент}$ ,  $W_{инд}$ ,  $W_{дисп}$ .

- $$W_{ориент} = \frac{\int W_D \exp\left(-\frac{W_D}{kT}\right) dv}{\int \exp\left(-\frac{W_D}{kT}\right) dv}, \text{ где } W_D = \frac{d_1 d_2 - 3(d_1 n)(d_2 n)}{r^3} \sim \frac{1}{r^3};$$

- $$W_{инд} = d_{инд} E = \chi E^2 = \chi \frac{-3(nd_1)^2 + d_1^2}{r^6} = -\frac{\chi d_1^2}{2r^6} \sim \frac{1}{r^6},$$

где  $d_{инд} = \chi E$  – индукционный момент;  $\chi$  - поляризуемость молекулы;

- $$W_{дисп} = -\frac{A_3}{r^6}, \text{ где } A_3 = \frac{3I_1 I_2}{2(I_1 + I_2)} \chi_1 \chi_2,$$

где  $I_1, I_2$  – потенциалы ионизации;  $\chi_1, \chi_2$  - поляризуемости молекул.

Эти составляющие обладают несколько отличной физической природой, но их потенциал зависит от расстояния между молекулами одинаково — как  $1/r^6$ . Это обстоятельство позволяет непосредственно сравнивать константы взаимодействия, соответствующие трем составляющим силы Ван-дер-Ваальса, причем по причине их одинаковой зависимости от расстояния, пропорция между компонентами будет сохраняться при различных  $r$ . Сами же константы при множителе  $1/r^6$  будут отличаться для разных веществ.

$$W = W_{ориент} + W_{инд} + W_{дисп} \sim 1/r^6.$$

В основе всех трех составляющих силы Ван-дер-Ваальса лежит взаимодействие диполей. Проводя оценки для типичных условий АСМ-эксперимента в режиме контакта сила ван-дер-ваальсовского притяжения:

$$F_{ВдВ} \sim 10^{-8} \div 10^{-9} \text{ Н.}$$

Сила соответствующая потенциалу парного взаимодействия равна производной по расстоянию  $r$ :

$$f = \frac{6A_3}{r^7},$$

где  $A_3$  — константа Гамакера.

Исходя из этого микроскопического описания, можно определить силу притяжения, возникающую между зондом и образцом в целом. Оно равно сумме всех парных взаимодействий молекул кантилевера и исследуемой поверхности:

$$F = \sum_{\text{по молекулам зонда}} \left( \sum_{\text{по молекулам образца}} f \right).$$

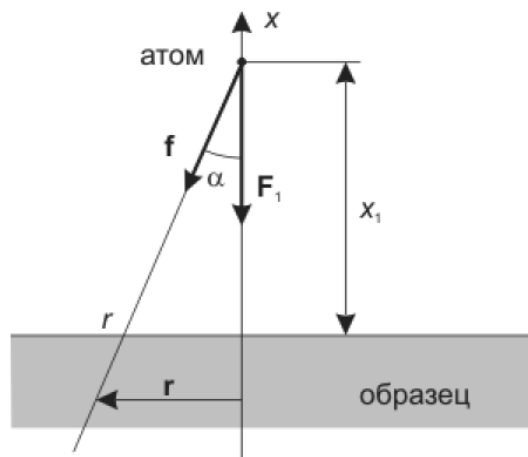
Понятно, что при этом результат будет существенно зависеть от пространственной геометрии задачи.

Пренебрегая дискретным распределением взаимодействующих центров (молекул), несложно от парного суммирования перейти к двойному интегралу:

$$F = \int n_1 dV_1 (\int n_2 dV_2 f),$$

где  $n_1$  и  $n_2$  - концентрации молекул зонда и образца (плотности).

Физический смысл внутреннего интеграла — сила взаимодействия одной молекулы с плоскостью. Сила притяжения очень быстро падает с увеличением расстояния ( $r^{-7}$ ), поэтому отдаленные части системы не вносят в интеграл существенного вклада. Благодаря этому, интегрирование можно распространить на все полупространство, как если бы оно было однородным образцом.



**Рис. 8.5.** Система атом - плоский образец

Для вычислений внутреннего интеграла (обозначив его через  $F_1$ ) вводится цилиндрическая система координат (рис. 8.5). Начало отсчета совмещается с молекулой. Из соображений симметрии понятно, что результирующая сила будет направлена вертикально вниз. При этом горизонтальные составляющие силы притяжения к двум молекулам, расположенным симметрично относительно оси  $x$ , компенсируются. Поэтому проще изначально учитывать лишь вертикальную компоненту силы

$$f_1 = \cos \alpha \frac{6A_3}{r^7}.$$

Эта сила будет одинакова для всех точек кольца с радиусом  $\rho$ , поэтому интегрирование по углу вокруг оси  $x$  сводится просто к умножению на  $2\pi$ . Дальнейшие вычисления довольно просты:

$$F_1 = n_2 \int_0^{\infty} 2\pi\rho d\rho \int_{x_0}^{\infty} dx \frac{z}{\sqrt{\rho^2 + x^2}} \frac{6A_3}{\left(\sqrt{\rho^2 + x^2}\right)^7} = \frac{\pi A_3 n_2}{2x_0^4}.$$

Чтобы взять внешний интеграл в (3.2.8) нужно проинтегрировать по объему зонда:

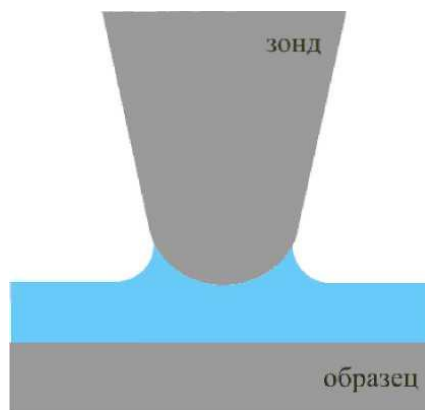
$$F = \int F_1 n_1 dV_1.$$

Дальнейшие вычисления необходимо проводить для конкретной модели острия кантилевера.

### 8.2.3. Капиллярные силы

В случае, если исследование поверхности проводится на воздухе, то на поверхности исследуемого образца обязательно присутствует жидкая пленка, состоящая в основном из воды.

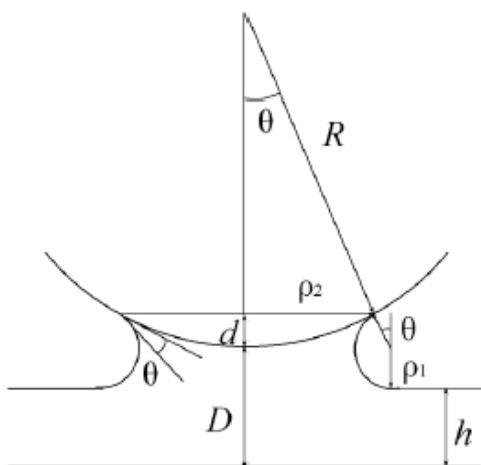
В тот момент, когда кантилевер касается жидкой пленки на плоской поверхности, происходит смачивание поверхности кантилевера (рис. 8.6), т.к. энергетически более выгоден контакт жидкости с материалом кантилевера (если он гидрофильный), чем с воздухом.



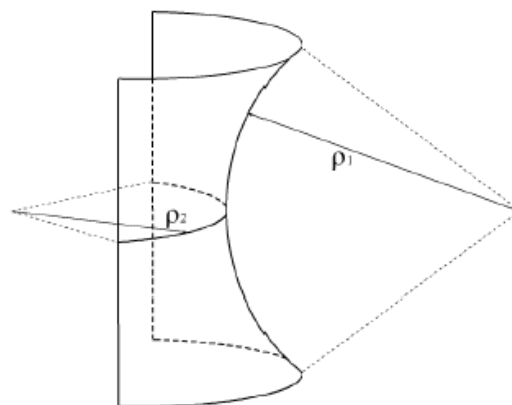
*Рис. 8.6. Образование «шейки»*

Искривленная поверхность шейки будет стремиться выпрямиться, что возможно только за счет опускания кантилевера. Это означает возникновение притяжения кантилевера к образцу.

Расчет этой силы довольно прост. Пусть радиус кривизны зонда намного превосходит все остальные характерные размеры задачи (рис. 8.7).



**Рис. 8.7.** К расчету капиллярной силы



**Рис. 8.8.** Пояснение к формуле Лапласа

На рис. 8.7 введены следующие обозначения:  $D$  — расстояние зонд-образец,  $d$  — «глубина погружения»,  $h$  — толщина пленки,  $\rho_1$  — меньший радиус кривизны жидкой поверхности,  $\rho_2$  — радиус площадки соприкосновения кантилевера с жидкостью.

Согласно формуле Лапласа, давление в жидкости будет меньше атмосферного на величину:

$$\Delta p = \sigma \left( \frac{1}{\rho_1} + \frac{1}{\rho_2} \right) \approx \frac{\sigma}{\rho_1}, \quad \rho_1 \ll \rho_2 .$$

Это давление приложено к поверхности соприкосновения зонда с жидкостью площадью  $\pi \rho_2^2 \approx 2\pi R d$  (рис. 8.8). Сила притяжения кантилевера к образцу, обусловленная капиллярным эффектом, составляет:

$$F = P \cdot 2\pi R d = 2\pi R d \frac{\sigma}{\rho_1} .$$

Для простоты можно считать краевые углы  $\theta$  для образца и зонда равными. Тогда  $\rho_1 \cos \theta + \rho_1 = D + d - h$ . Таким образом

$$F_{\text{кан}} = \frac{2\pi R \sigma (1 + \cos \theta)}{1 + \frac{D - h}{d}} . \quad (1)$$

Если можно пренебречь толщиной жидкой пленки ( $h=0$ ), то справедливо равенство  $2\rho_1 \cos \theta \approx D + d$ , и формула (1) упрощается до формулы

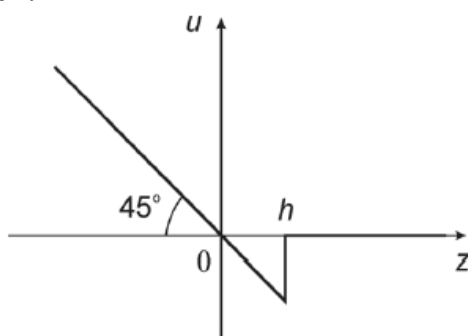
$$F_{\text{кан}} = \frac{4\pi R \sigma \cos \theta}{1 + \frac{D}{d}} .$$

Для оценок достаточно воспользоваться максимальным значением капиллярной притягивающей силы, которое достигается при  $D = 0$ . При этом из формулы исчезает неизвестный параметр  $d$ :

$$F_{\text{кан}} = F_{\text{макс}} = 4\pi R \sigma \cos \theta .$$

Из данной формулы можно установить, что для кантилевера радиус  $R$  которого равен  $10\text{ нм}$ , капиллярная сила  $F_{\text{кан}} \sim 10^{-8} - 10^{-9}$ . Таким образом, по порядку величины капиллярная сила совпадает с взаимодействием Ван-дер-Ваальса и электростатическими силами.

При подводе и отводе кантилевера от поверхности образца возникает гистерезис. При движении вверх перетяжка остается дольше. Поверхность кантилевера уже смочена, и жидкая шейка тянется за зондом. Когда связи прерываются, капиллярная сила притяжения перестает действовать и кантилевер резко выпрямляется.



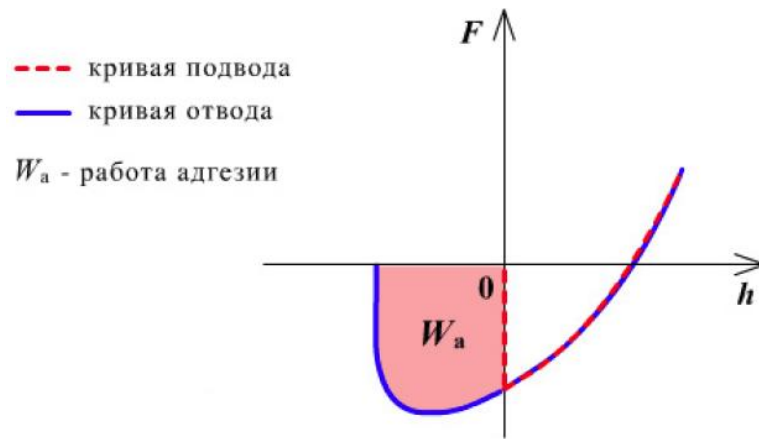
*Рис. 8.9. График зависимости деформации кантилевера от расстояния зонд-образец*

На рис. 8.9 по вертикальной оси отложена деформация кантилевера  $u$ , а по горизонтальной — расстояние  $z$ , на которое он подводится к образцу. На больших расстояниях деформация отсутствует, график горизонтален. В левой части зависимость линейна и имеет наклон  $45^\circ$ . Это означает контакт с поверхностью, т.к. острие уперлось в образец, и возникающая деформация равна величине прижатия зонда к поверхности.

При отведении кантилевера острие залипает: капиллярная сила удерживает его, и линейная зависимость продолжается уже ниже горизонтальной оси до тех пор, пока перетяжка не исчезнет. После чего кантилевер скачком выпрямляется до недеформированного состояния.

#### 8.2.4. Адгезионные силы

В промежуточной области, когда между одними парами молекул зонда и образца действуют силы притяжения (потенциал  $-1/r^6$ ), а между другими — отталкивания  $1/r^{12}$ , точное нахождение силы взаимодействия всего зонда с образцом невозможно. Более того, в переходной области возникает качественно новое явление — адгезия. Своим происхождением оно обязано наличию короткодействующих молекулярных сил. Следует различать две разновидности адгезии: зонд-жидкая пленка на поверхности и зонд-твердый образец. Если первый случай сводится к капиллярному взаимодействию, которому посвящен раздел 8.2.3, то причина адгезионных сил между зондом и твердым образцом — молекулярное электростатическое взаимодействие.



**Рис. 8.10.** Работа адгезии. Показаны зависимости адгезионных сил при подводе и отводе зонда от поверхности образца

Адгезия — неконсервативный процесс. Силы при подводе кантилевера к образцу отличаются от сил при отрыве зонда от поверхности. На подобную операцию будет потрачена работа, называемая работой адгезии (рис. 8.10).

Эта работа состоит из следующих компонентов:

$$W_a = W_a^d + W_a^p + W_a^i + W_a^h + W_a^\pi + W_a^{da} + W_a^e,$$

где  $d$  — дисперсионное взаимодействие Лондона,  $p$  — диполь-дипольное (ориентационное) взаимодействие,  $i$  — индукционное взаимодействие,  $h$  — водородная связь,  $\pi$  —  $\pi$ -связь,  $da$  — донорно-акцепторная связь,  $e$  — электростатическое взаимодействие. Причем первые три слагаемых — работа сил Ван-дер-Ваальса.

Причина адгезии — электростатические силы на границе раздела двух тел, возникающие благодаря образующемуся в контакте двойному заряженному слою. Его происхождение различно для материалов разных типов. Для металлов она определяется контактной разностью потенциалов, состояниями внешних электронов атомов поверхностного слоя и дефектами кристаллической решетки; для полупроводников — поверхностными состояниями и наличием примесных атомов; а для диэлектриков — дипольным моментом групп молекул на границе фаз.

Адгезия — необратимый процесс. При наличии контактной разности потенциалов, например, возникает дрейф электронов, который ведет к увеличению энтропии. Именно поэтому силы при сведении и разведении кантилевера и образца различаются (см. ниже), и процесс, тем самым, оказывается неконсервативным.

Для количественного описания адгезии используются приближенные модели.



### Модель DMT адгезии твердых тел

Модель DMT (*Derjagin, Muller, Toropov — 1975 г*) применима для зондов с малым радиусом закругления и большой жесткостью. При этом считается, что геометрия деформированных поверхностей мало отличается от той, что дает решение задачи Герца. Учет сил Ван-дер-Ваальса по периметру контактной площадки приводит к дополнительному притяжению между зондом и образцом, которое уменьшает силы упругого отталкивания.

Выражения для давления и глубины проникновения получают следующий вид:

$$F = \frac{Ka^3}{R} - 2\pi R\sigma, \quad h = \frac{a^2}{R},$$

где  $R$  - радиус закругления зонда,  $a$  — радиус контактной площадки,

$K \left( \frac{1}{K} = \frac{3}{4} \left( \frac{1-\mu'^2}{E'} + \frac{1-\mu^2}{E} \right) \right)$  — эффективный модуль Юнга,  $\sigma$  — работа адгезии (рис. 8.11).

### Модель JKR адгезии твердых тел

Модель JKR (*Johnson, Kendall, Roberts — 1964-1971 гг*) пригодна для зондов с большим радиусом кривизны и малым коэффициентом жесткости. Такие системы называются сильноадгезивными. Модель учитывает вклад ван-дер-ваальсовых сил в пределах контактной площадки.

За счет этого возникает притяжение, которое не только снижает силу упругого отталкивания (правая часть графика 8.11,в), но и приводит к образованию шейки ( $h < 0$ , левая часть графика) и отрицательной силе. График соответствует формулам:

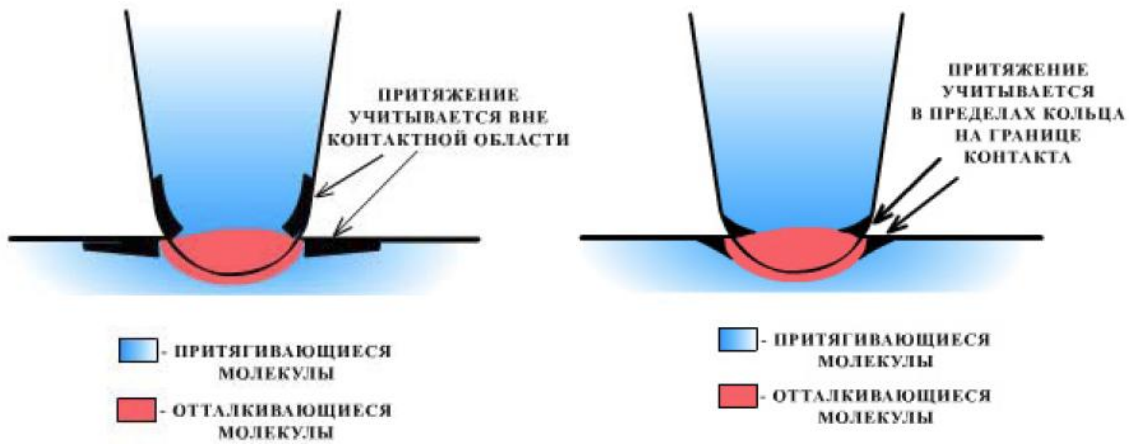
$$F = \frac{Ka^3}{R} - \sqrt{6\pi\sigma Ka^3}, \quad h = \frac{a^2}{R} - \frac{2}{3} \sqrt{\frac{6\pi\sigma a}{K}}.$$

### Модель Маугиса адгезии твердых тел

Механика Маугиса (*1992 г*) является наиболее сложным и точным подходом. Его можно использовать для любых систем: как с большой, так и с малой адгезией. Степень адгезии определяется параметром  $\lambda$ :

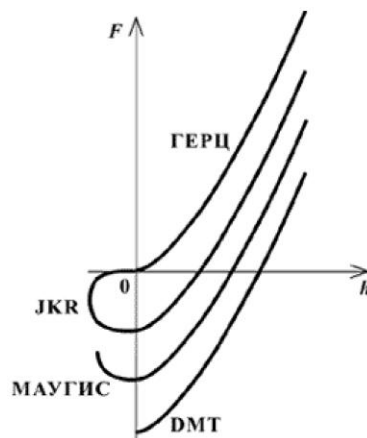
$$\lambda = \frac{2,06}{\xi_0} \left( \frac{R\sigma^2}{\pi K^2} \right)^{\frac{1}{3}},$$

где  $\xi_0$  — межатомное расстояние



**Рис. 8.11,а.** Применимость моделей JKR и DMT

**Рис. 8.11,б.** Применимость модели Маугиса



**Рис. 8.11,в.** Зависимость силы от глубины проникновения для задачи Герца, моделей JKR, DMT и Маугиса

Модели DMT и JKR являются предельными случаями механики Маугиса, отвечающие разным параметрам  $\lambda$ . Для жестких материалов (DMT)  $\lambda \rightarrow 0$ , для податливых (JKR)  $\lambda \rightarrow \infty$ .

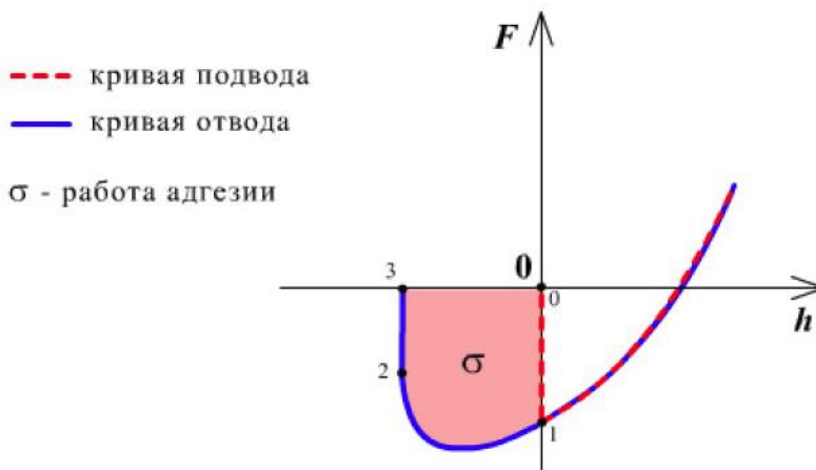
В модели Маугиса считается, что молекулярная сила притяжения действует в пределах кольцевой области на границе контактной площадки. Поправка Маугиса к решению задачи Герца выражается неявно, через параметр  $m$ :

$$1 = \frac{\lambda a^2}{2} \left( \frac{K}{\pi R^2 \sigma} \right)^{\frac{2}{3}} \left[ \sqrt{m^2 - 1} + (m^2 - 2) \arctan \sqrt{m^2 - 1} \right] + \frac{4\lambda a^2}{3} \left( \frac{K}{\pi R^2 \sigma} \right)^{\frac{1}{3}} \left[ 1 - m + \sqrt{m^2 - 1} \arctan \sqrt{m^2 - 1} \right];$$

$$F = \frac{Ka^3}{R} - \lambda a^2 \left( \frac{\pi \sigma K^2}{R} \right)^{\frac{2}{3}} \left[ \sqrt{m^2 - 1} + m^2 \arctan \sqrt{m^2 - 1} \right];$$

$$h = \frac{a^2}{R} - \frac{4\lambda a}{3} \left( \frac{\pi \sigma}{RK} \right)^{\frac{1}{3}} \sqrt{m^2 - 1}.$$

В модель JKR и механику Маугиса изначально заложено существование гистерезиса при подведении и отведении. Считается, что при подведении кантилевера сила притяжения возникает резко в момент касания, и система переходит от точки 0 в точку 1 (рис. 8.12). При отведении кантилевера, система «движется» по другой траектории 1-2 до момента отрыва, который происходит скачкообразно 2-3.



**Рис. 8.12.** Зависимость силы от глубины проникновения для моделей Герца, DMT, JKR и Маугиса при одном  $\lambda$

Наличие на графике петли 0-1-2-3 означает затрату работы на отрыв зонда от образца, которая равна площади петли. Это и есть работа адгезии  $\sigma$ .

### 8.2.5. Упругие взаимодействия зонда с образцом

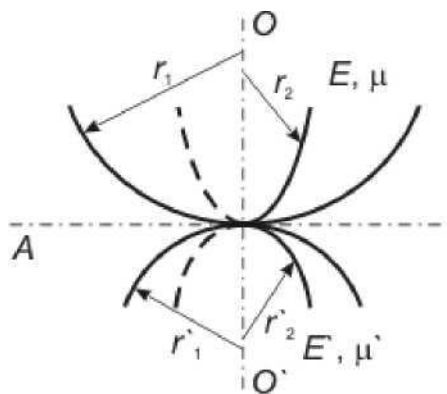
При исследовании любых приповерхностных свойств исследуемых объектов в зоне контакта зонда с образцом возникают механические напряжения, и для верной интерпретации получаемых результатов при сканировании необходимо учитывать деформации материала зонда и образца под воздействием таких механических напряжений. Для наноиндентирования задача определения механических нагрузок в зоне контакта зонда с исследуемой поверхностью и геометрических характеристик контактной зоны, является сутью данного метода исследований.

#### Задача Герца

Несмотря на то что инденторы номинируются как остроконечные пирамиды, острие реального индентора в первом приближении можно считать сферическим с эффективным радиусом  $R$ . Вследствии этого можно свести задачу определения нагрузок в контактной области к задаче Герца.

Решение задачи Герца в общем виде имеет неявный вид. Для получения численного решения контактной задачи Герца можно упростить условия данной задачи, сделав следующие предположения:

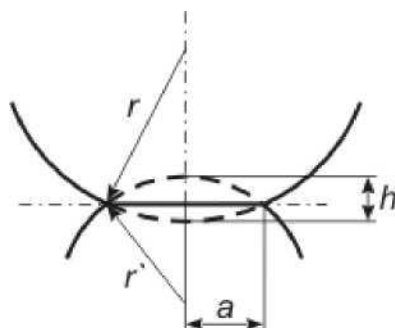
- Материалы, как образца, так и кантилевера, изотропны, а значит, все их упругие свойства описываются всего двумя парами параметров – модулями Юнга  $E, E'$  и коэффициентами Пуассона  $\mu, \mu'$ .
- Вблизи точки соприкосновения недеформированному участку поверхности каждого из тел можно приписать два радиуса кривизны  $r_1, r_2$  (для зонда) и  $r'_1, r'_2$  (для исследуемого участка образца) во взаимно перпендикулярных плоскостях, ортогональных самой поверхности в данной точке (рис. 8.13).
- Деформации малы по сравнению с радиусами кривизны поверхностей.



**Рис. 8.13.** Задача Герца ставится в предположении, что в точке касания до деформации поверхности описываются радиусами кривизны в двух взаимно перпендикулярных плоскостях, ортогональных касательной плоскости  $A$

### Решение задачи Герца

Чтобы получить общее представление о деформациях при упругом контакте и найти характерные числовые значения, достаточно ограничиться анализом взаимодействия сферических поверхностей — острия зонда и небольшого участка образца. Это означает, что  $r_1=r_2=r, r'_1=r'_2=r'$



**Рис. 8.14.** Связь между радиусом контактной площадки  $a$  и глубиной проникновения  $h$  в деформированном состоянии

При воздействии нагрузки соприкасающиеся тела деформируются так, что вместо точки касания образуется некоторая контактная площадка. Из осевой симметрии задачи следует, что эта площадка будет иметь форму круга с радиусом  $a$ .

Для расчетов будут удобны следующие величины:  $1/R=1/r+1/r'$ , а также эффективный модуль Юнга для заданной пары материалов:

$$\frac{1}{K} = \frac{3}{4} \left( \frac{1-\mu'^2}{E'} + \frac{1-\mu^2}{E} \right).$$

При малых деформациях (допущение 3 задачи Герца) выполняется следующее геометрическое соотношение между глубиной проникновения  $h$  и радиусом контактного круга  $a$ :

$$h = \frac{a^2}{2R},$$

которое можно понять из рис. 8.14.

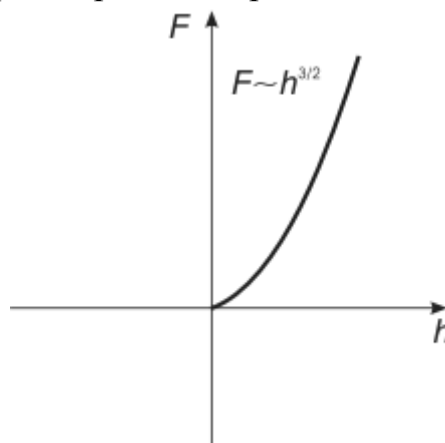
Решение задачи Герца дает соотношение между придавливающей силой  $F$  и глубиной проникновения  $h$ :

$$F = \frac{Ka^3}{R} = Kh^{\frac{3}{2}} R^{\frac{1}{2}}.$$

Соответственно, давление связано с силой следующим образом:

$$P = \frac{F}{\pi a^2} = \frac{1}{\pi} \sqrt[3]{\frac{FK^2}{R^2}}. \quad (2)$$

Приведенное решение для двух соприкасающихся сферических форм содержит важный частный случай: контакт плоского образца и зонда с радиусом кривизны  $R$  ( $r=R$ ,  $r'=\infty$ ). На рис. 8.15 изображено графическое решение задачи Герца. График определен при положительных  $h$ .



**Рис. 8.15.** Зависимость силы  $F$  от глубины проникновения  $h$

Решая задачу Герца в общем виде, можно получить решение для любых поверхностей, однако решение такой задачи будет иметь неявный вид. Тем не менее, результат по порядку величины будет такой же, что и в данном, крайне упрощенном случае. Поэтому приближенную оценку характерной величины давления в контакте можно провести на основе формулы (2). Типичные значения в АСМ: радиус зоны контакта до 10 нм, глубина проникновения до 20 нм, контактное давление до 10 ГПа.

Другое ограничение (условие 1 задачи Герца) состоит в рассмотрении соприкасающихся тел в рамках модели сплошной среды с изотропными характеристиками. Понятно, что на микроуровне молекулярная структура уже играет существенную роль, поэтому подобное приближение довольно условно. В связи с этим решение задачи Герца с более точными геометрическими характеристиками соприкасающихся поверхностей (в отличие от рассмотренного случая) лишено смысла, так как условие 1 раздела задачи Герца само по себе является весьма грубым приближением.

### Разрушение материала при сканировании

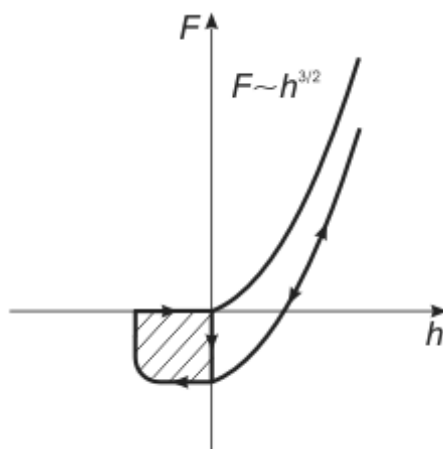
Оценив давление в месте контакта по формуле (2), нетрудно установить, разрушение каких материалов может происходить при сканировании. Достаточно сравнить предел прочности материалов (который измеряется в Па) и возникающее напряжение (давление  $P$ ). Впрочем, если в статическом режиме или при медленном сканировании нагрузка превышает критическую, то при большой скорости перемещения кантилевера разрушение может не происходить, даже если превышен порог прочности. Дело в том, что давление сверх критического должно воздействовать дольше, чем продолжается процесс разрушения (время релаксации упругих деформаций составляет примерно  $10^{-6}$  сек). При достаточно быстром сканировании больших площадей это условие может и не выполняться. Но эффект может неожиданно проявиться при изменении размеров кадра, когда его уменьшение приводит к разрушению образца.

Кроме того, при одинаковых условиях сканирования одним и тем же зондом может происходить разрушение твердого образца, в то время как более мягкий не разрушится. Это, как отмечалось, является следствием того, что для очень твердых веществ размеры контактной площадки малы, и возникает гораздо большее давление по сравнению с более мягкими образцами.

### Поправки к решению задачи Герца Неупругие консервативные силы при контакте

Помимо упругости при контакте зонда с образцом возникают и другие силы. Например, взаимодействие Ван-дер-Ваальса (которое

действует не только в соприкосновении, но и на некотором расстоянии) приводит к отрицательной добавке к давлению, так как силы Ван-дер-Ваальса, в отличие от сил упругости, являются притягивающими, а не отталкивающими.



**Рис. 8.16.** Зависимость силы  $F$  от глубины проникновения  $h$  (график решения задачи Герца, а также решение с петлей гистерезиса, учитывающее неконсервативные силы)

Это, наряду с другими притягивающими микроскопическими взаимодействиями, приводит к смещению вниз графика (рис. 8.15) — решения задачи Герца. Видно, что при  $h=0$  сила отрицательна. Это означает, что при незначительном касании зондом образца действует притягивающая сила.

### Неконсервативные эффекты

Кроме сил упругости и Ван-дер-Ваальса, существуют еще неконсервативные силы: от трения и капиллярной силы до рассеивания энергии возникающими упругими волнами — фононами. Их наличие приводит к еще большему искажению решения задачи Герца.

Благодаря неконсервативным силам, в частности, возникает адгезия (или прилипание) зонда к поверхности. При этом прикосновение и отрыв происходят по-разному, т.е. в системе появляется гистерезис.

Зонд, прилипший к поверхности, при движении вверх увлекает за собой небольшой «приклеившийся» участок образца, который в течение некоторого времени, пока не произойдет отрыв, поднимается, образуя «шейку» (см. рис. 8.6).

Таким деформациям можно сопоставить отрицательные значения глубины проникновения  $h$ . Это означает, что при обратном движении кантилевера вверх рассмотренный график может уйти левее вертикальной оси, пока не произойдет скачкообразный отрыв. На рис. 8.16 стрелками показана траектория в координатах  $F — h$  при перемещении зонда вниз и вверх.

Наличие петли гистерезиса на графике говорит о том, что чтобы прижать зонд к поверхности образца и затем его оторвать, вернув в первоначальное положение, необходимо затратить работу.

### **8.3. Эффекты, возникающие при исследовании приповерхностных свойств**

При исследовании приповерхностных свойств нанобъектов, гетероструктур и нанобластей с различным химическим и фазовым составом взаимодействие зонда микроскопа с исследуемым образцом приводит к появлению различных эффектов, которые могут также оказывать силовое воздействие на зонд, искажая тем самым получаемые результаты при исследовании какого-либо свойства приповерхностной области исследуемого образца.

#### **8.3.1. Эффект прилипания-скольжения**

Эффект прилипания-скольжения (stick-slip effect) является ключевым для контактной моды АСМ.

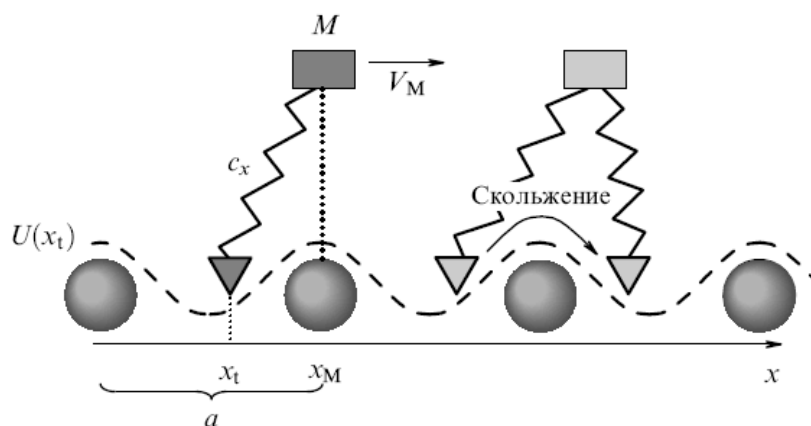
Упрощенная, но наглядная картина эффекта ПС выглядит следующим образом. Первоначально зонд находится в точке минимума потенциальной энергии системы зонд-поверхность. Это взаимодействие характеризуется периодическим потенциалом, отражающим трансляционную симметрию атомарной структуры поверхности. Существенной особенностью модели является предположение о применимости адиабатических условий на каждом шаге движения зонда. При латеральном нагружении контакта, вызванном сканированием кантилевера, энергия запасается в форме упругой энергии контакта, кантилевера и образца. Относительное скольжение зонда и поверхности начинается в тот момент, когда запасенная энергия достаточно велика для того, чтобы зонд мог "выскочить" из потенциальной ямы и зафиксироваться в другой точке поверхности. Затем система релаксирует, при этом "лишняя" энергия быстро диссипирует из области контакта через электрон-фононную подсистему. Время диссипации очень мало, поскольку характерные скорости электронов и фононов на много порядков превосходят типичные скорости сканирования зонда АСМ ( $10^{-7}$ – $10^{-4}$  мс<sup>-1</sup>).

Были проведены сравнительные исследования эффекта ПС для различных сочетаний материалов зонда и образца. В экспериментах наблюдавшаяся периодичность скольжения зонда по поверхности соответствовала топографии атомного рельефа поверхности, полученной в нормальной моде АСМ, но положения максимумов латеральных и нормальных сил были слегка смещены относительно друг друга.



## Механическая модель и результаты исследований

Механическая модель латерального движения зонда по поверхности основывается на ранней работе Томлинсона и приближении независимого осциллятора. Схема, иллюстрирующая ее, показана на рис. 8.17.



**Рис. 8.17.** Модель эффекта прилипания-скольжения (Слева: простая модель скольжения зонда по атомно-гладкой поверхности. Справа: схема движения зонда в потенциале поверхности)

Зонд упруго прикрепляется к телу массы  $M$ , связанному с кантилевером пружиной жесткостью  $c$ , и взаимодействует с образцом через периодический потенциал  $U(x_t)$ , где  $x$  – латеральная координата зонда. Движение кантилевера осуществляется с постоянной скоростью  $V_M$ . Диссипация энергии в единицу времени, независимо от микроскопических механизмов, определяется коэффициентом затухания и пропорциональна скорости, а сила трения от скорости не зависит.

Эффективные массы и жесткости пружин удовлетворяют условию критического демпфирования:  $\gamma = 2\sqrt{cm}$ . Без этого предположения регулярность эффекта ПС с периодом решетки нарушается, и наблюдается нерегулярное скольжение с кратными "прыжками".

В целом, наиболее слабым пунктом данной теории является несоответствие модели точечного осциллятора действительной ситуации с зондом АСМ, когда зона контакта имеет размеры порядка десятков нанометров и, следовательно, включает большое число атомов. В этом случае, как показывают расчеты, при различных положениях аперкса алмазного зонда на поверхности графита изменение потенциальной энергии системы после проведения релаксации структуры не превышает 0,2 эВ, поэтому нельзя говорить о том, что скачки зонда идут от одного положения минимума энергии до другого. Экспериментальный факт же заключается в том, что период, связанный с эффектом ПС, совпадает с периодом поверхностной атомной структуры. Таким образом, в режиме контактной латеральной моды АСМ следует говорить не об "истинном атомном разрешении" поверхности, как это имеет место в мо-

дуляционном режиме, а об "атомарном контрасте". Об этом свидетельствует и факт отсутствия разрешения точечных атомных дефектов поверхности (в контактной моде).

Используя синусоидальное приближение для периодической силы зонд-поверхность, уравнение движения зонда в одномерном случае можно записать в виде:

$$m \frac{d^2 x}{dt^2} = c(x_M - x) - U_0 \frac{2\pi}{a} \sin\left(\frac{2\pi}{a} x\right) - \gamma \frac{dx}{dt} , \quad (3)$$

где  $m$  — эффективная масса осциллятора,  $x_M = V_M t$  — равновесное положение недеформированной пружины в момент времени  $t$ ,  $a$  — период решетки,  $\gamma$  — постоянная затухания. Решая уравнение (3) относительно  $x(t)$ , величину латеральной силы, действующей на зонд, можно найти из уравнения  $F = c(x_M - x)$ . Диссипативная сила трения определяется путем усреднения по времени величины латеральной силы  $F_d = \langle F \rangle$ . Если  $V_M = 0$  и зонд находится в положении устойчивого равновесия, то критическим условием начала скачкообразного движения является малость коэффициента жесткости кантилевера по сравнению с латеральной жесткостью контакта:

$$c < - \frac{d^2 U}{dx^2} = \frac{4\pi^2 U_0}{a^2} .$$

В этой модели трение появляется только в результате движения и не зависит от скорости. Для больших скоростей доминирует "вязкое" затухание, и сила трения пропорциональна скорости скольжения,  $F_d \approx \gamma V_M$ .

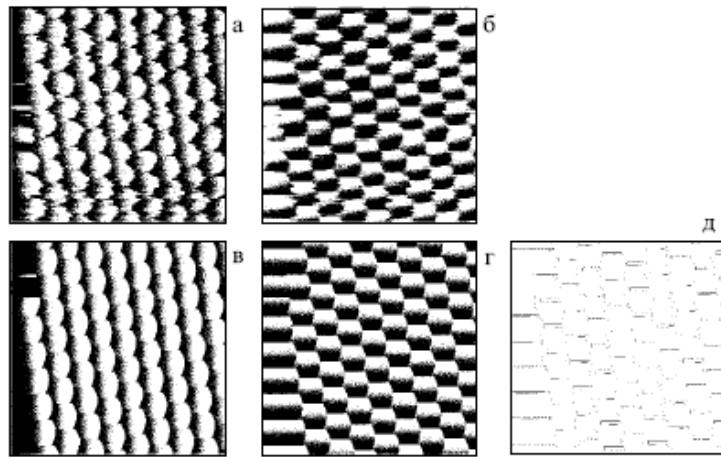
В случае двумерного движения зонда уравнение (3) обобщается путем введения двух эффективных масс, коэффициентов жесткости и постоянных затухания, соответствующих направлениям  $x$ ,  $y$ .

Для высокоориентированного пиролитического графита (ВОПГ), являющегося тестовым материалом АСМ, потенциал взаимодействия зонда с поверхностью аппроксимируется модельным выражением:

$$U_{НОРГ}(x, y) = -U_0 \left[ 2 \cos\left(\frac{2\pi}{a} x\right) \cos\left(\frac{2\pi}{a\sqrt{3}} y\right) + \cos\left(\frac{4\pi}{a\sqrt{3}} y\right) \right] ,$$

где  $a = 0,246$  нм,  $U_0 = 0,5$  эВ.

На рис. 8.18 показаны экспериментальные и полученные в результате моделирования распределения латеральных сил, действующих на зонд АСМ на поверхности ВОПГ.



**Рис 8.18.** Сравнение экспериментальных и полученных при моделировании силовых распределений

На рис. 8.18 кадры (а) и (б) соответствуют экспериментальным данным латеральных распределений сил в направлении сканирования  $F_x$  (а) и в перпендикулярном ему  $F_y$  (б), полученным с АФМ на поверхности свежего скола ВОПГ (размер кадра  $2 \times 2$  нм). Теоретические силовые распределения, соответствующие рассмотренной модели, приведены для  $F_x$  (в) и  $F_y$  (г). Сравнение рисунков показывает, что моделирование воспроизводит даже нерегулярности структуры в момент начала скольжения (левая часть изображений). Расстояние между отдельными максимумами (0,246 нм) совпадает как на экспериментальных, так и на теоретических изображениях. Направление сканирования составляет угол  $7^\circ$  по отношению к оси  $[1210]$  плоскости  $(0001)$  графита. Максимумы теоретических силовых распределений показаны точками для проекции  $F_x$  и сплошными линиями для  $F_y$ . Виден сдвиг между ними.

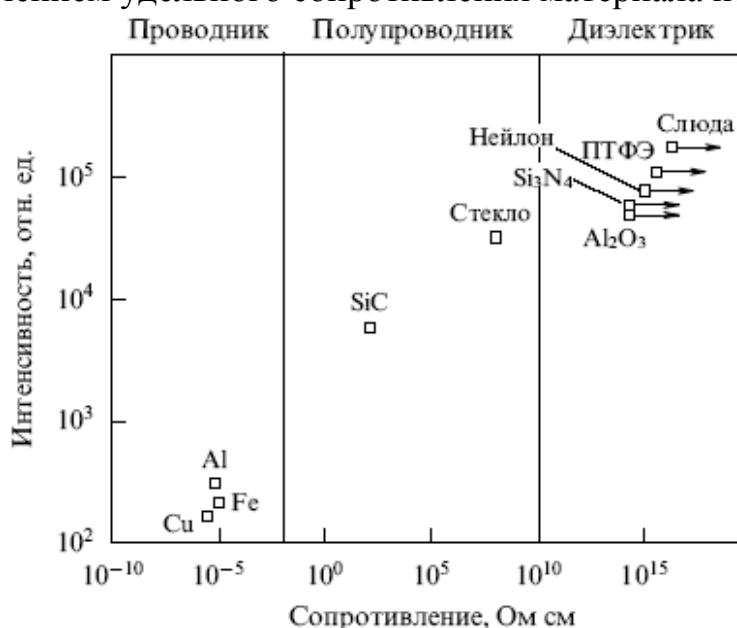
Из расчетов следует, что диссипативные силы трения не зависят от скорости скольжения при  $V < 10^{-2} \text{ см} \cdot \text{с}^{-1}$ . При больших скоростях сканирования начинает преобладать вязкое трение. Это согласуется с экспериментальными измерениями сил трения в алмазе, графите и на пленках аморфного углерода. Но экспериментальный интервал исследованных скоростей составлял  $2 \times 10^{-4} < V < 2,5 \times 10^{-3} \text{ см} \cdot \text{с}^{-1}$ , который в дальнейших исследованиях целесообразно увеличить.

Из экспериментов и модельных расчетов следует, что для наблюдения нестабильности, связанной с эффектом ПС, необходимо сочетание "мягкого" кантилевера с "жесткой" поверхностью при условии сильного взаимодействия с последней, причем больше энергии диссипирует в случае более мягких контактов. В моделях, однако, не учитываются возможные механизмы диссипации энергии, в которых сила трения пропорциональна скорости.

### 8.3.2. Трибоэмиссия частиц, электромагнитных и акустических волн

При трении в зоне контакта зонда с исследуемым образцом возникают трибоэлектромагнитные явления, которые могут сильно исказить результаты исследований различных параметров в АСМ.

Для различных контактов скольжения, используемых в устройствах магнитной записи, проводились исследования, в которых измерялся выход заряженных частиц и фотонов в процессе скольжения алмазного зонда с радиусом кривизны острия 10 мкм, при силе нормальной нагрузки 0,5 нН и скорости скольжения 2-7 см·с<sup>-1</sup> (в атмосферных условиях). Результаты этих экспериментов выявили корреляцию, характеризующую увеличение выхода частиц с увеличением удельного сопротивления материала поверхности.



**Рис. 8.19.** Корреляция между интенсивностью заряженных частиц, эмиттированных из зоны контакта зонд-поверхность, и проводимостью твердых тел. (Стрелки показывают, что значения сопротивления выходят за пределы чувствительности прибора)

Можно предположить, что образующиеся при трении зонда заряженные частицы оказываются способны покинуть поверхность и могут быть зарегистрированы лишь в том случае, если время разряда возникающих электрических полей достаточно велико. Для металлов это время примерно равно 10<sup>-15</sup> с, поэтому даже быстрые электроны, имеющие скорости порядка скорости Ферми (2·10<sup>8</sup> см·с<sup>-1</sup>) не имеют времени для того, чтобы покинуть поверхность. Следовательно, наблюдаемый в случае металлов выход положительных и отрицательных частиц, по всей видимости, связан с химической реактивностью поврежденной зондом поверхности, взаимодействующей с окружающими молекулами воздуха.

Для полупроводников и диэлектриков ситуация противоположная, и значительная часть частиц может уйти с поверхности в процессе полевой десорбции или испарения. Эти соображения дают качественное объяснение наблюдаемым зависимостям выхода частиц в случае проводников и диэлектриков (рис. 8.19)

Очевидно, необходимо проведение более детальных экспериментов такого рода для определения специфики соответствующих механизмов эмиссии в вакуумных условиях.

Эффекты нанотрибоэмиссии электронов, очевидно, имеют много общего с экзоэлектронной эмиссией, наблюдаемой в различных условиях, измерение которой является одним из перспективных методов изучения поверхностей.

Также при скольжении магнитных записывающих головок с промежуточным типом контакта, наблюдалась акустическая эмиссия. Такая эмиссия служит сигналом о контакте зонда с поверхностью при образовании царапин.

В целом, быстропротекающие процессы, происходящие в зоне контакта поверхностей, могут генерировать излучение разных типов, поэтому соответствующие эффекты целесообразно использовать для изучения нанотрибологических воздействий на поверхность.

### 8.3.3. Трибохимические эффекты

Трение может стимулировать и подавлять химические процессы, а последние, в свою очередь, влияют на трение, чувствительное к химическому составу зоны контакта. Зависимость нормальных и латеральных сил может быть, в частности, использована для получения соответствующих изображений и исследования химической реактивности поверхности.

Измеренные зависимости сил трения от силы нагрузки для поверхности KCl приведены на рис. 8.20, и качественно резко отличаются от аналогичных зависимостей в случае упругих контактов.

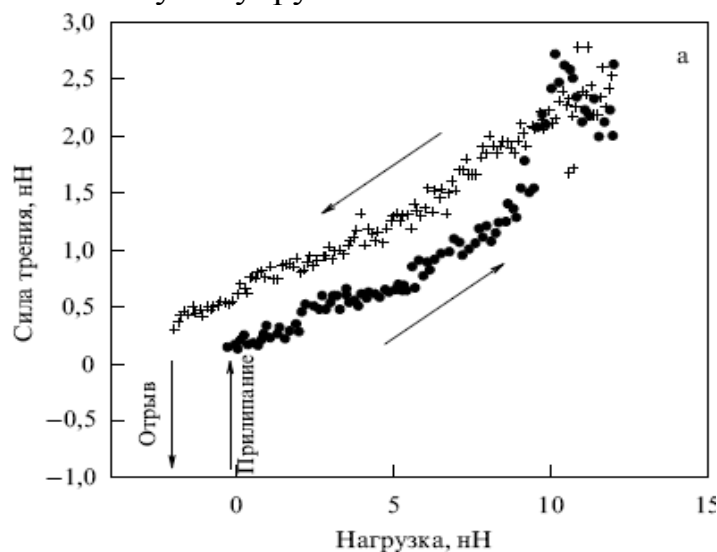


Рис. 8.20. Зависимость сила трения-сила нагрузки для KCl

Из экспериментальных данных видно, что сила нагрузки изменяется в относительно широких пределах. Сила трения плавно возрастает до тех пор, пока сила нагрузки не становится равной 5,5 нН, после чего наблюдается более резкий подъем. При силах нагрузки, превышающих 9 нН, наблюдаются нерегулярные флуктуации силы трения, причем при уменьшении нагрузки величина силы трения обнаруживает гистерезис и не возвращается к первоначальным значениям.

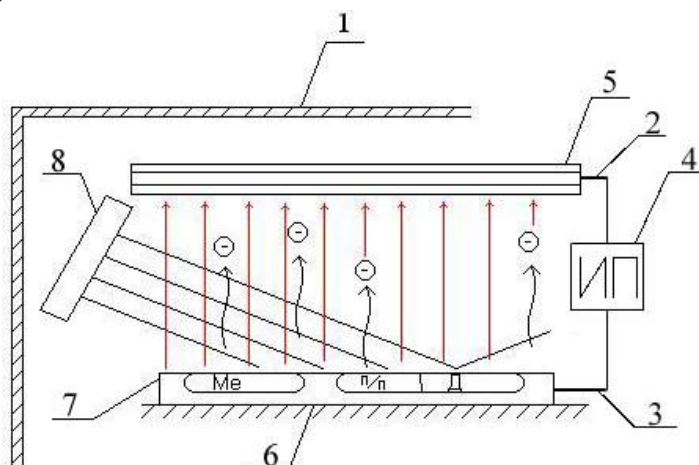
Для интерпретации данных кривых можно сделать предположение о структурной модификации контактной зоны. Аналогичные зависимости наблюдаются также на пленках золота, нитрида кремния и органических материалов. Возможно, именно такие (линейные) зависимости специфичны для трибохимического износа.

#### 8.4. Анализ химического состава поверхности

При проведении исследований приповерхностных свойств химический состав поверхности исследуемых образцов и распределение поверхностных потенциалов, учитывается косвенно или не учитывается совсем. А ведь именно эти свойства поверхности могут существенно влиять на получаемые экспериментальные данные и приводить к расхождениям между значениями параметров, предсказанных теоретически, и результатами эксперимента.

Для получения информации о химическом составе поверхности исследуемого образца а также о распределении поверхностных электрических потенциалов предлагается использовать устройство определения химического состава поверхности подложки.

В основу разработки положена задача обеспечить возможность определения химического состава на поверхности материала туннельно-зондовым (неконтактным) методом.



**Рис. 8.21.** Схема устройства регистрации химического состава подложки

Устройство для регистрации химического состава поверхности подложек (рис. 8.21) содержит вакуумную камеру 1, внутри которой расположены анод 2 и катод 3, связанные с источником питания 4. Анод 2 выполнен в виде люминесцентного экрана 5, а катод 3 в виде подложкодержателя 6, на котором установлена подложка 7, внутри вакуумной камеры 1 под углом к подложке 7 расположен источник электромагнитного излучения 8, с возможностью подачи квантов излучения на поверхность подложки 7 и отражения от неё без взаимодействия с люминесцентным экраном 5, который расположен параллельно подложке 7.

Устройство работает следующим образом.

При подаче электромагнитного излучения на подложку 7 кванты излучения начинают выбивать электроны с её поверхности, которые под действием электрического поля, создаваемого между анодом 2 и катодом 3, начинают «бомбить» люминесцентный экран 5. В результате того, что подложка 7 легирована, электроны, вылетая из разных областей и осаждаясь на люминесцентном экране 5, образуют на нём «картину», по которой можно определить химический состав и области легирования материала подложки 7.

Применение устройства для регистрации химического состава позволяет определять химический состав и области легирования материала подложки туннельно-зондовым (неконтактным) методом, что в свою очередь означает возможность применения данного устройства для исследования образцов, обладающих малой механической жесткостью (органические структуры и биологические объекты).

### **Заключение**

Рассмотренные выше свойства приповерхностных областей твердых тел являются лишь частью всех свойств, которыми обладают данные области. Также ясно, что для более полного описания этих свойств необходимо проводить дальнейшие исследования в данной области с использованием различных методов исследования, таких как сканирующая туннельная микроскопия, ближнепольная оптическая микроскопия, электронная спектроскопия и т.д.

Задача исследования приповерхностных свойств твердых тел является первостепенной при моделировании любых устройств, принцип действия которых основан на данных свойствах. Поэтому для составления полной картины о свойствах приповерхностных областей необходимо объединить исследования в различных направлениях, что исключит возможность повторений, ускорит процесс накопления знаний о данных свойствах, и в итоге приведет к ускорению темпов развития технологий.

## Литература

1. Binning G, Quate C F, Gerber C Phys. Rev. Lett. 56 930 (1986)
2. Mate C M et al. Phys. Rev. Lett. 69 1942 (1987)
3. Israelachvili J.N. Intermolecular and Surface Forces. –Academic Press, 1998. – 450с.
4. Зимон А. Д. Адгезия жидкости и смачивание. –М.: Химия, 1974. - 416с.
5. Derjaguin B. V., Muller V.M., Toporov Yu.P., J.Colloid. Interface Sci. 53, 314 (1975)
6. Дерягин Б.В., Чураев Н.В., Муллер В.М. Поверхностные силы. –М.: Наука, 1985.
7. Джонсон К.Л. Механика контактного взаимодействия. –М.: Мир, 1987.
8. Maugis D.J., J.Colloid. Interface Sci. 150, 243 (1992)
9. Tomlinson I Philos. Mag. 7 905 (1929)
10. Дедков Г. В., Дышеков М. Б. ЖТФ 70 (7) 96 (2000).
11. Ландау Л. Д., Лившиц Е. М. Теория упругости. –М.: Наука, 1987.–246с.
12. Головин Ю.И. А.И. Тюрин, В.И. Иволгин, В.В. Коренков. ЖТФ 70, 5 82 (2000)
13. М.О. Галлямов "Сканирующая зондовая микроскопия нуклеиновых кислот и тонких органических пленок". Диссертация на соискание ученой степени кандидата физико-математических наук. 1999г.
14. В.Л. Миронов "Основы сканирующей зондовой микроскопии". Учебное пособие для студентов старших курсов высших учебных заведений. 2004г.
15. Ивашов Е.Н., Степанчиков С.В., Мисюк Р.В. Приповерхностные свойства нанобъектов, гетероструктур и нанобластей с различным химическим и фазовым составом. М.: МИЭМ, 2006, 35с

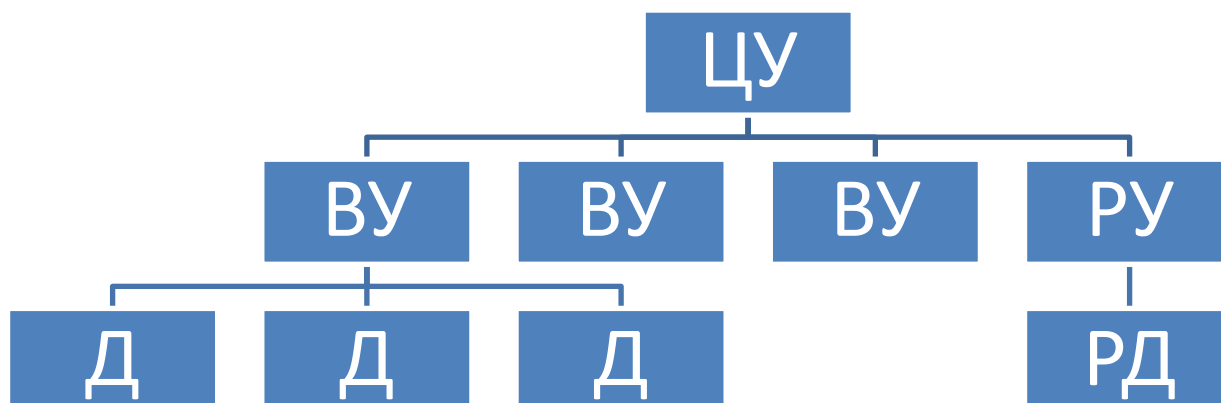


## Глава 9. Управление качеством распознавания образов в кластерных системах обработки информации

### Введение

В разделе рассматривается математическая модель функционирования кластерных систем обработки информации, алгоритмы и методы моделирования состояния подобных систем во времени, а так же метод определения порога функционирования системы с использованием эволюционной стратегии (генетического алгоритма).

Кластер функционирует как единая система, то есть для пользователя или прикладной задачи вся совокупность вычислительной техники выглядит как один компьютер. Именно это и является самым важным при построении кластерной системы. В настоящее время кластерные системы обработки информации получают все большее распространение в связи с удешевлением их компонентов и, как следствие, остро встает вопрос об управлении качеством функционирования подобных систем.



*Рис. 9.1. Древовидная структура кластерной системы*

К общим требованиям, предъявляемым к кластерным системам, относятся:

1. Высокая готовность.
2. Высокое быстродействие.
3. Масштабируемость.

## 9.1. Теоретический анализ кластерных систем

Кластерная система обработки информации описывается:

- Множеством состояний  $\Theta[a_1, \dots, a_n]$ , где  $a_i \in A_i$  - состояние отдельного элемента системы,  $A$  - дискретное множество состояний, которые может принимать  $i$ -й элемент системы;
- Целевой функцией  $F(\Theta_j) = [F_1, \dots, F_m]$  где  $\Theta_j \in \Theta$  - некоторое состояние системы,  $[F_1, \dots, F_m]$  - вектор целевых показателей, характеризующий систему в целом;
- Архитектурой кластерной системы.

В общем случае кластерную систему можно представить в виде графа, узлы которого представляют собой устройства сбора и обработки информации, а ветви - каналы передачи данных. Наиболее часто такой граф имеет древовидную структуру, представленную на рис. 9.1, где  $D_1, \dots, D_n$  - датчики (устройства ввода информации),  $PD$  - резервные датчики,  $BV_1, \dots, BV_n$  - вычислительные узлы (устройства обработки информации),  $PV$  - резервные узлы,  $ЦУ$  - центральный узел.

Архитектура системы находит отражение в целевой функции, т.к. целевые показатели вышестоящих элементов напрямую зависят от показателей нижестоящих элементов. Исходя из этого, состояние и целевые показатели  $i$ -го элемента будут являться функциями от состояний и показателей нижестоящих элементов:

$a_i = \Theta[b_1, \dots, b_n]$ , где  $a_i$  - состояние  $i$ -го элемента,  $\Theta[b_1, \dots, b_n]$  - функция перехода состояний, учитывающая весовой коэффициент  $j$ -го элемента, показывающий его важность для функционирования элемента верхнего уровня и системы в целом:

$$\Theta[b_1, \dots, b_n] = \frac{\sum_{j=1}^n k_j b_j}{n}, \quad \text{где } b_j \in B \quad - \text{ состояние } j\text{-го элемента нижнего}$$

уровня,  $k_j$  - весовой коэффициент элемента.

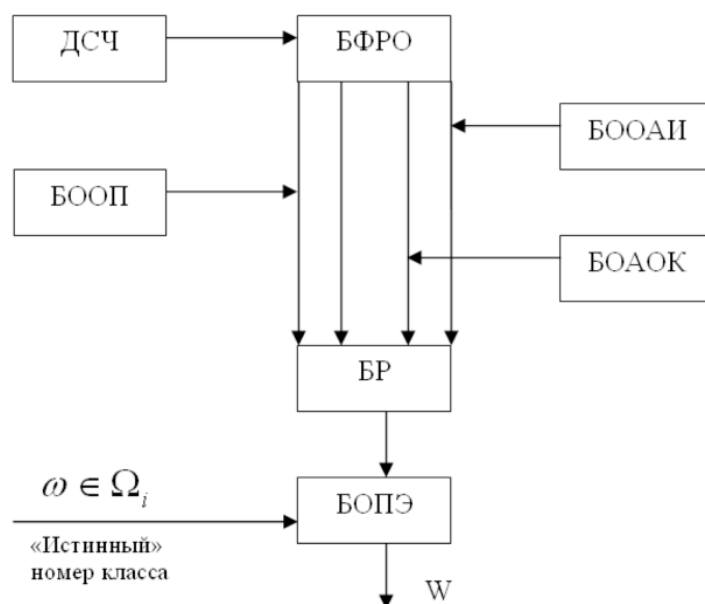
Целевая функция состояния элемента  $a_i$  имеет вид:

$f(a_i) = \sum_{j=1}^n k_j f(b_j)$ , где  $f(b_j)$  целевая функция состояния для элементов нижнего уровня:  $f(b_j) = [F_1, \dots, F_n]$

Рассмотрим целевые показатели обработки информации для кластерных информационных систем. Кластерную систему можно охарактеризовать следующими показателями:

- Вероятность ложного срабатывания  $W$  - это результирующая вероятность программной ошибки в каждом физическом элементе системы.
- Коэффициент доступности (или работоспособности) системы  $P$  - обусловлен вероятностью полной недоступности системы в связи с аппаратными или программными неполадками. Этот коэффициент выражает количественную меру работоспособности системы.

- Производительность системы – обусловлена временем, которое затрачивается системой на решение эталонного задания.
- Время отклика – это время, необходимое системе на обработку команды оператора или восприятие новой задачи.



*Рис. 9.2. Принципиальная схема системы обработки информации*

Рассмотрим подробнее каждый целевой показатель.

Чтобы оценить эффективность вероятностных систем обработки информации на основе математического моделирования, можно использовать метод статистических испытаний. Для проведения таких испытаний может служить математическая модель функционирования системы, принципиальная схема которой представлена на рис. 9.2, где БФРО – блок формирования распознаваемых объектов, БООП – блок ошибок определения признаков, БОАОК – блок ошибок априорного описания классов, БООАИ – блок ограничения объема апостериорной информации, БР – блок распознавания, БОПЭ – блок оценки показателя эффективности, ДСЧ – датчик случайных чисел.

Принцип действия модели следующий. Для проведения каждого испытания с помощью ДСЧ формируется модель объекта, принадлежность которого к определенному классу заранее известна. Формирование модели объекта производится заданием совокупности числовых значений признаков  $x_1, \dots, x_N$ , которые для объектов из класса  $\Omega_i$  генерируются как реализации многомерной случайной величины с заданным законом распределения  $f_i(x_1, \dots, x_N)$  по одному из известных алгоритмов.

Числовые значения параметров  $x_1, \dots, x_N$ , представляющие собой обрабатываемый объект, подвергаются случайному искажению, что имитирует результат воздействия различных помех в процессе определения признаков  $x_1, \dots, x_N$  при использовании соответствующих технических средств

с определенными точностными характеристиками. Искаженные значения параметров  $x'_1, \dots, x'_N$ , представляющие наблюдаемый объект в том виде, в каком его воспринимает система, поступают на вход БР, в котором определяется принадлежность объекта одному из классов  $\Omega_1, \dots, \Omega_N$ . Блок БОПЭ сопоставляет номер класса, к которому отнесен объект блоком распознавания, с «истинным» номером, т.е. с тем, который задавался на первом этапе формирования объекта, определяет правильность обработки информации и систематизирует соответствующую информацию для подсчета оценок вероятностей верных и ошибочных решений. При обработке объектов из класса  $\Omega_i$  оценкой  $p_i$  вероятности получения правильного решения служит отношение количества правильных ответов  $N_{пр}^i$  к общему числу испытаний  $N^i$  над объектами класса  $\Omega_i$ , т.е.  $p_i \approx N_{пр}^i / N^i$ . Число испытаний  $N^i$  определяется доверительной вероятностью, задаваемой при формулировке задачи исследования.

В зависимости от задачи исследования искажению могут подвергаться также априорные данные о классах объектов, т.е. функции распределения  $f_i(x_1, \dots, x_N)$  и  $P(\Omega_i)$ , информация о признаках  $x_1, \dots, x_N$  может урезаться, что соответствует отсутствию некоторых средств определения признаков, и т.п.

Если априорные вероятности  $P(\Omega_i)$  появления объектов из разных классов известны, то безусловная вероятность правильного решения задачи обработки информации данной системой может быть выбрана в качестве критерия эффективности системы обработки информации:

$$W = \sum_{i=1}^n p_i P(\Omega_i).$$

Рассмотренная статистическая модель позволяет найти зависимость  $W$  от вида и количества привлекаемых для обработки признаков  $x_1, \dots, x_N$  и точности  $\sigma_1, \dots, \sigma_s$  технических средств, которыми оснащается система обработки информации, т.е.  $W = W(x_1, \dots, x_n; \sigma_1, \dots, \sigma_s)$ .

Сведения, содержащиеся в этом равенстве, - исходные для задач об определении состава технических средств наблюдений системы обработки информации, необходимых точностей их работы, об оптимальном с точки зрения экономических соображений распределении точностей по средствам и т.д.

Перейдем к рассмотрению следующего целевого показателя. Коэффициент доступности кластерной системы можно определить с помощью метода соотношений. Суть метода сводится к определению вероятности безотказного функционирования сложной многоуровневой кластерной системы (КС).

Процесс функционирования кластерной системы организован таким образом, что система успешно решает свои задачи при условии, если в исправном состоянии находится хотя бы одно устройство ввода информации, все устройства обработки информации и центральный узел. Данное условие

выполнения целевой функции системы можно наглядно представить в форме логической функции:

$$\begin{aligned}
 F(\text{КС}) &= [F_1(D)] \wedge [F_2(\text{ВУ})] \wedge F_3(\text{ЦУ}), \\
 F_1(D) &= F_1(D_1 \vee \dots \vee D_n), \\
 F_2(\text{ВУ}) &= F_2(\text{ВУ}_1 \wedge \dots \wedge \text{ВУ}_n).
 \end{aligned}
 \tag{1}$$

Эти выражения означают, что устройство, указанное в скобках, работает исправно.

Представляет интерес также логическая зависимость, описывающая условия не выполнения системой своих целевых функций:

$$\begin{aligned}
 \neg F(\text{КС}) &= \neg[F_1(D)] \vee \neg[F_2(\text{ВУ})] \vee \neg F_3(\text{ЦУ}); \\
 F_1(D) &= F_1(D_1 \wedge \dots \wedge D_n); \\
 F_2(\text{ВУ}) &= F_2(\text{ВУ}_1 \wedge \dots \wedge \text{ВУ}_n).
 \end{aligned}
 \tag{2}$$

Последнее выражение может оказаться более удобным для решения поставленной задачи определения коэффициента доступности с учетом того, что:

$$P[F(*)] = 1 - P[\neg F(*)],$$

где  $P[F(*)]$  - вероятность истинности условия  $F(*)$ , а  $P[\neg F(*)]$  - вероятность истинности отрицания истинности данного условия. Перечисленные элементы КС имеют различное функциональное назначение и соединены так, что надежность каждого из них оказывает непосредственное влияние на работоспособность всей системы в целом. Поэтому в качестве факторов для оценки надежности функционирования КС следует взять вероятности  $P_i$  безотказного функционирования устройств в процессе решения системой поставленных задач. В общем случае вероятности  $P_i$  могут иметь различные значения. Вероятность  $P[F(\text{КС})]$  безотказного функционирования КС в целом есть функция от вероятностей безотказного функционирования всех ее элементов, вытекающая из рассмотренных выше логических условий.



**Рис. 9.3.** Обобщенная схема математической модели, характеризующей безотказность функционирования КС

Обобщенная схема математической модели, характеризующей безотказность функционирования КС, имеет вид, показанный на рис. 9.3.

Проблема состоит в том, как из логических условий получить соответствующее выражение для количественного значения вероятности  $P[F(КС)]$ . Дело в том, что вероятность  $P[F(КС)]$  определяется на множестве состояний  $\Theta(t)$ . Число состояний в данном множестве равно  $2^n = N$ , где  $n$  – число структурных элементов КС. Условие функционирования (1) определяет подмножество состояний системы, обеспечивающих выполнение системой заданных целевых функций, а условие (2) определяет подмножество состояний, в которых система оказывается неработоспособной.

Очевидно, для решения данной задачи таким способом придется осуществить полный перебор всех  $N$  состояний системы, или же придумать более эффективный способ определения работоспособных состояний, особенно если учесть, что число состояний системы находится в степенной зависимости от числа ее элементов. Наиболее перспективными методами решения этой проблемы представляются метод имитационного моделирования и формализованный переход от логических функций к соответствующим формулам вероятностей сложных событий.

Рассмотрим производительность систем обработки информации. Общая производительность кластерной системы обработки информации обусловлена производительностью каждого вычислительного элемента системы и определяется экспериментальным путем. Для этого каждому из элементов вычислительной системы дается эталонное задание и определяется время, затраченное на его решение. Исходя из затраченного на решение задачи времени узлам назначаются весовые коэффициенты, характеризующие производительность вычислительного узла и системы в целом.

В реальных условиях производительность кластерной системы обработки данных зависит не только от производительности вычислительных узлов, но и от надежности и пропускной способности каналов передачи данных. Таким образом, ко времени, затрачиваемому на решение эталонного задания всей системой, добавляется время, необходимое системе на подтверждение принятия задания, и время, затрачиваемое системой на передачу данных между вычислительными узлами к центру.

Производительность систем в реальных условиях можно вычислить следующим образом:

$$П = \sum_{i=1}^M w_i + 2L ,$$

где  $П$  – производительность системы,  $w_i$  – весовой коэффициент производительности вычислительного узла,  $N$  – общее количество элементов системы,  $M$  – количество вычислительных узлов.  $L$  – время прохождения сигнала по каналам связи, определяемое по формуле:

$$L = \min \left\{ \sum_{j=1}^{N-M} w_j \right\} ,$$

где  $w_j$  – весовой коэффициент пропускной способности канала связи.

Одним из важных параметров, описывающих кластерную систему обработки информации, является порог функционирования, т.е. такое значение целевой функции, при переходе через которое система перестает функционировать. Для относительно простых систем обработки информации это значение может быть получено экспериментальным или эмпирическим путем. Однако, для систем с большим числом разнородных элементов это представляется затруднительным. Выходом в подобной ситуации может быть моделирование системы с использованием эволюционной стратегии, где критерием отбора будет являться наиболее функциональное состояние системы при максимальном количестве неисправностей. Эволюционные алгоритмы базируются на коллективном обучающем процессе внутри популяции индивидуумов, каждый из которых представляет собой поисковую точку в пространстве допустимых решений данной задачи. Наиболее известными из класса эволюционных алгоритмов являются генетические алгоритмы. Генетический алгоритм (рис. 9.4) может быть легко применен для безусловной оптимизации функций, т.е. для задачи отыскания значений параметров, которые минимизируют или максимизируют заданную целевую функцию и для безусловной комбинаторной оптимизации, т.е. для задачи отыскания наилучшей комбинации вариантов, которая оптимизирует заданную целевую функцию. Их основные адаптивные процессы концентрируются на идее системы, получающей сенсорную информацию от окружающей среды через бинарные детекторы. В генетических алгоритмах существует строгое различие между фенотипом (решением) и генотипом (представлением решения). Генетический алгоритм работает только с генотипом, поэтому требуется процесс декодирования генотипа в фенотип и обратно («обобщенный» рост). Вещественные параметры могут быть представлены числами с фиксированной точкой или целыми числами путем масштабирования и дискретизации. Для вещественных параметров имеет место конфликт между желанием иметь как можно более короткий ген для обеспечения хорошей сходимости и необходимостью получить результат с определенной точностью.

## 9.2. Комментарии к генетическому алгоритму

Шаг 1. Генерация начальной популяции. Случайным образом генерируется  $n$  уникальных состояний системы (индивидов), для каждого состояния вычисляется значение целевой функции и показатель работоспособности системы.

Шаг 2. Кодирование состояний системы в бинарный код (составление хромосом).

Шаг 3. Оценка пригодности каждого состояния. Для этого состояния ранжируются по значениям показателя работоспособности.

Шаг 4. Репликация состояний, т.е. генерация новой популяции: из  $m < n$  состояний попарно генерируются потомки. В нашем случае потомком будет

являться результирующее состояние, являющееся следствием событий состояний-родителей.



**Рис. 9.4.** Генетический алгоритм

Шаг 5. Оценка пригодности всех состояний, включая потомков.

Шаг 6. Селекция. Для имитации естественной селекции состояния с более высокой пригодностью должны выбираться с большей вероятностью, поэтому из получившихся состояний выбирается  $n$  самых пригодных.

Шаг 7. Проверка конечного условия: если номер поколения  $n_{\text{поко.}}$  не равен заложенному на этапе инициализации конечному числу поколений  $n_{\text{кон.}}$ , то увеличение  $n_{\text{поко.}}$  на единицу и переход на Шаг 4. Если  $n_{\text{поко.}} = n_{\text{кон.}}$ , то переход на Шаг 8.

Шаг 8. Декодирование и отображение полученного результата.



### 9.3. Кластеры повышенной производительности

Кластеры повышенной производительности обозначаются англ. аббревиатурой HPC (*High performance cluster*). Позволяют увеличить скорость расчетов, разбивая задание на параллельно выполняющиеся потоки. Используются в научных исследованиях. Одна из типичных конфигураций - набор серверов с установленной на них операционной системой Linux, такую схему принято называть кластером Beowulf. Для HPC создается специальное ПО, способное эффективно распараллеливать задачу.

Эффективные связи между серверами в кластере позволяют им поддерживать связь и оперативно обмениваться данными, поэтому такие кластеры хорошо приспособлены для выполнения процессов, использующих общие данные.

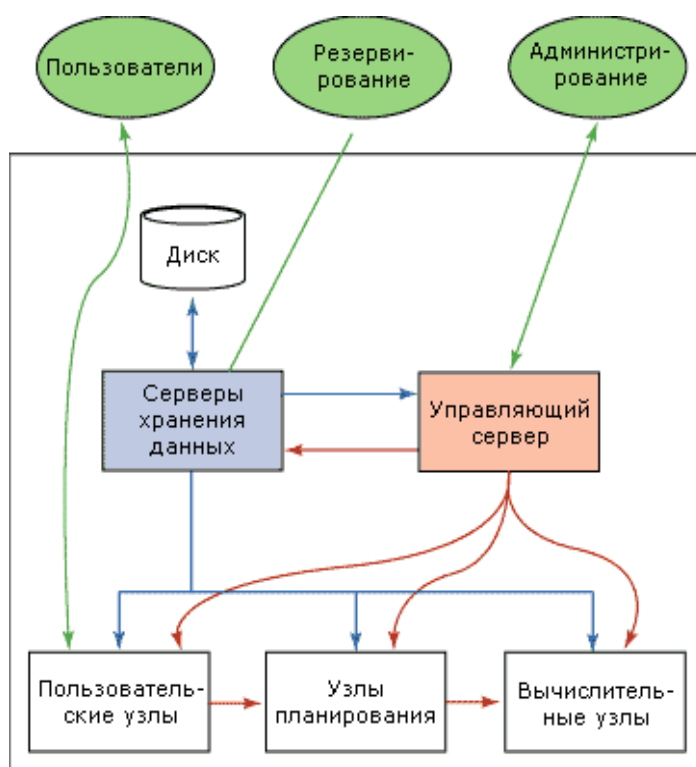


Рис. 9.5. Пример кластера

#### Физическая схема:

- План стойки для каждого их типа (например, управляющие и вычислительные стойки).
- Поэтажный план расположения стоек во время процесса установки системы и при рабочем использовании, если они отличаются.
- Схемы внутренних соединений стоек для сети, цепей питания, пульта оператора и т.д.
- Схема внешних соединений для серверов системы хранения, терминальных серверов и т.д.

### **Логическая схема:**

- Схема сети, включая диапазоны IP-адресов, конфигурацию подсетей, соглашения по наименованию компьютеров и т.д.
- CSM-конфигурация по расположению пользовательских сценариев, аппаратные настройки и требования по мониторингу.
- Требования к операционным системам, список специализированных пакетов и параметры конфигурации системы.
- Схема системы хранения данных, включая схему файловой системы, разбиение дисков, параметры репликации и т.д.

Кластер (рис. 9.5) состоит из компьютеров, работающих на процессорах Intel или AMD, с подключенными подсистемами TotalStorage.

Для простоты соединения в кластере выполнены медным кабелем стандарта гигабитный Ethernet. Этот кабель обеспечивает хорошую скорость в большинстве случаев.

Сетевая топология имеет форму звезды - все стойки подключены к основному коммутатору управляющей стойки. В примере используется три сети: одна для управления/данных (вычислительная сеть), одна для кластерной файловой системы (сеть хранения данных) и одна для администрирования устройств. Первые две сети - это обычные IP-сети. Для большинства задач, включая межпроцессные взаимодействия (например, MPI) и управление кластером, используется вычислительная сеть. Сеть хранения данных используется исключительно для доступа и взаимодействия с кластерной файловой системой.

### **Управляющий сервер**

Функция управляющего сервера может выполняться одним сервером или на нескольких. В среде с одним сервером управляющий сервер функционирует в автономном режиме. Можно настроить также управляющие серверы с высокой готовностью. Для этого можно использовать программное обеспечение CSM для поддержки высокой готовности (high-availability - HA), которое будет выдавать тактовые импульсы ("heartbeat") между двумя серверами и поддерживать динамическое восстановление после сбоев при возникновении аварийных ситуаций. Другим возможным методом организации нескольких управляющих серверов является использование репликации, если поддержка HA не важна для вашей среды. В этом случае вы можете резервировать данные управляющего сервера на другую рабочую систему, которую можете, при необходимости, перевести в оперативный режим вручную. Управляющий сервер - это CSM-сервер, использующийся исключительно для внутреннего управления кластером при помощи CSM-функций: управление установкой системы, мониторинг, обслуживание и другие задачи. В данном кластере присутствует только один управляющий сервер.

## **Серверы хранения данных и дисковые накопители**

Вы можете подключить несколько серверов хранения данных к организованному на дисковых накопителях хранилищу данных при помощи различных механизмов. Подключить систему хранения данных к серверу можно напрямую: либо через SAN-коммутатор (storage area network - сеть хранения данных) по оптическому волокну или медному кабелю, либо используя оба типа соединений (рис. 9.5). Эти серверы предоставляют совместный доступ к системе хранения данных другим серверам кластера. Если необходимо резервирование базы данных, подключите резервное устройство к серверу хранения данных, используя дополнительное медное или оптическое соединение. В примере кластера хранилище представляет собой единую сущность, обеспечивающую доступ к общей файловой системе в пределах кластера. В следующей статье данной серии будет более подробно описана установка, настройка и реализация аппаратного обеспечения системы хранения данных и кластерной файловой системы.

### **Пользовательские узлы**

В идеальном случае вычислительные узлы кластера не должны принимать внешние подключения и должны быть доступны только для системных администраторов через управляющий сервер. Пользователи системы могут зарегистрироваться на вычислительных узлах (или узлах регистрации) для выполнения своей работы в кластере. Каждый пользовательский узел состоит из образа с возможностями любого редактирования, необходимых библиотек разработчика, компиляторов и всего, что необходимо для создания кластерного приложения и получения результатов.

### **Узлы планирования**

Для запуска рабочей нагрузки на кластере пользователи должны передать свою работу узлу планирования. Фоновый процесс-планировщик (scheduler daemon), работающий на одном или нескольких узлах планирования, применяет предопределенную политику для запуска рабочих нагрузок в кластере. Аналогично вычислительным узлам, узлы планирования не должны принимать внешних подключений от пользователей. Системные администраторы должны управлять ими при помощи управляющего сервера.

### **Вычислительные узлы**

Эти узлы выполняют рабочую нагрузку кластера, принимая задания от планировщика. Вычислительные узлы - это самые свободные части кластера.

Системный администратор может легко переустанавливать или перенастраивать их при помощи управляющего сервера.

### **Ethernet-коммутаторы**

Имеется две физических сети: одна для вычислений, а вторая для хранения данных. Стандартная емкость стойки в 32 узла требует применения двух 48-портовых коммутаторов на каждую стойку, по одному на каждую сеть. В маленьких кластерах в управляющей стойке также необходимо использовать два одинаковых коммутатора. Для больших кластеров 48-портов может оказаться недостаточно, и может потребоваться более мощный центральный коммутатор.

Терминальные серверы играют важную роль в больших кластерных системах, использующих версии CSM, ниже чем CSM 1.4. Кластеры, использующие старые версии, нуждаются в терминальных серверах для сбора MAC-адресов при установке. При совместимости CSM и системных UUID, терминальные серверы становятся не так важны для установки более современного кластера. Однако, если в большом кластере у вас имеется немного устаревшее оборудование или программное обеспечение, терминальные сервера все еще остаются жизненно важными во время установки системы. Обеспечение корректной настройки терминального сервера само по себе может сэкономить значительное количество времени в дальнейшем в процессе установки системы. Кроме сбора MAC-адресов терминальные серверы могут также использоваться для просмотра терминалов из одной точки во время процедуры начального самотестирования (POST) и запуска операционной системы.

Когда компьютер получает DHCP-адрес во время PXE, конфигурационные файлы в `/tftpboot/pxelinux.cfg` ищутся в определенном порядке, и первый найденный файл используется в качестве загрузочной конфигурации для запрашивающего компьютера. Порядок поиска определяется путем преобразования запрашиваемого DHCP-адреса в шестнадцатиричные цифры и поиска первого подходящего имени файла в конфигурационном каталоге методом расширения подсетей - удаления цифр справа налево на каждом цикле поиска. [3]

#### **9.4. Коммуникационные библиотеки**

Наиболее распространенным интерфейсом параллельного программирования в модели передачи сообщений является MPI. Рекомендуемая бесплатная реализация MPI - пакет MPICH, разработанный в Аргоннской Национальной Лаборатории. Для кластеров на базе коммутатора Myrinet разработана система HPVM, куда также входит реализация Message Passing Interface (MPI, интерфейс передачи данных).

MPI – программный интерфейс (API) для передачи информации, который позволяет обмениваться сообщениями между компьютерами, выполняющими одну задачу.

MPI является наиболее распространённым стандартом интерфейса обмена данными в параллельном программировании, существуют его реализации для большого числа компьютерных платформ. Основным средством коммуникации между процессами в MPI является передача сообщений друг другу. Стандартизацией MPI занимается MPI Forum. В стандарте MPI описан интерфейс передачи сообщений, который должен поддерживаться как на платформе, так и в приложениях пользователя. В настоящее время существует большое количество бесплатных и коммерческих реализаций MPI. Существуют реализации для языков Фортран 77/90, Си и Си++.

## 9.5. Стандарты MPI

Большинство современных реализаций MPI поддерживают версию 1.1. Стандарт MPI версии 2.0 поддерживается большинством современных реализаций, однако некоторые функции могут быть реализованы не до конца.

В MPI 1.1 поддерживаются следующие функции:

- передача и получение сообщений между отдельными процессами;
- коллективные взаимодействия процессов;
- взаимодействия в группах процессов;
- реализация топологий процессов;

В MPI 2.0 дополнительно поддерживаются следующие функции:

- динамическое порождение процессов и управление процессами;
- односторонние коммуникации;
- параллельный ввод и вывод.

Для эффективной организации параллелизма внутри одной SMP-системы возможны два варианта:

1. Для каждого процессора в SMP-машине порождается отдельный MPI-процесс. MPI-процессы внутри этой системы обмениваются сообщениями через разделяемую память (необходимо настроить MPICH соответствующим образом).

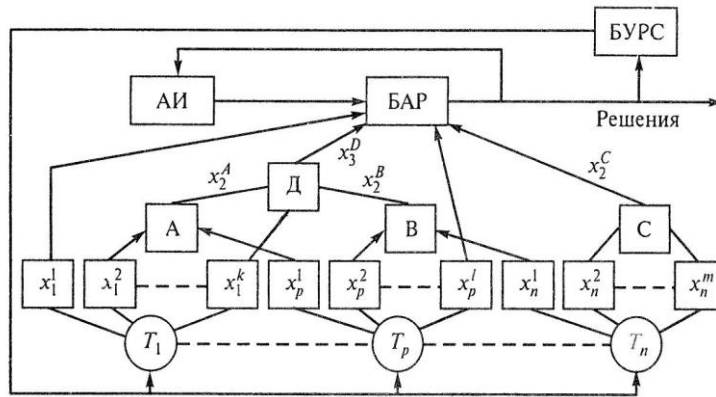
2. На каждой машине запускается только один MPI-процесс. Внутри каждого MPI-процесса производится распараллеливание в модели "общей памяти", например с помощью директив OpenMP.

После установки реализации MPI имеет смысл протестировать реальную производительность сетевых пересылок.

Кроме MPI, есть и другие библиотеки и системы параллельного программирования, которые могут быть использованы на кластерах.

Следует понимать, что использование для программирования OpenMosix — расширение (патч) ядра Linux, позволяющее создать единый кластер. Превращает сеть обычных персональных компьютеров в

суперкомпьютер для Linux-приложений. Представляет собой полнофункциональную кластерную среду с единой операционной системой (SSI), автоматически распараллеливающую задачи между однородными узлами. Это позволяет миграцию процессов между машинами — узлами сети.

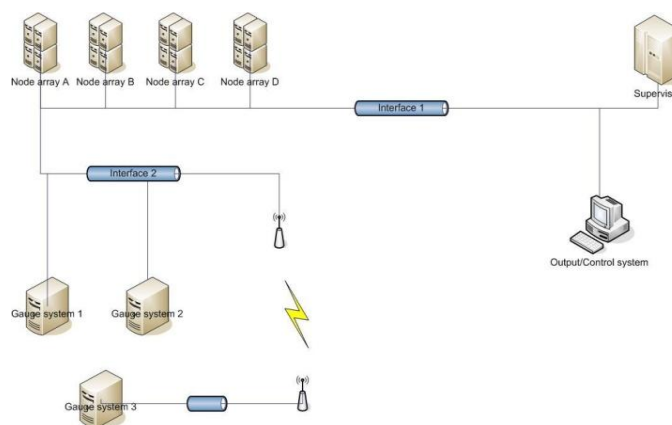


**Рис. 9.6.** Сложная многоуровневая система распознавания

Кластер ведёт себя подобно SMP-машине (за исключением любых видов разделяемой памяти). При этом возможно наращивание до тысяч узлов, которые тоже могут быть SMP-машинами. Добавление новых узлов возможно параллельно работе кластера, добавленные ресурсы будут задействованы автоматически. OpenMosix также предоставляет оптимизированную файловую систему (OMFS) для HPC-приложений, которая, в отличие от NFS, поддерживает кэширование, отметки о времени и ссылки.

OpenMosix — это проект, являющийся продолжением проекта MOSIX, но под свободной лицензией GPL. Последние релизы MOSIX стали проприетарными (закрытыми) в конце 2001 года, а проект openMosix стартовал 10 февраля 2002. Инициатор проекта — Moshe Bar.

OpenMosix поставляется с набором утилит для администрирования кластера. Для этого имеется также удобное GUI-приложение openMosixview.



**Рис. 9.7.** Пример архитектурного решения кластера

При обработке результатов сканирующей зондовой микроскопии часто встает вопрос о принадлежности объекта исследования тому или иному классу объектов. Подобные задачи решаются применением систем распознавания.

Распознавание представляет собой задачу преобразования входной информации, в качестве которой уместно рассматривать некоторые параметры, признаки распознаваемых объектов, в выходную, представляющую собой заключение о том к какому классу объектов принадлежит распознаваемый объект.

Важнейшей характеристикой принимаемых в процессе распознавания решений и основным показателем качества распознавания является достоверность.

В многоуровневых системах распознавания апостериорная информация о признаках определяется на основе косвенных измерений. Для таких измерений используются специализированные локальные системы распознавания. По данным технических средств  $T_1, \dots, T_r, \dots, T_n$  определяются признаки  $x_{11}, \dots, x_{k1}; x_{r1}, \dots, x_{lp}; x_{1n}, \dots, x_{mn}$  (первичные признаки), которые используются локальными системами распознавания для определения признаков более высокого уровня, которые, в свою очередь, используются в процессе распознавания неизвестных объектов (рис. 9.6).

Результаты полученные при исследовании объектов подвергаются статистическому анализу и фильтрации. Исключаются аномальные значения. Далее модель, полученная после фильтрации и статистического анализа в модуле распознавания нижнего уровня (Gauge system, рис. 9.7), представляется в виде карты высот. Карта высот подвергается нормализации и направляется в модуль распознавания верхнего уровня (Node array, рис. 9.7), где на основе данных от различных систем нижнего уровня происходит распознавание и моделирование свойств объекта. Результаты исследования, а так же диагностическая информация представляются пользователю посредством системы контроля и управления (Output/Control system, рис. 9.7).

## 9.6. Оценка достоверности

Оценка достоверности результатов моделирования может быть произведена различными методами, в частности, непараметрическим.

Непараметрический метод выявления отклонений заключается в ранжировании анализируемых значений  $X_1 \leq X_2 \leq \dots \leq X_n$  и вычислении статистики  $r$  для крайних значений. Формулы расчета зависят от числа анализируемых значений  $n$  и от того, значение с какого края проверяется на аномальность (самое большее или меньшее).

Таблица 1

### Формула для работы отклонений

Число значений	Проверяется максимальное значение $X_n$	Проверяется минимальное значение $X_1$
$3 \leq n \leq 7$	$r_{10} = (X_n - X_{n-1}) / (X_n - X_1)$	$(X_2 - X_1) / (X_n - X_1)$
$8 \leq n \leq 10$	$r_{11} = (X_n - X_{n-1}) / (X_n - X_2)$	$(X_2 - X_1) / (X_{n-1} - X_1)$
$11 \leq n \leq 13$	$r_{20} = (X_n - X_{n-2}) / (X_n - X_2)$	$(X_3 - X_1) / (X_{n-1} - X_1)$
$14 \leq n \leq 25$	$r_{21} = (X_n - X_{n-2}) / (X_n - X_3)$	$(X_3 - X_1) / (X_{n-2} - X_1)$

Полученное значение  $r$  сравнивается с критическим  $r$  и считается аномальным, если  $r > ar$

После этого процедура проверки повторяется для оставшихся  $n - 1$  значений.

Данный критерий оптимален для малых серий наблюдений (данных) и не зависит от числа имеющихся аномальных значений. В то время, как приведенный ранее параметрический критерий (особенно на для небольших серий) оптимален, когда имеется лишь одно аномальное значение.

### Заключение

Рассматриваемая математическая модель в совокупности с эволюционной стратегией позволяет оценивать критические ситуации для кластерных систем обработки информации и выявлять их последствия, а так же моделировать и оптимизировать адаптивными методами показатели качества подобных систем.

### Литература

1. Кравчук И.С., Тихоглаз Ю.С., Занг Н.Ч. Математическая модель и алгоритм управления качеством в кластерных системах сбора и обработки информации. – Системы управления и информационные технологии, 2008, 1.2(31), - с. 299-303.
2. Кравчук И.С., Тихоглаз Ю.С., Ву Тхи Тует Ланг. Эволюционная стратегия управления в задачах распознавания образов// Системы управления и информационные технологии, 2008, 2.3(32). – с.т 358-360.
3. Слободин М.Ю., Царев Р.Ю. Компьютерная поддержка многоатрибутивных методов выбора и принятия решения при проектировании корпоративных информационно-управляющих систем. – СПб.: Инфо-да, 2004. – 223 с.
4. Бусленко Н.П. Моделирование сложных систем. – М.: Наука, 1978.
5. Резников Б.А. Методы и алгоритмы оптимизации на дискретных моделях сложных систем. Л.: ВИКИ им. А.Ф. Можайского, 1983. – 215 с.



6. Schwefel H.-P. Evolution and Optimum Seeking. – N.Y.: Wiley Publ., 1995. – 612 pp.
7. Goldberg D.A. Genetic algorithm in search, optimization and machine learning. Addison-Wesley, Reading MA, 1989.
8. Вальков В.М., Вершин В.Е. Автоматизированные системы управления технологическими процессами. М., Машиностроение, 1977.
9. Кравченко В.А., Цидилин С.М., Федосеева Т.Л. Алгоритмы решения задач многокритериальной оптимизации: Учебное пособие. – М.: Издательство МИЭМ, 1988. – 74 с.
10. Рапопорт Э.Я. Анализ и синтез систем автоматического управления с распределенными параметрами: Учеб. Пособие / Э.Я. Рапопорт. – М.: Высш. шк., 2005. – 292 с.: ил.
11. Оптимальное управление. Сборник. – М.: «Знание», 1978. – 144 с. (Нар. Ун-т. Естественнонаучный фак. Издается с 1961 г.).
12. Производство тонкопленочных структур в электронном машиностроении: Учебник для вузов в 2-х томах. Т.2/ Н.В.Василенко, Е.Н. Ивашов, С.В.Степанчиков и др. – М.: Машиностроение, 2006. – 427с.

## **Глава 10. Оборудование молекулярно-лучевой эпитаксии для выращивания метастабильных наноструктур**

### **Введение**

Производство БИС и СБИС с повышенным быстродействием и высокой степенью интеграции требует разработки технологических процессов и оборудования, которые обеспечивали бы формирование микроструктур с размерами элементов до 1 мкм, толщиной эпитаксиальных слоев от 0,01 до 1 мкм и шириной межфазной границы примерно 1 нм, а также выращивания различных метастабильных структур и их модификаций. Улучшение вакуумных условий в технологическом оборудовании не только обуславливает высокую чистоту всех используемых материалов, но и открывает возможность контроля за ходом технологического процесса с помощью аналитических средств.

Под технологией молекулярно-лучевой эпитаксии (МЛЭ) подразумевают процесс выращивания в условиях сверхвысокого вакуума (при давлении  $10^{-8}$  Па и ниже) полупроводниковых структур путем взаимодействия одного или нескольких молекулярных (атомарных) пучков с монокристаллической поверхностью подложки, нагретой до определенной температуры, при контроле и управлении химическим составом и интенсивностью пучков, направленных на подложку.

### **10.1. Основные сведения о молекулярно-лучевой эпитаксии**

Молекулярно-лучевая эпитаксия по сравнению с парогазовой (ПГЭ) и жидкофазной (ЖФЭ) обладает рядом достоинств, основными из которых являются:

- возможность наращивания слоев атомарной толщины требуемого стехиометрического состава;
- легкость воспроизводства при относительно невысоких температурах совершенных по структуре и однородных по толщине слоев.
- универсальность, позволяющая наращивать слои как одного вещества, так и двух.

Кроме того, методом МЛЭ можно выполнять селективную эпитаксию, а также выращивать слои, содержащие гетеропереходы с сопряженными кристаллическими решетками или с постепенно изменяющимися периодами.

Выращиваемые слои контролируют высокочувствительной аналитической аппаратурой (оже-анализаторами, квадрупольными масс-спектрометрами, электронными дифрактометрами и др.), располагаемой в рабочей зоне. Применение молекулярно-лучевой эпитаксии в технологическом процессе существенно повышает выход годных структур с электронно-дырочными, гомо- и гетеропереходами.

Низкая температура и малая скорость роста слоев при молекулярно-лучевой эпитаксии требуют особого внимания к подготовке поверхностей подложек. Обычно подложки после химического или многократного ионного травления отжигают в сверхвысоком вакууме. Высококачественные слои арсенида галлия получают после отжига подложек при 570 °С в течение 30 мин в атмосфере паров мышьяка. При такой обработке удаляются не только различные примеси, но и восстанавливаются нарушения кристаллической решетки подложек.

Качество наращиваемых слоев существенно зависит от соотношения компонентов в молекулярных потоках. При нанесении слоев бинарных соединений из разных источников не удастся получить заданный состав пленки.

Хорошие результаты получают при использовании общего источника, содержащего смесь двух компонентов, например галлия и индия. При температуре источника 800 °С скорость роста пленки составляет 1000 нм/ч. При этом скорость испарения индия в 10 раз выше, чем галлия, а соотношение их масс примерно равно 10, что способствует созданию равных потоков, необходимых для нормального роста слоя.

Так как с ростом температуры давление паров галлия изменяется быстрее давления паров индия, то при увеличении температуры общего источника возрастает и соотношение масс галлия и индия в потоках. Например, при повышении температуры на 1 °С соотношение масс веществ в потоках из общего источника возрастает лишь на 0,5 %, в то время как при раздельном испарении оно увеличивается до 3 %.

В высоком и сверхвысоком вакууме молекулярно-лучевой эпитаксией можно осаждать полупроводниковые, диэлектрические металлические слои.

Молекулярно-лучевой эпитаксией осаждают проводящие слои алюминия, а также изолирующие слои его оксида испарением алюминия в среде кислорода на поверхности подложек из арсенида галлия.

Как уже отмечалось, достоинством молекулярно-лучевой эпитаксии является возможность анализа и контроля с помощью специальных приборов и методов параметров слоев в процессе наращивания.

О Ж Э - спектрометр — прибор, работа которого основана на ОЖЭ эффекте, названном по имени французского ученого П.В. Оже, открывшего его в 1925 году. Этот прибор применяют для контроля наличия примеси и неоднородности состава очень тонких осажденных поверхностных и приповерхностных слоев. Причем с его помощью по изменению энергии электронной оболочки электронов обнаруживают в осажденных слоях любой химический элемент, кроме водорода и гелия. Исследуемые подложки необходимо предварительно тщательно очищать, так как только при этом условии кинетическая энергия электронов будет достаточна для их выхода с поверхностных или из приповерхностных слоев.

Квадрупольный масс-спектрометр служит для контроля спектра масс (до 200 а.е.м.) газовых и молекулярных потоков, а также остаточной среды.

Принцип действия этого прибора основан на разделении ионов по удельным зарядам при их движении в квадрупольных (прямоугольных, квадратных) высокочастотных полях, создаваемых специальным датчиком.

Методы дифракции быстрых и медленных электронов позволяют выполнять соответственно кристаллографический анализ подложек и наращенных слоев.

Применяют также другие методы контроля параметров слоев. Так, концентрацию примесей в слоях контролируют вторично-ионным масс-спектрометрическим методом (ВИМС), а их химический состав - методом рентгено-электронной спектроскопии.

## **10.2. Сверхвысоковакуумные установки и комплексы для молекулярно-лучевой эпитаксии**

Оборудование для молекулярно-лучевой эпитаксии разделяют на одно- и многокамерное.

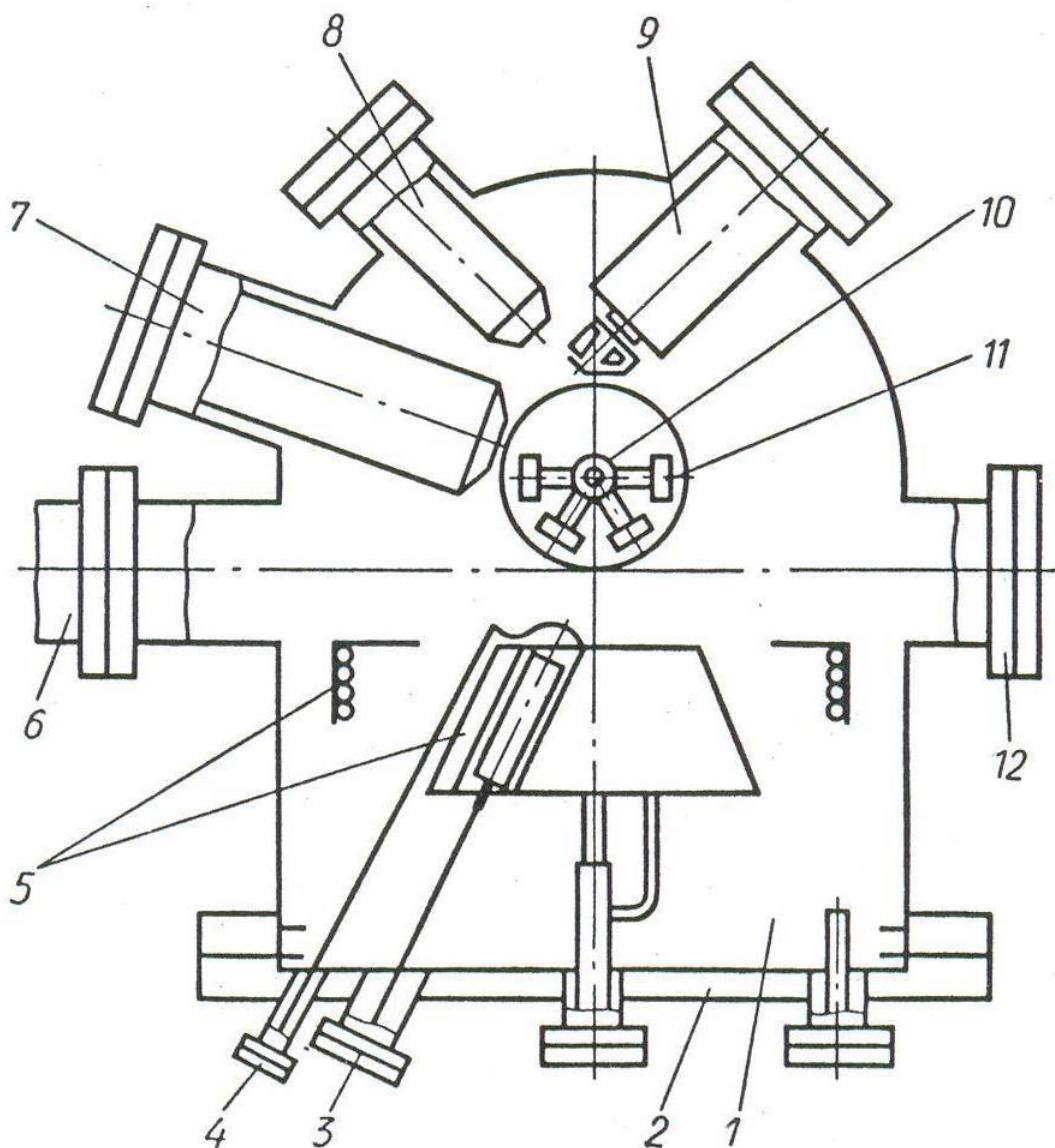
Однокамерные установки (рис.10.1) состоят из вакуумной рабочей камеры со средствами откачки, блока молекулярных источников, устройства нагрева подложек, ионной пушки для подготовки поверхности подложек, оже-спектрометра для анализа структуры поверхности подложек и выращенной эпитаксиальной пленки, а также устройств загрузки подложек в рабочую камеру, юстировки ее положения и разгрузки.

Процесс наращивания эпитаксиального слоя осуществляется в определенной последовательности. Предварительно перед загрузкой на подложкодержателе закрепляют расплавом индия, галлия, висмута или механическими прижимами подложку. Затем при атмосферном давлении в рабочей камере подложкодержатель устанавливают на держатель прецизионного манипулятора. Создание сверхвысокого вакуума и обезгаживание всей системы обычно длится 20 - 48 часов при температуре 250 – 400<sup>0</sup> С.

При комбинированной тепловой и ионной обработке поверхности подложек, одновременно методами дифракции быстрых электронов и оже-спектрометрии контролируют их структуру и химический состав на атомарном уровне.

Тепловую обработку подложек выполняют косвенным нагревом, а ионную- с помощью ионной пушки. Для осаждения эпитаксиального слоя молекулярный пучок вещества направляют на подложку, нагретую до определенной температуры. Одновременно квадрупольным масс-спектрометром контролируют плотность и химический состав молекулярного пучка, а методом дифракции быстрых электронов с оперативной оценкой сглаживания поверхности подложки в начале наращивания – кристаллогеометрию растущего слоя. После окончания процесса эпитаксии на участке подложек площадью 1 - 2 см<sup>2</sup> контролируют методом вторично-

ионного масс-спектрометрического анализа распределения примесных атомов.



**Рис. 10.1.** Схема сверхвысоковакуумной однокамерной установки для молекулярно-лучевой эпитаксии:

1 - рабочая камера; 2 - блок молекулярных источников; 3 - молекулярный источник; 4 - привод заслонки; 5 - криопанель блока молекулярных источников; 6 - пушка для дифракции быстрых электронов; 7 - оже-спектрометр; 8 - ионная пушка для очистки поверхности подложек; 9 - квадрупольный масс-спектрометр; 10 - держатель подложек прецизионного манипулятора; 11 - подложка; 12 - флуоресцентный экран для дифракции быстрых электронов

После разгерметизации рабочей камеры и выдвижения манипулятора снимают подложкодержатель с подложкой.

Многокамерные установки для молекулярно-лучевой эпитаксии являются более современным оборудованием для выращивания и исследования тонкопленочных структур.

К преимуществам многокамерных установок относятся:

- высокая производительность технологического процесса эпитаксиального роста за счет установившихся вакуумных условий в камере роста и проведения операций предварительной и финишной очистки подложки в других камерах;
- стабильность и надежность работы аналитических приборов за счет исключения взаимодействия чувствительных поверхностей аналитической аппаратуры с материалами, испаряемыми молекулярными источниками во время проведения процесса эпитаксии;
- возможность использования различных физико-химических воздействий для подготовки поверхностей подложек;
- увеличение производительности оборудования за счет одновременного производства различных операций в отдельных камерах.

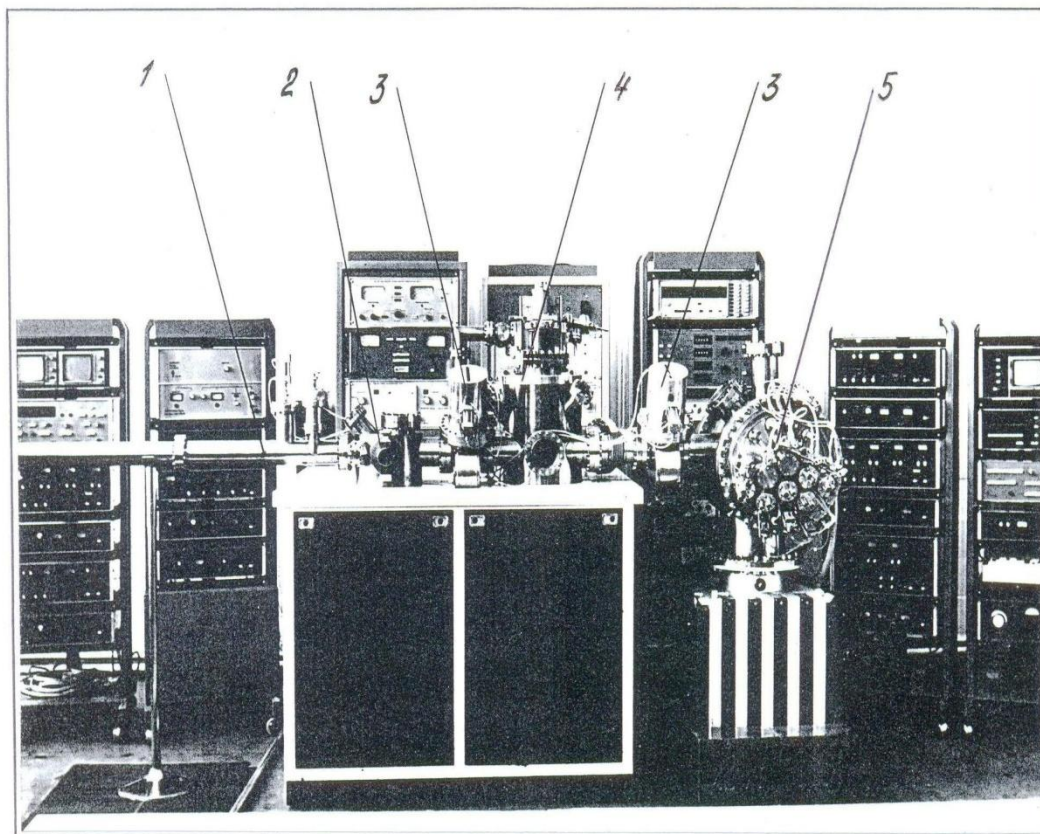
Известные многокамерные установки для молекулярно-лучевой эпитаксии отличаются друг от друга по компоновке, составу и конструкции отдельных устройств.

На рис. 10.2 представлен внешний вид и структурно-компоновочная схема многокамерной установки для молекулярно-лучевой эпитаксии с системой транспортирования через вакуумные камеры. Установка выполнена в виде двух модулей: аналитического и технологического, представляющих собой функционально законченные единицы, имеющие полный комплект устройств и средств для выполнения операций.

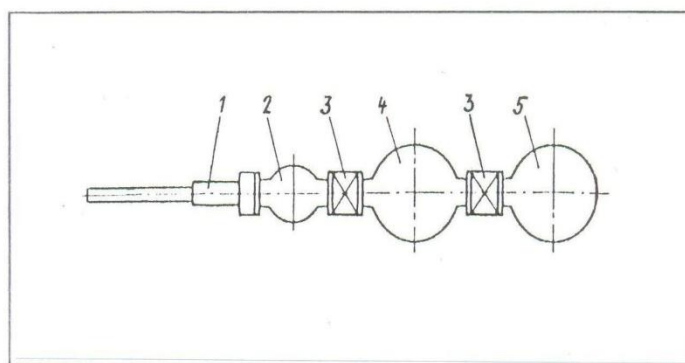
Вакуумные объемы модулей соединяются между собой с помощью сильфонного переходника и могут изолироваться вакуумным шибером.

В состав аналитического модуля входят вакуумная камера с системой предварительной и сверхвысоковакуумной откачки, прецизионный манипулятор, камера загрузки с передающим манипулятором и вакуумный шибер. Вакуумная камера аналитического модуля содержит ионную пушку для очистки подложек, прецизионный манипулятор и оснащена различными методами анализа поверхности.

Камера загрузки имеет люк для загрузки подложек, смотровое окно и несколько фланцев для размещения различных устройств (электронной пушки для обезгаживания подложек, нагревательного устройства, клапана для обезгаживания подложек, нагревательного устройства, клапана для напуска газа и др.).



а)



б)

**Рис. 10.2.** *Сверхвысоковакуумная многокамерная установка для молекулярно-лучевой эпитаксии ( а – внешний вид; б – структурно-компоновочная схема): 1 – передающий манипулятор; 2 – камера загрузки; 3 - шиберные затворы; 4 - аналитический модуль; 5 - технологический модуль*



Передающий манипулятор обеспечивает перемещение и передачу подложек из камеры загрузки в аналитическую и ростовую камеры и обратно.

Сверхвысоковакуумная система аналитического модуля состоит из сублимационного титанового насоса с азотной криопанелью и магниторазрядного насоса. Эта система обеспечивает предельный вакуум порядка  $1 \cdot 10^{-8}$  Па и скорость откачки около 2000 л/с.

Система предварительной откачки включает два адсорбционных цеолитовых насоса и механический насос с сорбционной ловушкой.

Для откачки камеры загрузки до вакуума порядка  $1 \cdot 10^{-5}$  Па применяются магниторазрядный насос, диффузионный насос с азотной ловушкой и механический насос.

Аналитический модуль обеспечивает выполнение следующих операций:

- загрузку и выгрузку подложек, перемещение и передачу их в камеру анализа;
- обезгаживание подложек;
- контроль химического состава тонкопленочных структур и поверхности подложек методом электронной оже-спектроскопии;
- исследование структур методом дифракции медленных электронов;
- исследование структур и поверхности подложек вторично-ионным масс-спектрометрическим методом;
- юстировку положения подложки с помощью сверхвысоковакуумного манипулятора.

В состав технологического модуля входят: технологическая камера со сверхвысоковакуумной системой откачки, блок молекулярных источников, прецизионный манипулятор с держателем и устройством нагрева, вакуумный щитер, а также системы анализа на основе квадрупольного масс-спектрометра и дифрактометра быстрых электронов.

Технологическая камера расположена горизонтально. На одном из фланцев установлен блок молекулярных источников, на другом фланце – сублимационный насос, в центре которого установлен прецизионный манипулятор. Снизу камеры присоединен магниторазрядный насос НМД – 1,0. Фланцы на цилиндрической поверхности камеры служат для монтажа приборов анализа и вакуумного щитера. Предусмотрена возможность установки передающего манипулятора.

Сверхвысоковакуумная система состоит из сублимационного насоса с азотной криопанелью и магниторазрядного насоса НМД – 1,0 и обеспечивает предельный вакуум  $1 \cdot 10^{-8}$  Па и скорость откачки около 10000 л/с.

Предварительная откачка технологической камеры производится за счет вакуумной системы аналитического модуля.

Горизонтальный блок молекулярных источников состоит из восьми источников, расположенных в несколько рядов под углом к горизонту в пределах от  $8^{\circ}$  до  $24^{\circ}$ .

Технологический модуль обеспечивает:



- управление процессом эпитаксиального роста полупроводниковой структуры;
- контроль структуры подложек и пленок в процессе роста методом дифракции быстрых электронов;
- контроль химического состава пучков и спектра остаточных газов в камере методом квадрупольной масс-спектрометрии;
- температуру нагрева молекулярных источников до  $1250^{\circ}\text{C}$  с точностью регулирования  $\pm 0,5^{\circ}\text{C}$ .

Также разработана модификация установки для молекулярно-лучевой эпитаксии с двумя передающими манипуляторами. При такой компоновке второй передающий манипулятор устанавливается на технологическую камеру, что обеспечивает уменьшение длины хода первого манипулятора, проведение операций одновременно в обеих камерах и сокращение потерь времени на загрузку и выгрузку подложек.

Стремление к повышению производительности оборудования путем совмещения технологических операций во времени и повышению надежности, благодаря упрощению маршрутизации, привело к созданию сверхвысоковакуумных технологических комплексов.

Один из вариантов комплекса – магистрально-модульный комплекс для молекулярно-лучевой эпитаксии представлен на рис. 10.3.

Комплекс предназначен для получения в условиях сверхвысокого вакуума методами молекулярно-лучевой эпитаксии, электронно-лучевого испарения и ионно-лучевого осаждения:

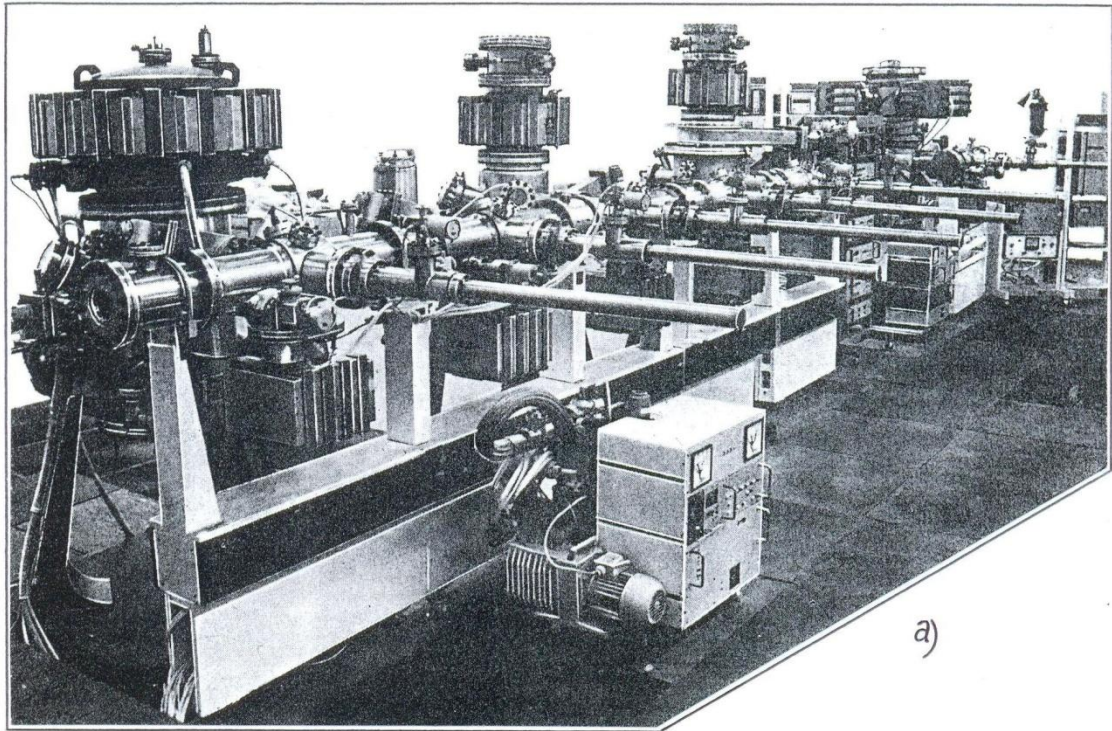
- полупроводниковых структур (Si, Ge соединений  $A^{III}B^V$ ,  $A^{II}B^{VI}$ ,  $A^{IV}B^{VI}$ );
- металлических и диэлектрических покрытий, магнитных слоев;
- пленок субмикронных толщин с любым профилем легирования, а также многослойных периодических структур (сверхрешеток), которые не могут быть синтезированы традиционными технологиями.

В состав комплекса входят модули:

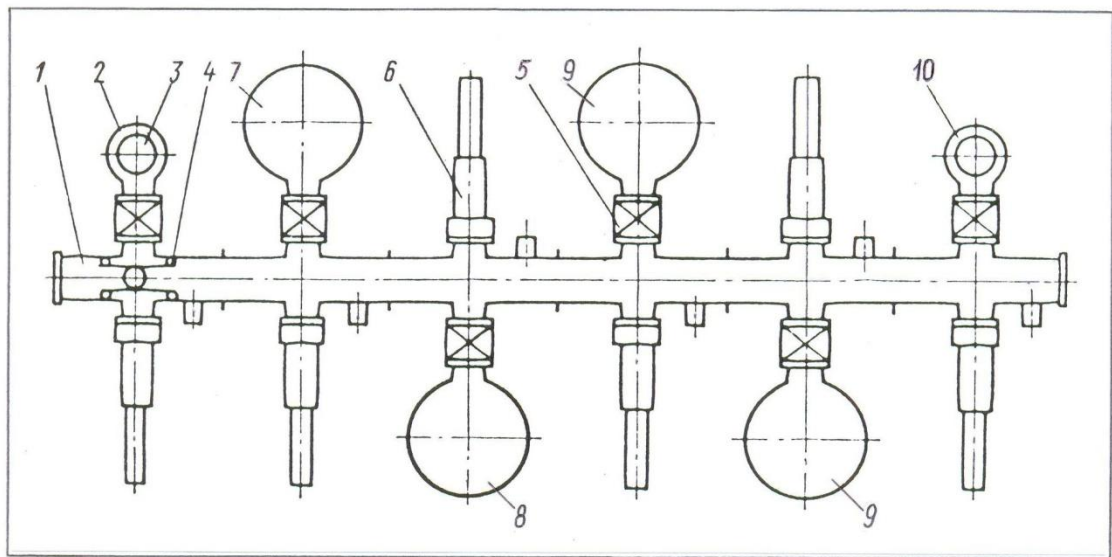
- подготовки подложек;
- молекулярно-лучевой эпитаксии;
- получения диэлектрических покрытий и металлизации.

Модули объединены в единый комплекс загрузочно-транспортной системой с выносным магистральным транспортером, позволяющей производить загрузку, выгрузку и транспортирование подложек и готовых структур между модулями.

Каждый технологический модуль соединяется с магистральным транспортером посредством шиберного затвора, что позволяет полностью разделить вакуумные рабочие объемы в процессе транспортирования и передачи подложек, проводить операции их загрузки и выгрузки параллельно с выполнением основных операций.



a)



б)

**Рис.10.3.** *Сверхвысоковакуумный магистрально-модульный комплекс для молекулярно-лучевой эпитаксии: а - внешний вид; б - структурно – компоновочная схема: 1 - магистральный транспортер; 2 - модуль загрузки; 3 - кассета с подложками; 4 - каретка транспортера; 5 - шиберные затворы; 6 – передающие манипуляторы; 7 - модуль подготовки подложек; 8 - модуль молекулярно-лучевой эпитаксии; 9 – модуль получения диэлектрических покрытий и металлизации; 10 - модуль выгрузки*

Магистральный транспортер выполнен в виде секций, герметично соединенных между собой и образующих вакуумный канал параллельно технологическим камерам, расположенными в линию. Внутри канала по направляющим перемещается каретка, несущая одну или две подложки. Передача подложек с каретки транспортера в технологический модуль, в кассету-накопитель или обратно осуществляется с помощью передающих манипуляторов, установленных на каждой транспортной секции. На крайних (загрузочных) секциях установлены устройства для кассетной загрузки выгрузки подложек (по 10 подложек в кассете).

С целью снижения загрязнений, привносимых на подложки в вакуумно-транспортных системах оборудования, последние могут быть установлены в пылезащитных подложкодержателях (рис. 10.4) которые открываются только на позициях обработки и контроля.

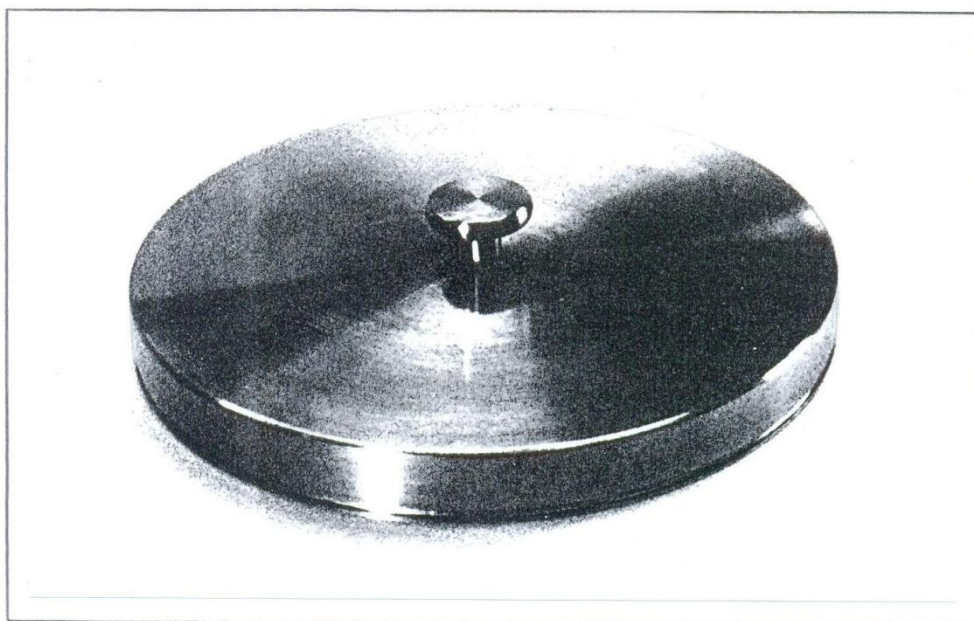
Вакуумный транспортер снабжен агрегатами вакуумной откачки на позициях загрузки и выгрузки и магниторазрядными вакуумными насосами, расположенными на каждой секции.

Магистрально-модульный принцип конструкции комплекса обеспечивает возможность наращивания его модулями другого функционального назначения, реализовать различные варианты компоновки модулей в зависимости от поставленных задач.

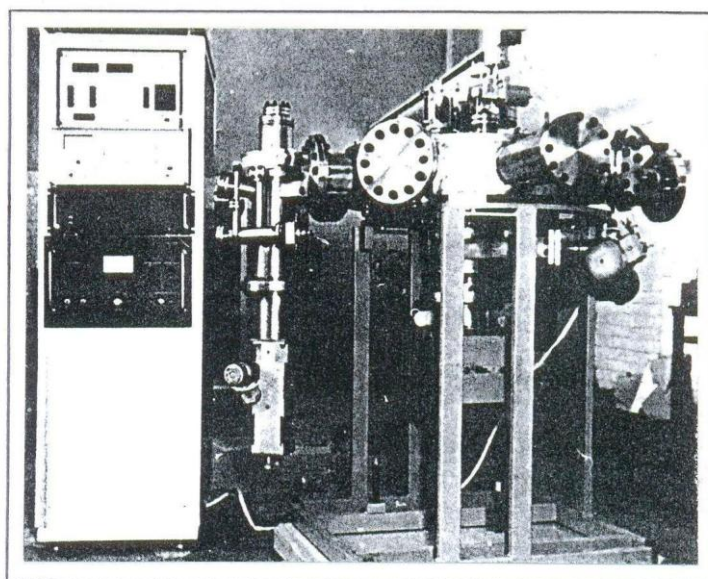
Помимо разработанного варианта магистрально-модульного комплекса для молекулярно-лучевой эпитаксии, существует вариант комплекса, обслуживаемого вакуумным роботом (рис. 10.5). При небольшом числе операций сверхвысоковакуумный роботизированный комплекс молекулярно-лучевой эпитаксии выполнен с центральной камерой, где функционирует один робот, выполняющий все необходимые транспортные и загрузочные перемещения кассет с подложками, включая обслуживание загрузочной и разгрузочной камер (рис. 10.6).

## Литература

- 1.Моряков О.С. Эпионная обработка. М.: Высш. школа, 1990. 128 с.
- 2.Панфилов Ю.В., Рябов В.Т., Цветков Ю.Б. Оборудование производства интегральных микросхем и промышленные роботы. М.: Радио и связь, 1988. 320 с.
- 3.Волчкевич Л.И. Автоматизация производства электронной техники. М.: Высшая школа, 1988. 287 с.
- 4.Ивашов Е.Н., Степанчиков С.В. Метрологическая навигационная система вакуумного оборудования кластерного типа: Материалы НТК «Датчик – 1994». Гурзуф, 1994. 4 с.
- 5.Ивашов Е.Н., Степанчиков С.В. Сверхвысоковакуумное оборудование для молекулярно-лучевой эпитаксии. М.: МИЭМ, 2006, 18с.

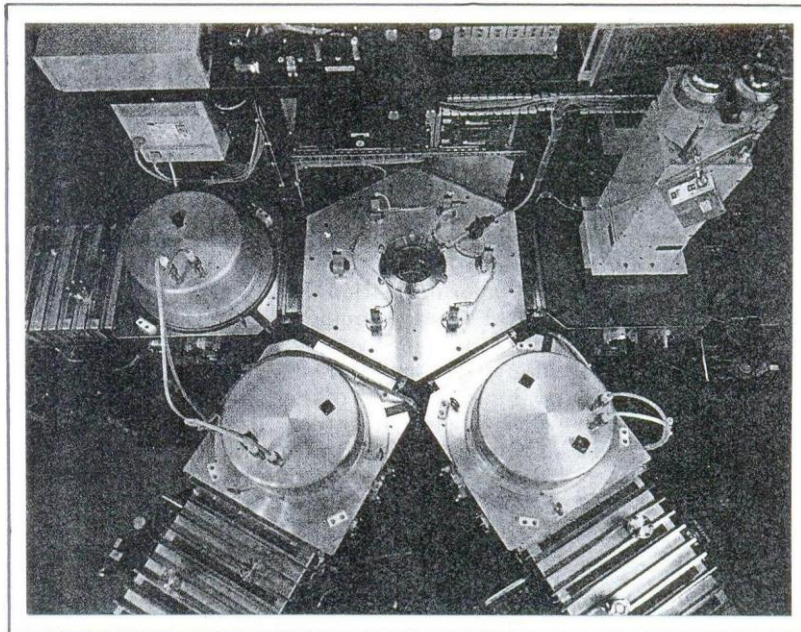


*Рис. 10.4. Пылезащитный подложкодержатель*

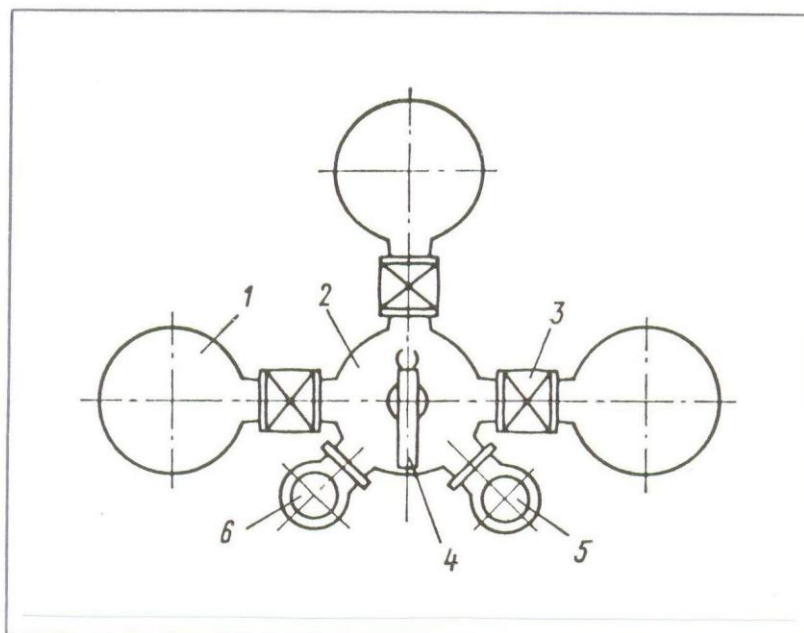


*Рис. 10.5. Вакуумный робот;  
Управление процессорное;  
Предельное остаточное давление в камере робота –  $1,33 \cdot 10^{-7}$  Па;  
Количество обслуживаемых технологических камер – 3;  
Количество степеней подвижности робота – 3;  
Перемещение исполнительного органа робота – 750 мм;  
Количество кассет – 2;  
Диаметр транспортного канала – 125 мм*





a)



б)

**Рис. 10.6.** *Сверхвысоковакуумный роботизированный комплекс для молекулярно-лучевой эпитаксии (а – внешний вид ; б – структурно-компоновочная схема):*

*1 – технологические камеры; 2 – центральная транспортная камера; 3 – шиберный затвор; 4 – вакуумный робот; 5 – шлюзовое загрузочное устройство; 6 – шлюзовое разгрузочное устройство*



**ВАСИН Владимир Анатольевич**

**ИВАШОВ Евгений Николаевич**

**СТЕПАНЧИКОВ Сергей Валентинович**

## **НАНОТЕХНОЛОГИЧЕСКИЕ ПРОЦЕССЫ И ОБОРУДОВАНИЕ ЭЛЕКТРОННОЙ ТЕХНИКИ**

**Редактор С. П. Клышинская**

**Технический редактор О. Г. Завьялова**

Подписано в печать 23.10.2009. Формат 60x84/16

Бумага типографская. Печать – ризография.

Усл. печ. Л. 16,5. Уч.-изд. л. 14,8. Тираж 300 экз.

Заказ 221. Изд. № 81

Московский государственный институт электроники и математики  
109028 Москва, Б. Трехсвятительский пер., 3/12.

Отдел оперативной полиграфии Московского государственного  
института электроники и математики.

113054 Москва, ул. м. Пионерская, 12.